Information Processing and Management 63 (2026) 104438

Contents lists available at ScienceDirect

Information Processing and Management

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

Unraveling urban problem patterns in Zhejiang: Evolutionary trajectories and driving factors based on public complaint data

Xiangfu Kong [*], Bo Dong

Zhejiang Philosophy and Social Sciences Pilot Laboratory, Zhejiang Lab, Hangzhou, Zhejiang 311100, China

ARTICLE INFO

Keywords:
Urban problems
Urban governance
Public complaint data
Evolutionary trajectories
Influencing factors

ABSTRACT

This study systematically investigates the formation, evolution, and drivers of urban problem patterns across 87 county-level administrative units in Zhejiang Province, China. By classifying approximately 3.5 million public complaints into 37 distinct categories, urban problem distributions in these counties are quantified from 2018 to 2022. The cluster analysis reveals seven distinct problem patterns: pollution and municipal infrastructure, consumer protection, land use, housing purchase, garbage management, financial service, and pandemic control, which exhibit marked temporal persistence and spatial clustering. Using fuzzy logic to calculate membership degrees, the evolution of problem patterns exhibits predictable trajectories: emerging patterns showed consistent growth, while dominant patterns gradually declined, with occasional deviations. Analysis of eight socioeconomic factors indicates that development-driven patterns—land use, housing purchase, and consumer protection—are strongly correlated with socioeconomic indicators such as land acquisition area, property sales volume, and retail activity, necessitating enhanced regulatory oversight of economic activities. Trend-driven patterns, exemplified by pollution and municipal infrastructure problems, correlate less with short-term fluctuations and more with stable socioeconomic conditions, underscoring the importance of systematic planning and coordinated interventions to mitigate structural regional imbalances. Event-driven patterns, including garbage management, financial service, and pandemic control, demonstrate minimal association with socioeconomic factors, indicating their susceptibility to notable events and highlighting the need for rapid crisis management.

1. Introduction

Urban problems are deeply embedded in residents' daily lives, yet they often receive insufficient attention from policymakers, researchers, and the public. Issues such as environmental pollution, illegal land use, and selective law enforcement, while persistent and impactful, rarely escalate into crises that directly threaten social stability (Wang et al., 2025; Sun, 2015). Consequently, most residents tolerate or adapt to these issues, leaving them unaddressed. However, scholars have provided rich empirical evidence on the gradual deterioration of both physical and mental health caused by urban problems, ultimately undermining overall well-being (Nobile et al., 2023; Xu et al., 2018; Zhang et al., 2023). For example, a 10 dB(A) increase in road traffic noise raises the relative risk of ischemic heart disease by 8_% (Münzel et al., 2021).

In China, economic development often takes precedence over people's well-being, leading to a range of quality-of-life problems,

* Corresponding author at: Building 8B, Zhejiang Lab, Yuhang Subdistrict, Hangzhou, Zhejiang 311100, China.
E-mail addresses: kongxf@zhejianglab.edu.cn (X. Kong), dongb@zhejianglab.edu.cn (B. Dong).

https://doi.org/10.1016/j.ipm.2025.104438
Received 27 March 2025; Received in revised form 5 October 2025; Accepted 5 October 2025
Available online 15 October 2025
0306-4573/

such as food safety concerns, environmental degradation, and contentious land expropriation (Pei et al., 2011; Sha, 2023; Zhao et al., 2022). These problems increasingly challenge central and local governance (Cui et al., 2015; Lin et al., 2019; Wu et al., 2017). For example, rapid urbanization has driven large-scale land expropriation, affecting nearly 19 % of Chinese villages from 2012 to mid-2016 (Zhao et al., 2022), a practice linked to diminished political trust and heightened citizen-authority tensions (Sha, 2023; Zhao & Xie, 2022). Acknowledging the threat these problems pose to social stability, the Chinese government has deployed top-down measures such as campaign-style enforcement (Wang et al., 2022), zero-tolerance policies (Zhai, 2023), large-scale surveillance networks (Xie & Yuan, 2023), and social credit system (Zhao & Liu, 2024). While these strategies are effective in rapidly mobilizing resources to achieve short-term governance targets, they are primarily unsustainable, resource-intensive solutions that fail to address the root causes of urban problems.

Effectively addressing urban problems requires not only substantial governance resource investments but also a systemic comprehension of how their distributions evolve with socioeconomic development (Engin et al., 2020; Xie et al., 2023). Critical to this effort are two questions: what patterns characterize the evolution of a region's urban problem distribution, and which factors drive this evolution? Answering these questions enables authorities to differentiate between the problems that may be resolved organically via socioeconomic development and those requiring targeted governance investment. Such distinctions help policymakers design context-appropriate governance strategies.

Despite its significance, systematic analysis of urban problem distributions, including their typical patterns, evolutionary trajectories, and underlying drivers, remains limited. A major barrier is restricted data accessibility, which hinders the quantification of urban problem distributions. Existing studies typically address single problems rather than multiple coexisting problems, leaving many underexplored (Jiao et al., 2024). This fragmented perspective also constrains systemic analysis and obscures interrelationships among problems. While individual problems may appear to fluctuate randomly, their collective dynamics likely follow discernible trends shaped by socioeconomic or institutional forces.

This study bridges existing research gaps by leveraging public complaint data to systematically analyze urban problems. It pursues two main objectives: methodologically, to demonstrate the value of complaint data in urban studies; and theoretically, to reveal that the distribution and evolution of urban problems follow identifiable patterns influenced by exogenous factors. Accordingly, the study proposes three research questions:

RQ1. What patterns characterize the urban problem distributions across counties and years?
RQ2. How do these problem patterns evolve over time?
RQ3. What socioeconomic factors drive the evolution of these patterns?

To answer these questions, we analyzed 3.5 million complaint records from 87 counties in Zhejiang spanning 2018–2022. Using text classification, complaints were categorized into 37 problem types, yielding 5 × 87 year-county urban problem distributions. Cluster ensemble methods identified seven typical patterns from the distributions. Each pattern was characterized by multiple representative problems that are interrelated. Fuzzy logic was then applied to estimate each distribution's degree of membership across the patterns, capturing the evolutionary trajectories. A key finding is that emerging patterns exhibited steady growth in memberships, whereas dominant patterns gradually declined. To identify the drivers of these evolutions, we employed multinomial logit regression to model the relationships between socioeconomic factors and problem patterns. Results show that urban problem patterns can be broadly classified as development-driven, trend-driven, or event-driven, highlighting the heterogeneous forces shaping urban problems.

2. Backgrounds

The concept of *urban problems* remains inconsistently defined across scholarly works. To anchor this study, we first delineate the scope of urban problems before conducting a literature review. First, urban problems encompass problems not confined to urban cores but also prevalent in rural regions, such as land degradation and unclear land titling. Second, while scholars often frame macro-level issues, like urbanization-driven carbon emissions and housing affordability, as urban problems, this study prioritizes tangible problems that directly affect personal well-being and daily activities. Third, we exclude personal disputes, commercial conflicts, and ideological debates (e.g., social justice) as these typically fall outside administrative governance mandates. In summary, *urban problems* are defined here as non-urgent public affair problems that directly and tangibly impact residents' daily lives and necessitate government-led interventions.

Academic research on urban problems has traditionally emphasized isolated problems rather than systemic analysis, reflecting a fragmented perspective (Jiao et al., 2024; Liu et al., 2024). This limitation is largely due to the restricted data availability. Although government agencies collect extensive data through patrols, inspections, and monitoring systems, public access is limited. Only select departments—such as the Market Regulation Bureau and the Housing and Urban-Rural Development Bureau—occasionally release partial data on issues like product quality violations or unauthorized land use. Moreover, China's urban governance framework is institutionally fragmented, with responsibility for problem resolution dispersed across multiple bureaucratic departments (Peng et al., 2022), hindering systematic data aggregation and cross-sector analysis. These barriers ultimately constrain holistic understanding of urban problems and their interdependencies.

Recently, public complaint data has garnered increasing scholarly and policy interest for its capacity to provide fine-grained, real-time insights into urban problems through direct resident feedback. Early research emphasized optimizing complaint-handling mechanisms (Brewer, 2007; Gelbrich & Roschk, 2011; HaCohen-Kerner et al., 2019; Orsingher et al., 2010) and assessing governmental responsiveness (Chatfield & Reddick, 2018; Pan & Chen, 2018). Recent advances in Natural Language Processing (NLP) have enabled sophisticated extraction of user-defined insights from unstructured complaint narratives via text classification (Zhang et al., 2025; Madyatmadja et al., 2023; Peng et al., 2022), topic modeling (Bastani et al., 2019; Ma et al., 2016; Zhang et al., 2023), and sentiment analysis (Feng et al., 2023; Guo et al., 2023; Jiao et al., 2024; Liu et al., 2024). However, these studies remain largely methodological, prioritizing technical innovation, such as model accuracy and computational efficiency, over advancing urban governance theory.

Despite increasing use of complaint data in digital governance practices, systematic theoretical research on the mechanisms underlying the emergence, evolution, and resolution of urban problems remains critically underexplored. Existing spatial-temporal analyses often remain superficial, emphasizing descriptive mappings of *where* and *when* problems cluster rather than interrogating *how* and *why* they persist or transform (He et al., 2024; Sun et al., 2021; Wang et al., 2022). Furthermore, studies linking urban development to urban problems—such as Jiao et al. (2024), who correlated land use types with urban problems in Sanya City, and Peng et al. (2024), who traced facility distribution impacts on problem evolution in Chengdu—are constrained by a single-city perspective. Such localized approaches limit the identification of cross-regional patterns and hinder the development of generalizable insights applicable across diverse demographic, geographic, and infrastructural contexts.

3. Materials and methods

3.1. Study area

Zhejiang Province, a highly urbanized and economically vibrant region in eastern China, exemplifies the dual-edged nature of rapid economic development. While its urbanization and industrialization have significantly elevated living standards, they have also engendered complex urban challenges including environmental degradation, consumer protection violations, and land disputes. These problems are compounded by pronounced intra-provincial disparities: core cities such as Hangzhou and Ningbo contrast sharply with less developed peripheral cities. Zhejiang's large-scale 12,345-hotline dataset provides a unique empirical foundation to quantify urban problem distributions, model their spatial-temporal evolution, and identify key socioeconomic drivers. This makes Zhejiang Province a representative and valuable case for advancing theoretical studies on urban problems and urban governance.

Zhejiang's administrative division comprises 11 prefecture-level cities subdivided into 90 county-level units,[1] including 37 urban districts, 33 counties, and 20 county-level cities (Fig. 1). Districts represent urban cores with dense populations and extensive infrastructure, while counties manage rural or suburban areas, overseeing local administration. County-level cities, similar in governance to counties, exhibit higher urbanization, economic output, and administrative autonomy. For simplicity, we collectively referred to districts, counties, and county-level cities as "*counties*," unless otherwise specified. County-level administrative regions were selected as the basic areal unit to balance the trade-off between sample size and reliability: prefecture-level units offer reliability but a limited sample size ($n = 11$), whereas town-level units ($n = 1300$) increase sample size but reduce reliability.

3.2. Data sources

The Zhejiang 12,345-hotline platform (see https://zxts.zjzwfw.gov.cn) began publicly disclosing complaint data in 2017. We collected five years of public complaint data from this platform, spanning from January 1, 2018, to December 31, 2022. After deduplicating entries with identical narratives, the final dataset comprised nearly 3.5 million unique records. Each record includes metadata such as the order number, date, county-level location, responsible department, title, and an anonymized complaint narrative, with all personal identifiers removed prior to public release (Table 1).

Due to data security protocols, the dataset disclosed less than 10 % of the total complaints, consisting of user-authorized entries and randomized samples released by the petition departments. To evaluate the representativeness of our dataset, we compared it with official hotline reports from local petition departments. Cities like Hangzhou and Ningbo occasionally released aggregate complaint counts or proportions of common urban problems, which served as valuable references. These reports typically employed coarse-grained categories and reported data at the city rather than county level. To ensure comparability, we reclassified our problem categories (see Table A.1 in Appendix A) to match the official taxonomy and aggregated our data at the city level. Fig. 2 compares urban problem distributions derived from our dataset with those reported by four petition departments. The distributions show only minor differences, indicating that our sample is broadly representative of the actual urban problem distributions.

3.3. Methods

First, unstructured complaint narratives were transformed into structured labels by text classification, with each label corresponding to a distinct urban problem. Complaint volumes were then aggregated by county, year, and label, generating annual county-level urban problem distributions. Next, a cluster ensemble method partitioned these distributions into clusters, with the centroids representing the typical patterns of urban problem distributions. This clustering facilitated the analysis of problem coexistence and

[1] Between 2018 and 2022, three county-level regions, including Linping District, Qiantang District, and Longgang City, underwent administrative restructuring. As these counties were newly established during the study period, their longitudinal data were incomplete and were excluded from the analysis to ensure temporal consistency and reliability.

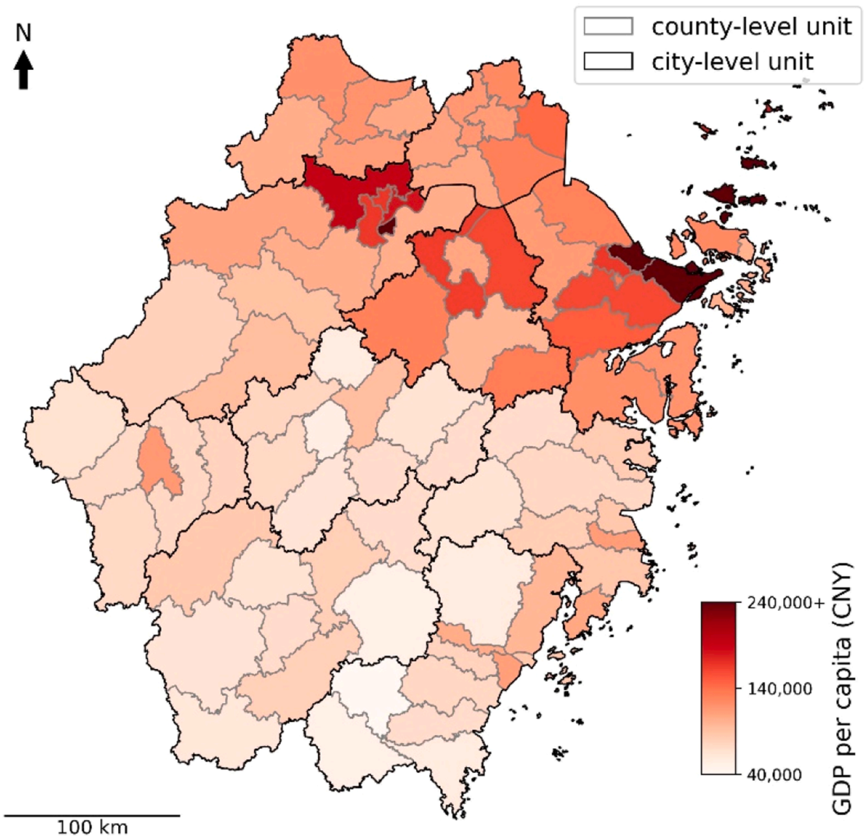


Fig. 1. Administrative division of Zhejiang Province.

Table 1
A typical example of a complaint record.

Order number	1557756880012546049
Date	2019–07–25 16:38:28
County	Wuxing District
Responsible department	Urban Management Bureau of Binhu Subdistrict
Title	莫太路路边摆摊没人管Unregulated Street Vendors on Motai Road
Complaint narrative	莫太路路边早上有人摆摊卖早点, 无证经营 In the morning, unlicensed vendors operate stalls selling breakfast along Motai Road

spatiotemporal dynamics. A fuzzy logic method then quantified the membership of each distribution relative to the identified patterns, enabling the tracking of evolutionary trajectories in these patterns. Temporal variations in memberships were visualized using stacked bar plots, and key evolutionary trajectories were manually summarized. Finally, to assess the factors influencing pattern evolutions, two multinomial logistic models were employed to analyze both the correlative and causal relationships between socioeconomic factors and pattern memberships. Comparative results were visualized using Directed Acyclic Graphs (DAGs). The full methodological workflow is illustrated in Fig. 3.

3.3.1. Urban problem distribution

The original complaint data for Zhejiang Province lacked publicly available labels categorizing urban problems.[2] To address this, we transformed the unstructured narratives into labeled datasets through single-label text classification.[3] Text classification is a common task in NLP, with many well-established deep learning classifiers available. First, we iteratively developed a classification

[2] Although the petition bureau manually tagged complaints, these labels were not accessible on the complaint platform in Zhejiang.

[3] Classifying resident complaints is fundamentally a multi-label classification task, as some complaints may involve multiple concurrent problems. However, empirical annotation showed that fewer than 0.5% of annotated complaints exhibit multi-label characteristics. Therefore, simplifying the task to single-label classification introduces negligible distortion to analytical outcomes while significantly reducing computational complexity.

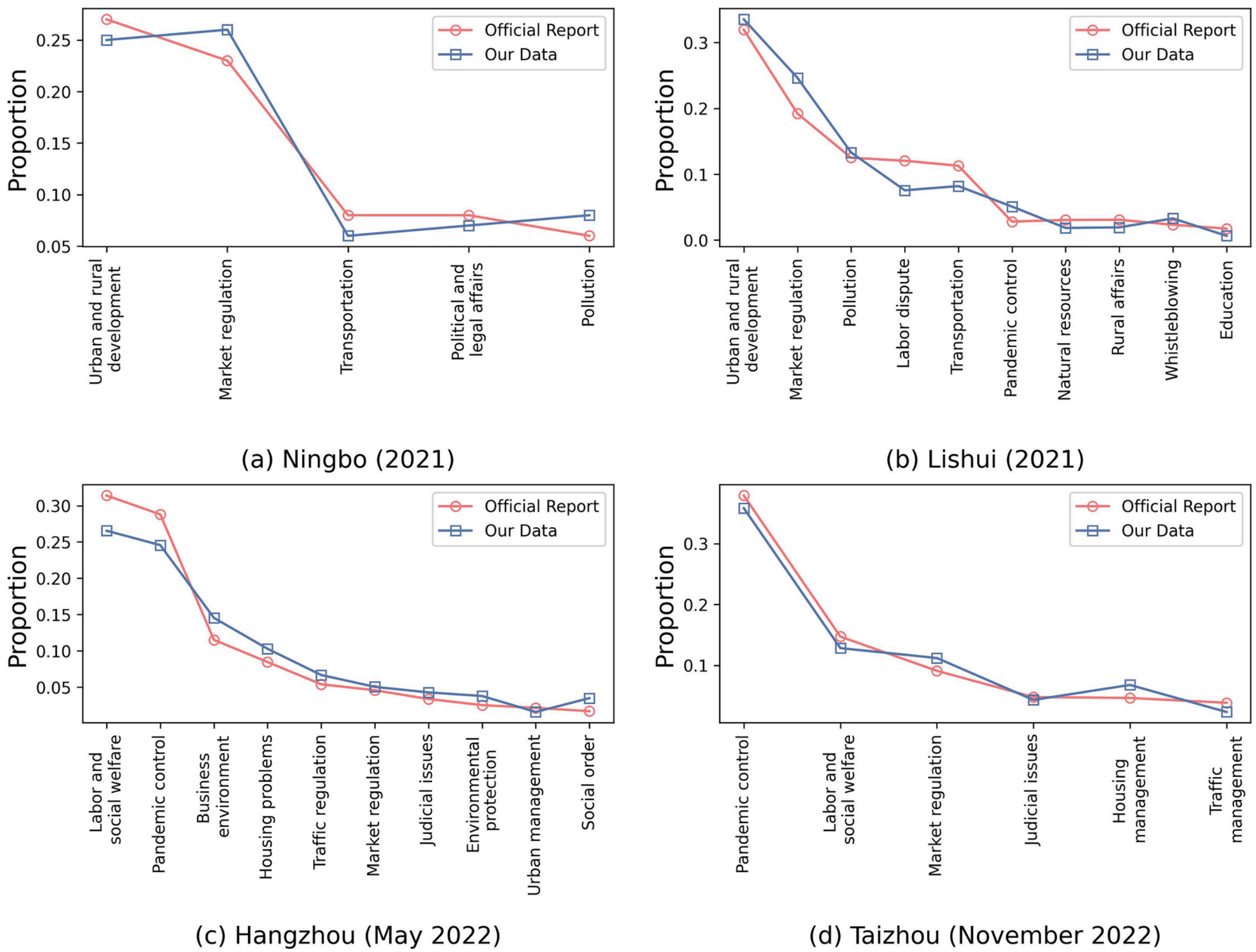


Fig. 2. Comparison of urban problem distributions based on our dataset and official reports from (a) Ningbo in 2021, (b) Lishui in 2021, (c) Hangzhou in May 2022, and (d) Taizhou in November 2022. **Note:** In each subplot, the blue line with square markers represents the urban problem distribution derived from the study's complaint data, while the red line with circle markers depicts the distribution reported by local petition departments. Because the official reports vary in classification granularity and problem types, the x-axis categories differ across subplots. Data sources for the official reports are listed in Appendix B.

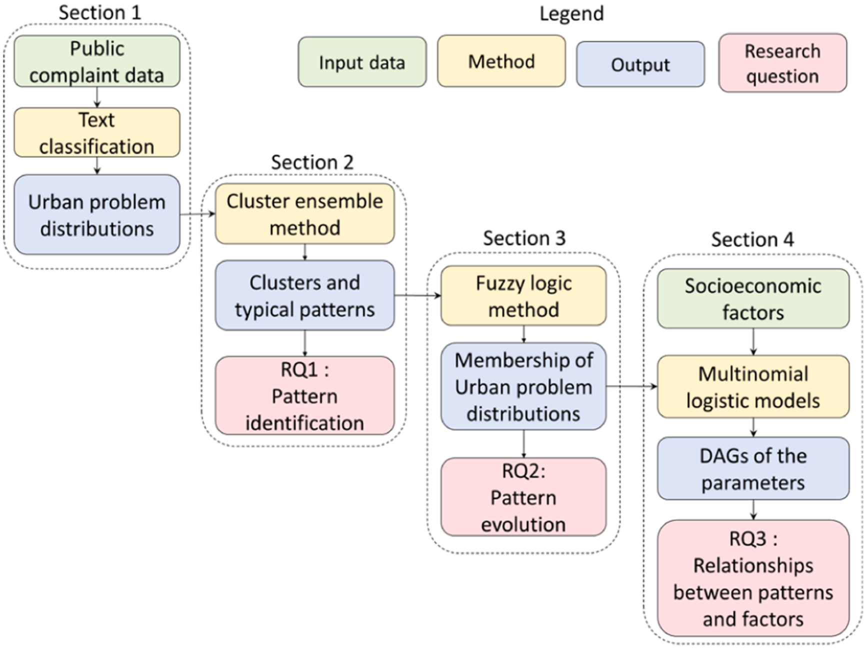


Fig. 3. Methodological workflow.

scheme of 37 urban problem categories (see Table A.1 in Appendix A for details) by manually reviewing 10,000 randomly sampled complaints. Five professional annotators then labeled 50,000 complaints according to this classification scheme. Each complaint was cross-labeled by three annotators, achieving an inter-annotator agreement rate of 94.3 %. Due to class imbalance, we oversampled minority classes in the training data. The final dataset comprised 73,274 samples, ensuring a minimum of 1000 samples per label, and was divided into training and test sets via a stratified 7:3 split to preserve label distributions. The text classifiers were subsequently trained on this corpus.

To automate label assignment, we fine-tuned a large language model (LLM), Qwen2.5–1.5B, for the downstream text classification task. Despite having significantly fewer parameters than larger models such as DeepSeek-R1 (671 billion), Qwen2.5–1.5B demonstrates strong performance in natural language understanding and question answering (Yang et al., 2024), making it a suitable choice for classification. However, full fine-tuning demands roughly 64 GB of GPU memory to store model parameters, optimizer states, gradients, and activations—exceeding the capacity of most GPUs. To mitigate this, we implemented the Low-Rank Adaptation (LoRA), freezing pretrained weights and injecting trainable low-rank matrices to reduce computational demands (Hu et al., 2022). Given a pre-trained weight matrix $\boldsymbol{W}_0 \in \mathbb{R}^{d\times d}$, LoRA represents updates as $\boldsymbol{W}_0 + \boldsymbol{BA}$, where $\boldsymbol{B} \in \mathbb{R}^{d\times r}$, $\boldsymbol{A} \in \mathbb{R}^{r\times d}$, and $r \ll d$. During training, $\boldsymbol{W}_0$ is frozen and only $\boldsymbol{A}$ and $\boldsymbol{B}$ are optimized. The forward pass becomes $\boldsymbol{h} = \boldsymbol{W}_0\boldsymbol{x} + \boldsymbol{BAx}$, and because $\boldsymbol{A}$ and $\boldsymbol{B}$ contain far fewer parameters than $\boldsymbol{W}_0$, LoRA substantially reduces the computational burden of fine-tuning.

Another challenge in fine-tuning Qwen2.5–1.5B for text classification arises from its decoder-only architecture. Unlike encoder-only models such as BERT or RoBERTa, which map input texts directly to class labels, Qwen2.5–1.5B generates outputs sequentially and is better suited for text generation tasks. To align with this architecture, text classification must be reformulated as a question answering task. This can be achieved through instruction-tuning, as proposed by Wei et al. (2021), which trains an LLM to follow task-specific instructions. While initially designed to improve zero-shot performance on unseen tasks, instruction-tuning also facilitates effective fine-tuning for specific downstream applications.

The fine-tuning framework in this study is illustrated in Fig. 4. Given a complaint text, it is first appended with an instruction, which tells Qwen2.5–1.5B that the task is text classification. After tokenization and embedding, the input is converted into a sequence of 1536-dimensional dense vectors, which are then input to 28 decoder layers. Each decoder layer contains four attention projections ($\boldsymbol{W}_Q$, $\boldsymbol{W}_K$, $\boldsymbol{W}_V$, and $\boldsymbol{W}_O$) in the multi-query attention module and two feed-forward projections ($\boldsymbol{W}_{up}$ and $\boldsymbol{W}_{down}$) in the MLP module. We applied LoRA only to $\boldsymbol{W}_Q$ and $\boldsymbol{W}_V$, following the efficient setup proposed by Hu et al. (2022). Specifically, LoRA injects trainable matrices $\boldsymbol{A}_Q$ and $\boldsymbol{B}_Q$, representing the composition of $\boldsymbol{W}_Q$, and $\boldsymbol{A}_V$ and $\boldsymbol{B}_V$, representing the composition of $\boldsymbol{W}_V$, to the decoder layer. These additional parameters constitute 7.4 % of the total model parameters, with the remaining parameters kept frozen during fine-tuning.

After fine-tuning and evaluation, the classifier was applied to 3.5 million complaints, assigning a problem label to each complaint. The labeled complaints were then aggregated by year t, county i, and label j, with n_{tij} denoting the complaint count for each group. Since n_{tij} reflects sampled rather than actual complaint volumes, we normalized these counts into proportional terms p_{tij}, calculated as:

$$p_{tij} = \frac{n_{tij}}{\sum_{l=1}^{37} n_{til}} \quad (1)$$

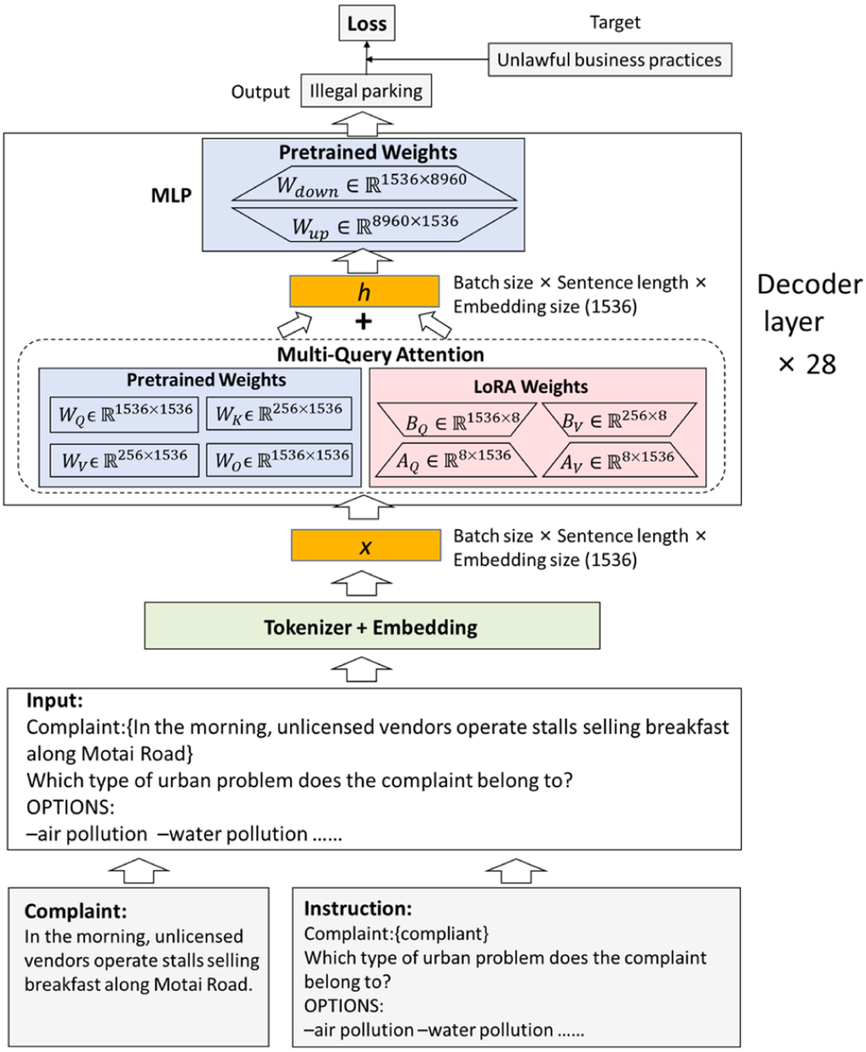


Fig. 4. The proposed framework of fine-tuning Qwen2.5–1.5B for text classification.

The urban problem distribution for the t-th year ($t = 1, 2, \ldots, 5$) in the i th county ($i = 1, 2, \ldots, 87$) is represented by a 37-dimensional vector, $\boldsymbol{P}_{ti} = (p_{ti1}, p_{ti2}, \ldots, p_{ti37})$, where elements sum to 1.

3.3.2. Pattern identification

Urban problem distributions may display distinct patterns, with certain problems disproportionately concentrated in specific counties or years. For example, economically underdeveloped counties often experience higher proportions of infrastructure deficits compared to more developed counties. Distributions with similar patterns exhibit analogous shapes in the 37-dimensional space, enabling clustering methods to group them effectively, with the cluster centroids representing the underlying patterns.

Let $\boldsymbol{P} = \{\boldsymbol{P}_{ti} \mid t\epsilon[1,\ 5],\ i\epsilon[1,\ 87]\}$ denote a dataset of 5×87 points, each corresponding to an urban problem distribution. Our objective is to partition $\boldsymbol{P}$ into K clusters, $C_1,\ C_2, \ldots,\ C_K$. After partitioning, cluster centroids are computed to represent the typical patterns of urban problem distributions. Each centroid $\boldsymbol{c}_k$ is also a 37-dimensional vector, where each element represents the average proportion of an urban problem across all distributions within Cluster C_k:

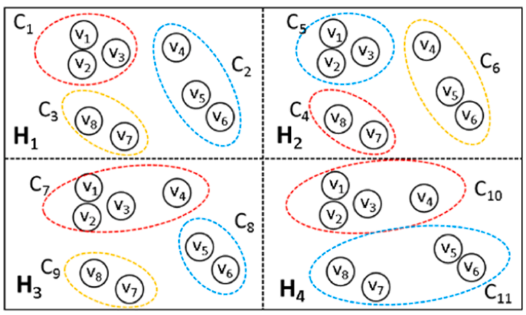
$$\boldsymbol{c}_k = \frac{\sum_{\boldsymbol{P}_{ti} \in C_k} \boldsymbol{P}_{ti}}{|C_k|} \tag{2}$$

where $|C_k|$ is the number of members in Cluster C_k. Given the inherent complexity of urban problem distributions, selecting a single clustering algorithm is challenging due to the varying strengths and weaknesses of different clustering methods (Boongoen & Iam-On, 2018). To mitigate algorithmic biases and improve robustness, we employed a cluster ensemble approach, which consolidates results from multiple base clustering algorithms into consensus-based groupings.

Commonly used cluster ensemble methods include Cluster-based Similarity Partitioning Algorithm (CSPA), Hyper-Graph Partitioning Algorithm (HPGA), Meta-Clustering Algorithm (MCLA), and Hybrid Bipartite Graph Formulation (HBGF). CSPA, HGPA, and MCLA were proposed by Strehl and Ghosh (2003), and HBGF was introduced by Fern and Brodley (2004). Although these methods aim to achieve maximum consensus among multiple base clusterings under different principles, they are not designed to effectively exclude samples that exhibit inherently low consensus. In our context, some urban problem distributions deviate significantly from typical patterns and exhibit low agreement across clusterings; such samples should be grouped into a dedicated “outlier cluster.” The primary goal of filtering outliers is to enhance the accuracy of the identified urban problem patterns. As previously discussed, these patterns are represented by cluster centroids. Including excessive outliers in clusters can distort the centroids toward anomalous instances.

This study proposes a cluster ensemble method that assigns high-consensus samples to regular clusters and low-consensus samples to an outlier cluster. Building upon MCLA, our method introduces an entropy-based threshold to identify low-consensus samples prior to cluster assignment. Given only minimal modifications to the original MCLA, we omit a detailed algorithmic description and instead illustrate the core idea through a simplified example in Fig. 5.

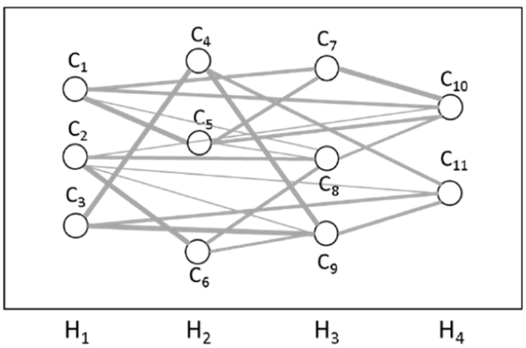
Fig. 5a presents eight data points ($v_1, v_2, \ldots, v_8$) partitioned by four base cluster methods (H_1, H_2, H_3, H_4). For example, H_1 partitions



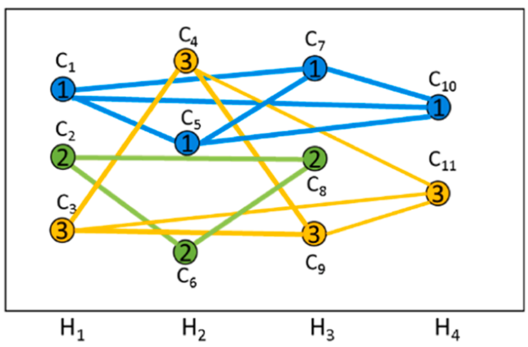
(a)

	H_1			H_2			H_3			H_4	
	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9	C_{10}	C_{11}
v_1	1	0	0	0	1	0	1	0	0	1	0
v_2	1	0	0	0	1	0	1	0	0	1	0
v_3	1	0	0	0	1	0	1	0	0	1	0
v_4	0	1	0	0	0	1	1	0	0	1	0
v_5	0	1	0	0	0	1	0	1	0	0	1
v_6	0	1	0	0	0	1	0	1	0	0	1
v_7	0	0	1	1	0	0	0	0	1	0	1
v_8	0	0	1	1	0	0	0	0	1	0	1

(b)



(c)



(d)

	CE_1 C_1, C_5, C_7, C_{10}	CE_2 C_2, C_6, C_8	CE_3 C_3, C_4, C_9, C_{11}
v_1	1	0	0
v_2	1	0	0
v_3	1	0	0
v_4	0. 50	0. 67	0
v_5	0	1	0. 25
v_6	0	1	0. 25
v_7	0	0	1
v_8	0	0	1

(e)

	Outlier	CE_1 C_1, C_5, C_7, C_{10}	CE_2 C_2, C_6, C_8	CE_3 C_3, C_4, C_9, C_{11}
v_1	0	1	0	0
v_2	0	1	0	0
v_3	0	1	0	0
v_4	0. 68	0. 43	0. 57	0
v_5	0. 50	0	0. 80	0. 20
v_6	0. 50	0	0. 80	0. 20
v_7	0	0	0	1
v_8	0	0	0	1

(f)

Fig. 5. A simplified example illustrating the proposed method: (a) results of four base clustering methods, (b) binary cluster association matrix, (c) four-partite graph capturing consensus among base clustering results, (d) meta-clusters representing the ensemble clusters, (e) results generated by the original MCLA, and (f) results generated by the proposed method.

these points into three clusters: C_1, C_2, and C_3. In Fig. 5b, the results of four methods are encoded into a binary cluster-association matrix, where a value of 1 indicates membership of a data point in a given cluster, and 0 otherwise. In Fig. 5c, each vertex represents a base cluster label. The edge weights are proportional to the similarity between clusters, measured by the Jaccard index. Since clusters within the same method are non-overlapping, no edges exist between vertices from the same method, forming a four-partite graph. In Fig. 5d, the four-partite graph is partitioned into three balanced meta-clusters using the METIS algorithm. Each meta-cluster represents an ensemble cluster label, indicating a group of base clusters with relatively high mutual consensus.

In Fig. 5e, the association vectors of the meta-clusters are computed by averaging the binary association degrees of their constituent clusters. For instance, the association vector of CE_1 is the average of those of C_1, C_5, C_7, C_{10}. Subsequently, the meta-clusters compete for the data points. Specifically, a data point is assigned to the meta-cluster with the highest entry in the association vector. Consequently, CE_1 wins the data points v_1, v_2, and v_3. The original MCLA outputs $\{(v_1, v_2, v_3), (v_4, v_5, v_6), (v_7, v_8)\}$ as the final cluster ensemble result. Notably, the membership of v_4 is relatively uncertain, as it is located far from the three dominant clusters and can be considered an outlier.

In Fig. 5f, we adopt MCLA to separate such outliers from the typical clusters. We first normalize the association vectors of three meta-clusters by rows and then calculate the entropy values for each point. Points whose entropy exceeds a predefined threshold (0.6 in this example), indicating low consensus among different cluster methods, are recognized as outliers, while the rest are assigned to their most associated meta-clusters. The proposed method outputs $\{(v_4), (v_1, v_2, v_3), (v_5, v_6), (v_7, v_8)\}$ as the final cluster ensemble result, where (v_4) represents the outlier cluster.

3.3.3. Evolution analysis

Evolution analysis investigates the nuanced temporal shifts in urban problem patterns across counties. Rather than experiencing abrupt transitions, a county's urban problem distribution may gradually shift from one pattern to another. Particularly during periods of dominant pattern shifts, the urban problem distribution may not fully align with any established pattern but instead reflects a probabilistic mixture of multiple patterns. To trace the evolution of problem patterns in the i th county, we first calculated the similarities between $\boldsymbol{P}_{ti}$ and $\boldsymbol{c}_k$ $(k = 1, 2, \ldots, K)$ for each year. These similarities were then normalized to facilitate the analysis of temporal trajectories, as illustrated in Fig. 6. Assuming a given similarity metric, Fig. 6a displays five bar plots, each showing the similarity between $\boldsymbol{P}_{ti}$ and three problem patterns $(\boldsymbol{c}_1, \boldsymbol{c}_2, \boldsymbol{c}_3)$ across five years $(t = 1, 2, \ldots, 5)$. However, similarity scores alone lack sufficient discriminative power for capturing evolutionary trends. Thus, the similarity scores were converted into pattern memberships, as illustrated in Fig. 6b Finally, Fig. 6c visualizes the annual memberships as a stacked bar plot, highlighting the temporal dynamics of problem pattern evolution in the county.

Common metrics for assessing similarity between probability distributions include the Kullback-Leibler (KL) divergence, Jensen-Shannon (JS) divergence, Wasserstein distance, and total variation distance (TVD). However, KL divergence is asymmetric, and Wasserstein distance is unsuitable for categorical variables due to the difficulty in defining meaningful inter-category distances. Consequently, both JS divergence and TVD emerge as suitable choices for measuring the similarity between urban problem distributions. Preliminary analysis demonstrated a strong linear correlation (coefficient: 0.974) between TVD and JS divergence in this context, indicating that the metric selection had negligible influence on analytical outcomes. We ultimately adopted TVD as our

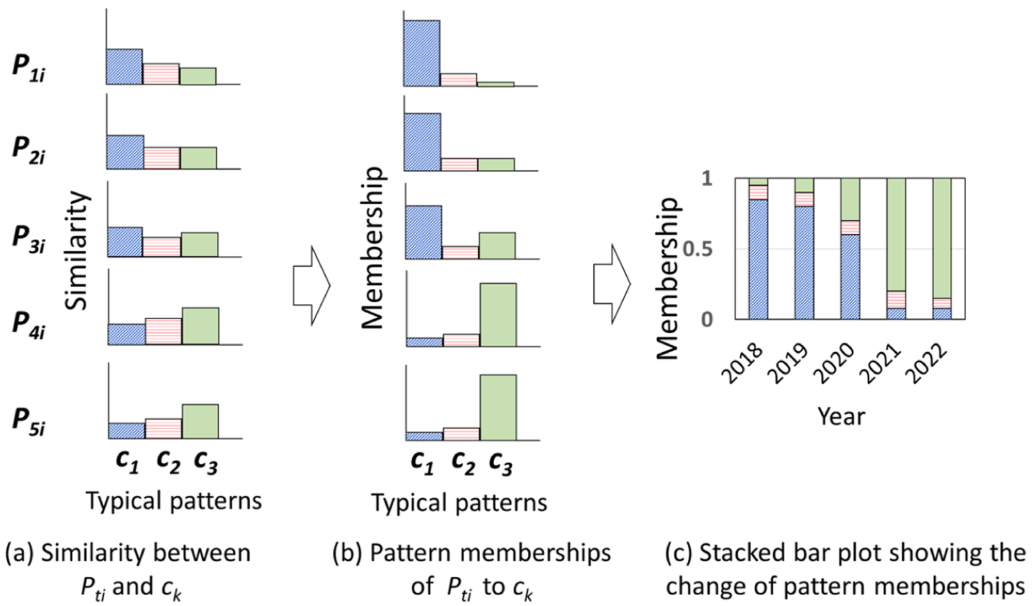


Fig. 6. Evolution analysis of problem patterns: (a) similarity between $\boldsymbol{P}_{ti}$ and $\boldsymbol{c}_k$, (b) pattern memberships of $\boldsymbol{P}_{ti}$ to $\boldsymbol{c}_k$, and (c) stacked bar plot showing the change of pattern memberships.

similarity measure, as it quantifies the maximum difference between two distributions and offers more intuitive interpretability. Specifically, the TVD score between $\boldsymbol{P}_{ti}$ and $\boldsymbol{c}_k$ can be calculated as:

$$TVD(\boldsymbol{P}_{ti}, \boldsymbol{c}_k) = \frac{1}{2}\sum_{j=1}^{37}\left|p_{tij} - c_{kj}\right| \tag{3}$$

As previously discussed, similarity scores are inadequate for tracking evolutionary trajectories of problem patterns. To address this limitation, we converted the similarity scores into partial memberships using fuzzy logic. After calculating $\text{TVD}(\boldsymbol{P}_{ti}, \boldsymbol{c}_k)$ for all combinations of $\boldsymbol{P}_{ti}$ and $\boldsymbol{c}_k$, we constructed K fuzzy sets. Each fuzzy set F_k consisted of 5×87 ordered pairs:

$$F_k = \{(\boldsymbol{P}_{ti},\ \mu_k(\boldsymbol{P}_{ti}) |\ \boldsymbol{P}_{ti} \in \boldsymbol{P})\ (k \in [1,\ K]) \tag{4}$$

Here, $\mu_k(\boldsymbol{P}_{ti}) \in [0, 1]$ is the membership degree of $\boldsymbol{P}_{ti}$ to $\boldsymbol{c}_k$, constrained by $\sum_{k=1}^{K}\mu_k(\boldsymbol{P}_{ti}) = 1$ to enable probabilistic interpretation of pattern membership. Membership degrees were computed using a Gaussian kernel:

$$\mu_k(\boldsymbol{P}_{ti}) = \exp\left(-\frac{\text{TVD}(\boldsymbol{P}_{ti},\ \boldsymbol{c}_k)}{2{\sigma_k}^2}\right) \tag{5}$$

where σ_k is the standard deviation of the distribution of $\text{TVD}(\boldsymbol{P}_{ti}, \boldsymbol{c}_k)$, with $\boldsymbol{c}_k$ held constant.

Using the described method, we computed the memberships of each urban problem distribution relative to the identified patterns. Subsequently, we constructed 87 stacked bar plots—one for each county—illustrating the annual changes in pattern memberships. The generalized evolutionary trajectory was manually summarized through visually inspecting these stacked bar charts. Notably, we avoided automated pattern recognition techniques (e.g., dynamic time warping or principal component analysis), as the limited sample size impedes reliable pattern construction. Manual analysis, conversely, combines domain expertise with direct observation to identify salient evolutionary trends and explicitly link them to interpretable conclusions.

3.3.4. Influencing factor analysis

The evolutionary direction of a county's problem patterns may be closely tied to its socioeconomic development. To investigate this relationship, we selected representative socioeconomic factors and analyzed their correlations and causal influences on pattern memberships using multinomial logit regression. Eight variables, listed in Table 2, were chosen based on four criteria: 1. Minimal collinearity: To avoid inflated standard errors and ensure model interpretability, variables with strong interdependencies were excluded. For example, population density was excluded due to its strong correlation with GDP per capita. 2. High spatiotemporal heterogeneity: The variables should show significant variation across space and time. For example, GDP per capita was preferred over urbanization as a measure of a county's development level due to its higher spatiotemporal heterogeneity. 3. Data accessibility and reliability: All variables were drawn from authoritative sources such as Zhejiang Provincial Government statistical yearbooks, while factors lacking county-level data (e.g., public service quality, enforcement capacity) were excluded. 4. Scale independence: Scale-adjusted metrics (e.g., GDP per capita, road network density) were used instead of absolute values to mitigate county-size biases.

To formalize the relationship between socioeconomic factors and problem patterns, we employed multinomial logit regression models to analyze both the correlational and causal effects. These two models predict membership probabilities by comparing each

Table 2
Eight socioeconomic factors used to analyze their impacts on problem patterns.

Factor	Meaning	Source	Mean	STD
1. GDP per capita	Gross Domestic Product (GDP) per permanent resident.	https://tjj.zj.gov.cn/col/col1525563/index.html	10.30×10^4 (CNY/person)	5.49×10^4 (CNY/person)
2. Mountainous counties	If a county belongs to the "26 Mountainous Counties," it is assigned a value of 1; otherwise 0.	List of 26 mountainous counties	29.9 (%)	45.8 (%)
3. Land acquisition area per capita	Land expropriated or requisitioned per permanent resident.	https://zd.zrzyt.zj.gov.cn/sszq-xxgk/#/HomePage/zdxm	2.98×10^{-4} (ha/person)	2.52×10^{-4} (ha/person)
4. Housing sales volume per capita	The number of presale housing projects per permanent resident.	https://data.zjzwfw.gov.cn/jdop_front/index.do	7.06×10^{-5} (project/person)	5.36×10^{-5} (project/person)
5. Total retail sales of consumer goods per capita	Total retail sales of consumer goods per permanent resident.	https://tjj.zj.gov.cn/col/col1525563/index.html	4.06×10^4 (CNY/person)	1.55×10^4 (CNY/person)
6. Urban-rural income gap	Ratio of average urban income to average rural income.	https://tjj.zj.gov.cn/col/col1525563/index.html	1.78	0.24
7. Road network density	Total road length divided by built-up area.	https://tjj.zj.gov.cn/col/col1525563/index.html	4.65 km/km^2	2.10 km/km^2
8. Administrative penalty rate	Volume of administrative penalties per permanent resident.	https://data.zjzwfw.gov.cn/jdop_front/index.do	7.69×10^{-4} (case/person)	4.12×10^{-4} (case/person)

Note: The "26 Mountainous Counties" are a group of underdeveloped counties designated by the Zhejiang provincial government to reduce regional disparities in development. Located primarily in mountainous areas, they are mainly administered by the cities of Quzhou, Lishui, and Wenzhou. The list of these 26 counties remained unchanged between 2018 and 2022, making this variable time-invariant in this study.

pattern to a predefined reference pattern. The core equation for the correlation model is:

$$\frac{\mathrm{Prob}(\boldsymbol{P}_{ti} \in C_k)}{\mathrm{Prob}(\boldsymbol{P}_{ti} \in C_{Ref})} = \frac{\mu_k(\boldsymbol{P}_{ti})}{\mu_{Ref}(\boldsymbol{P}_{ti})} = exp(\beta_{k0} + \beta_{k1}x_{1ti} + \ldots + \beta_{k8}x_{8ti} + \epsilon_{ti}) \ (k \in [1, K]) \tag{6}$$

where $\mathrm{Prob}(\boldsymbol{P}_{ti} \in C_k)$ is the probability of $\boldsymbol{P}_{ti}$ belonging to Cluster C_k, equivalent to its membership value $\mu_k(\boldsymbol{P}_{ti})$. Likewise, $\mathrm{Prob}(\boldsymbol{P}_{ti} \in C_{Ref})$ represents the probability of belonging to the reference cluster C_{Ref}, equal to $\mu_{Ref}(\boldsymbol{P}_{ti})$. The independent variables x_{1ti},…, x_{8ti} (detailed in Table 2) refer to the i th county in the t-th year, with β_{k1}, …, β_{k8} as the corresponding coefficients for Pattern $\boldsymbol{c}_k$. For K patterns, the model estimates $(K-1) \times 9$ coefficients.

To analyze the causal relationships, we employed a difference-in-differences (DiD) model, specified as:

$$\frac{\mathrm{Prob}(\boldsymbol{P}_{ti} \in C_k)}{\mathrm{Prob}(\boldsymbol{P}_{ti} \in C_{Ref})} = \frac{\mu_k(\boldsymbol{P}_{ti})}{\mu_{Ref}(\boldsymbol{P}_{ti})} = exp(\beta_{kt} + \beta_{kc} + \beta_{k1}x_{1ti} + \ldots + \beta_{k8}x_{8ti} + \epsilon_{ti}) \ (k \in [1, K]) \tag{7}$$

where β_{kt} and β_{kc} represent the time and county fixed effects, respectively.

In multinomial logit models, coefficient analysis is more complex than in binary logit models because the interpretation of the coefficients depends on the selection of the reference pattern. Rather than directly quantifying the probability of selecting a pattern, these coefficients reflect the relative odds of choosing a specific pattern over the reference pattern. To analyze the effects of an independent variable, say x_p, across all patterns, we should focus on the coefficients including β_{1p}, β_{2p},…, β_{Kp}, and 0 (for the reference pattern), collectively forming the set $S_p = \{\beta_{1p}, \ \beta_{2p}, \ldots, \ \beta_{Kp}, \ 0\}$. Unlike standard regression models, statistical significance tests for individual coefficients have limited utility. Instead, meaningful inference arises from comparing differences between pairs of coefficients within S_p. For example, if β_{1p} is significantly greater than 0 in the DiD model, it suggests that increasing x_p increases the odds of choosing the first pattern over the reference pattern, while it reveals nothing about other pairwise comparisons. To evaluate the effect of x_p on the first and second patterns, one must test $\beta_{1p} - \beta_{2p}$, and a statistically significant positive difference implies that increasing x_p raises the odds of selecting the first pattern over the second by a multiplicative factor of $\exp\left(\beta_{1p} - \beta_{2p}\right)$.

However, conducting exhaustive pairwise comparisons among the coefficients in S_p is computationally burdensome and analytically redundant. A more efficient approach involves ranking the coefficients and emphasizing the differences between adjacent pairs. This ranking, however, may not yield a strict hierarchy, as some coefficient pairs might lack statistical significance. To resolve this ambiguity, we employed DAGs to visualize the relationships: nodes represent problem patterns, and directed edges denote hierarchical dependencies. A directed path from nodes k_1 to k_2 indicates that β_{k_1p} is significantly greater than β_{k_2p}. With eight socioeconomic variables and two DAGs per variable, we constructed a total of 16 distinct DAGs.

Fig. 7 illustrates two schematic DAGs—one for the correlation model (Fig. 7a) and one for the DiD model (Fig. 7b)—to visualize the relationships between x_p and four problem patterns. Both DAGs contain four nodes, each representing a problem pattern. In the correlation DAG, a directed path from Node 1 to Node 4 indicates that β_{1p} is significantly larger than β_{4p}, suggesting that counties with higher x_p values are more likely to exhibit Pattern 1 than Pattern 4. In the causation DAG, the path from Node 1 to Node 2 suggests that β_{1p} significantly exceeds β_{2p}, indicating that an increase in x_p raises the likelihood of Pattern 1 relative to Pattern 2. Meanwhile, the absence of a path between Node 2 and Node 3 in Fig. 7b reflects an insignificant difference between β_{2p} and β_{3p}, implying that x_p does not significantly affect the relative likelihood of Patterns 2 and 3.

4. Results

4.1. Basic description of urban problem distributions

To assign problem labels to the complaints, we fine-tuned Qwen2.5–1.5B for text classification. For comparison, we also evaluated three widely used encoder-only models, including BERT (Devlin et al., 2019), RoBERTa (Liu et al., 2019), and XLNet (Yang et al.,

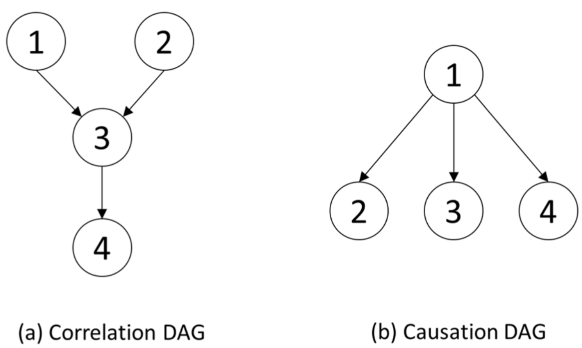


Fig. 7. Two schematic DAGs illustrating the impacts of variable x_p on four problem patterns: (a) correlation DAG and (b) causation DAG.

2019), and two decoder-only LLMs, namely the original Qwen2.5–1.5B and DeepSeek-R1, in zero-shot settings. Table 3 summarizes the model sizes, training details, and macro-averaged Precision, Recall, and F1 scores. Without fine-tuning, LLMs significantly underperformed fine-tuned encoder-only models, primarily because LLMs relied on the semantic interpretation of label names to perform text classification. When the distinctions between label names are subtle or the label descriptions are imprecise, the model often struggles to assign accurate labels. While incorporating the interpretations of label names in prompts can partially alleviate this issue, limited interpretive coverage restricts classification of diverse complaints. After LoRA fine-tuning, Qwen2.5–1.5B achieved the best performance on our dataset, highlighting the value of task-specific adaptation.

The classified complaints were then aggregated by year, county, and problem type to construct the urban problem distribution for each county-year pair. Fig. 8 shows the average proportions of 37 urban problems across 87 counties from 2018 to 2022. The most frequent problems include illegal land use (6.68 %), housing purchase (5.33 %), product quality (5.04 %), whistleblowing (5.03 %), unlawful business (4.54 %), and land expropriation (4.03 %). The predominance of land-, property-, and business-related problems in Zhejiang is likely due to its buoyant market economy, characterized by frequent transactions, active private enterprises, and large-scale land development.

4.2. Typical patterns of urban problem distribution

To identify typical problem patterns from the distributions, we compared the proposed cluster ensemble method with four commonly used alternatives, including CSPA, HPGA, MCLA, and HBGF, using the same sets of base clusterings: k-means, spectral clustering, agglomerative clustering, Gaussian mixture, and PCA-based Gaussian mixture.[4] As shown in Table 4, the proposed method outperformed the others in both the Silhouette coefficient and average Rand index. However, its Silhouette coefficient of 0.236 indicates limited inter-cluster separation. For cluster visualization, we used t-SNE, a dimensionality reduction technique that projects high-dimensional data into a low-dimensional space while preserving local relationships. Fig. 9 reveals that Clusters 1 and 2 are strongly interconnected, as are Clusters 3 and 4, suggesting shared distributional characteristics. In contrast, Clusters 5, 6, and 7 are spatially isolated. The outliers predominantly occupy boundary regions, particularly between Clusters 3 and 4, suggesting the presence of transitional or hybrid patterns.

Fig. 10 illustrates the urban problem distributions for seven identified problem patterns, represented by cluster centroids. In Fig. 10a, rows represent problem patterns and columns denote urban problem types. Cell colors reflect the proportion of each problem within a given pattern, with darker shades indicating higher proportions. Black-bordered cells highlight the representative problems, defined as those with the highest relative proportions across patterns. Some problems, however, are not representative of any pattern, as they show a relatively uniform distribution across all patterns (coefficient of variation $<$ 0.35). In Fig. 10b, proportions are normalized across patterns to better highlight the distinctiveness of representative problems within each pattern.

Pattern 1 is primarily associated with pollution problems—air, water, and noise—coupled with deficient infrastructure and inadequate municipal services. Pattern 2 highlights consumer protection problems, such as car purchases, poor product quality, poor consumer service, prepaid consumption, and unlawful charges. Pattern 3 centers on land-related problems, including illegal land use, unclear land titling, contentious land expropriation, and forced demolition. Pattern 4 is characterized by disputes over housing purchases. Pattern 5 focuses on garbage management problems, with community security and NIMBY problems also prominent. These three problems are frequently observed in Chinese urban communities. Pattern 6 is marked by financial service problems, which comprise 25.2 % of the distribution, approximately 25 times higher than in other patterns. Whistleblowing is also prominent in this pattern. The coexistence of these two problems is complex yet explainable. Many financial service complaints stemmed from investment losses due to peer-to-peer (P2P) platform collapses, while whistleblowing often targets government inaction in resolving the P2P crisis. Pattern 7 is dominated by pandemic control problems, which account for 19.0 % of the centroid's distribution—roughly ten times higher than in other patterns.

Fig. 11 depicts the spatiotemporal distribution of the identified patterns. Pattern 1, characterized by pollution and infrastructure deficiencies, was concentrated in southwestern Zhejiang (Quzhou, Lishui, and Wenzhou) and Ningbo. The spatial clustering in Quzhou and Lishui correlates with their economic underdevelopment, geographical constraints, and limited infrastructure investment. Notably, select counties within the economically advanced Ningbo and Wenzhou also exhibited Pattern 1. This disparity reflects pronounced urban-rural divides, where peripheral counties lag substantially behind urban cores. From 2018 to 2021, most counties classified as Pattern 1 remained stable, although several in Ningbo shifted to Pattern 2, characterized by consumer protection problems. Pattern 2 was predominantly located in Ningbo's municipal districts, and, to a lesser extent, in Wenzhou. From 2018 to 2021, the number of counties classified as Pattern 2 increased slightly.

In 2018, Pattern 3, characterized by land-related problems, was also a common pattern. Counties exhibiting this pattern were mainly concentrated in central and northern Zhejiang, including Hangzhou, Huzhou, Shaoxing, Jinhua, and Taizhou. Unlike Pattern 1, which showed only a slight decrease, Pattern 3 experienced a marked decline in coverage from 2018 to 2020. This decline was counterbalanced by the rise of Pattern 4, marked by housing purchase issues, signaling a shift from land-related to housing-related concerns.

Pattern 6, marked by financial service problems, decreased from 2018 to 2022, with most cases concentrated in Hangzhou. This reduction stemmed largely from the government-mandated shutdown of P2P lending platforms. Once praised for democratizing

[4] These clustering methods were sourced from Scikit-learning (https://scikit-learn.org/1.5/modules/clustering.html)

Table 3
Configurations and performances of six models in the text classification task.

Pretrained Model	Model configurations	Training hyperparameters	Macro precision	Macro recall	Macro F1
1. BERT	Hidden size = 768	Learning rate = 2e-5	91.1	90.5	90.7
2. RoBERTa	Hidden layers = 12	Batch = 16	91.5	90.4	91.1
3. XLNet	Attention heads = 12	Epoch = 3	91.3	89.9	90.4
4. Qwen2.5–1.5B (zero-shot)	Hidden size = 1536 Hidden layers = 28 Attention heads = 12	Not trained	73.4	71.0	72.2
5. DeepSeek-R1 (zero-shot)	Hidden size = 7168 Hidden layers = 61 Attention heads = 128	Not trained	82.5	81.4	81.6
6. Qwen2.5–1.5B (LoRA)	Hidden size = 1536 Hidden layers = 28 Attention heads = 12	Learning rate = 1e-4 Batch = 8 Epoch = 3 Lora rank = 8	**93.3**	**91.5**	**92.2**

Note: The parameter configurations and training settings for BERT, RoBERTa, and XLNet were identical.

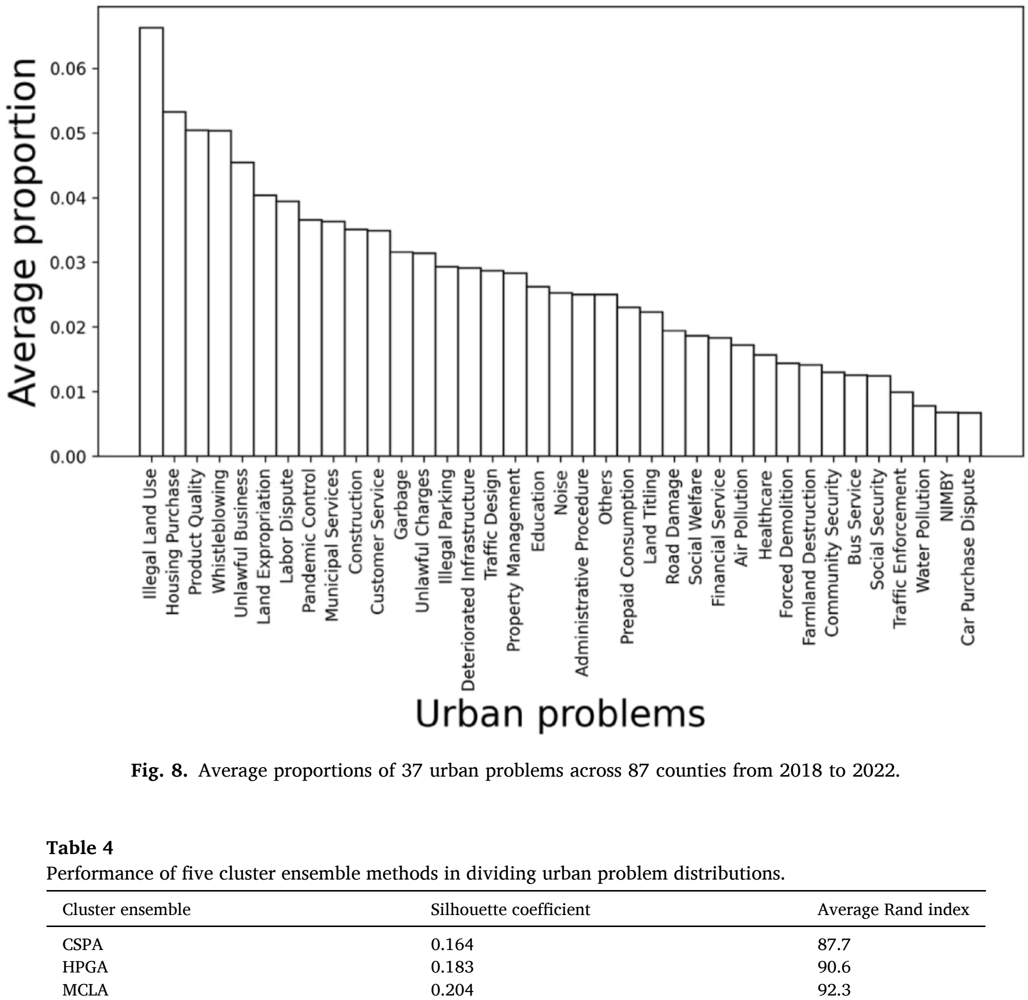


Fig. 8. Average proportions of 37 urban problems across 87 counties from 2018 to 2022.

Table 4
Performance of five cluster ensemble methods in dividing urban problem distributions.

Cluster ensemble	Silhouette coefficient	Average Rand index
CSPA	0.164	87.7
HPGA	0.183	90.6
MCLA	0.204	92.3
HBGF	0.189	89.4
Proposed method	**0.236**	**94.7**

Note: The Silhouette coefficient measures the quality of clustering by assessing both cohesion and separation. The Rand index evaluates the similarity between two clustering outcomes, and the average Rand index refers to the mean similarity between the cluster ensemble results and each base clustering.

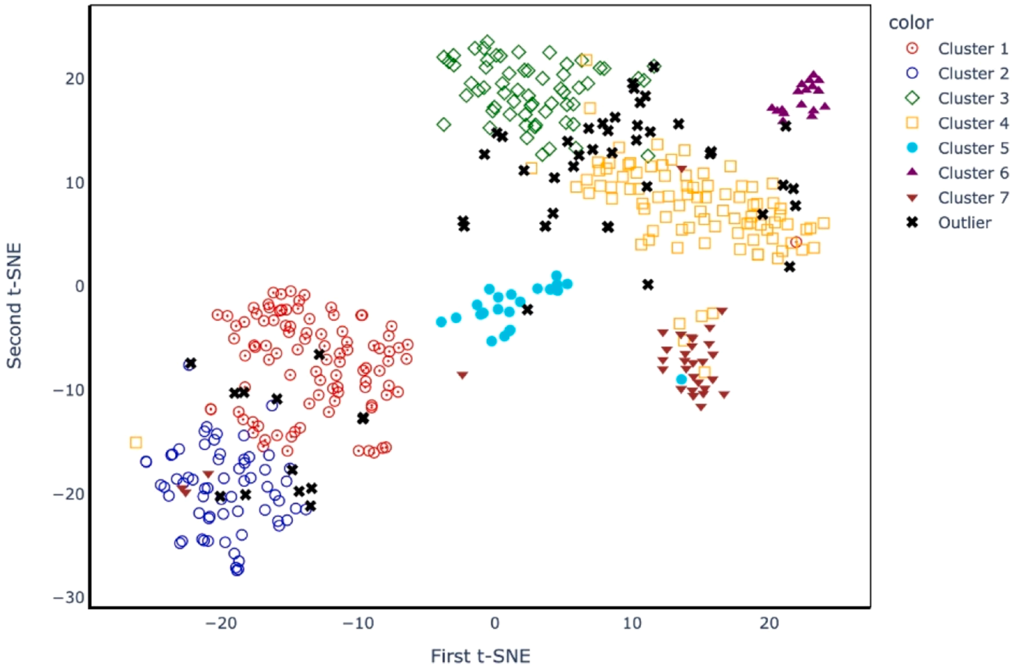


Fig. 9. T-SNE visualization of the identified clusters and outliers.

finance, China's P2P market expanded rapidly between 2007 and 2017 but collapsed due to regulatory failures. As a major hub, Hangzhou became the crisis epicenter in 2018 due to widespread defaults, Ponzi schemes, and mismanagement (Huang & Pontell, 2023). Investor losses and public unrest triggered national intervention, leading to the shutdown of most platforms by 2020. Consequently, related complaints dropped sharply by 2021, and fewer counties were classified under Pattern 6.

Patterns 5 and 7 exhibited limited spatial concentration and temporal continuity, setting them apart from the other patterns. The proportion of counties in Pattern 7, characterized by pandemic control problems, surged in 2022, overshadowing other urban problems and leading to abrupt county reclassifications into Pattern 7. Although pandemic complaints emerged with the initial COVID-19 outbreak in 2020, the 2022 surge likely originated from intensified enforcement of China's Zero-COVID policy. Compared to Pattern 7, Pattern 5, characterized by garbage problems, demonstrated more erratic spatiotemporal variation. Episodic and acute garbage management crises in specific years led to the sporadic reclassification of counties into Pattern 5.

4.3. Evolution of problem patterns

The spatiotemporal map reveals that most problem patterns display temporal continuity. On average, a pattern persists with a probability of 73.2 % between consecutive years within a county, indicating a strong tendency toward self-replication. However, relying solely on discrete patterns may obscure the nuanced temporal dynamics. Several key questions remain unresolved. First, in counties where pattern transitions occurred from 2018 to 2022, were these shifts abrupt or gradual? Second, in counties where patterns were stable, how did the dominance of the prevailing patterns evolve? Discrete representations of urban problem distributions may mask gradual changes and fail to account for distributions not captured by the discrete patterns. To address this limitation, we incorporated fuzzy logic into the analytical framework. Fuzzy logic enables a more precise assessment of the degree to which a distribution aligns with each typical pattern, enabling analysis of continuous evolution. Temporal changes are visualized using stacked bar plots of membership degrees. Fig. 12 presents these evolutions, with counties manually grouped into five categories to highlight distinct evolutionary trajectories.

Fig. 12a depicts the counties with pronounced pattern persistence, where a dominant pattern persists throughout the five-year study period. Specifically, twelve counties consistently aligned with Pattern 1 (indexed 1 to 12), and seven with Pattern 2 (indexed 13 to 19). While these dominant patterns persisted, their membership degrees slightly declined, accompanied by gradual increases in secondary pattern memberships. For example, counties dominated by Pattern 1 saw a slight rise in Pattern 2 membership, while those dominated by Pattern 2 experienced a continued decline in Pattern 1 membership. On the other hand, Fig. 12b illustrates counties undergoing gradual pattern transitions, characterized by a significant decline in the original dominant pattern and a corresponding increase in the emerging pattern. Specifically, eight counties shifted from Pattern 1 to Pattern 2 (indexed 20 to 27), nine from Pattern 3 to Pattern 4 (indexed 28 to 36), and two from Pattern 6 to Pattern 4 (indexed 37 to 38). Notably, during these transitions, the original pattern's membership consistently declined while the emerging pattern's membership rose, suggesting a predictable evolutionary trajectory for problem patterns.

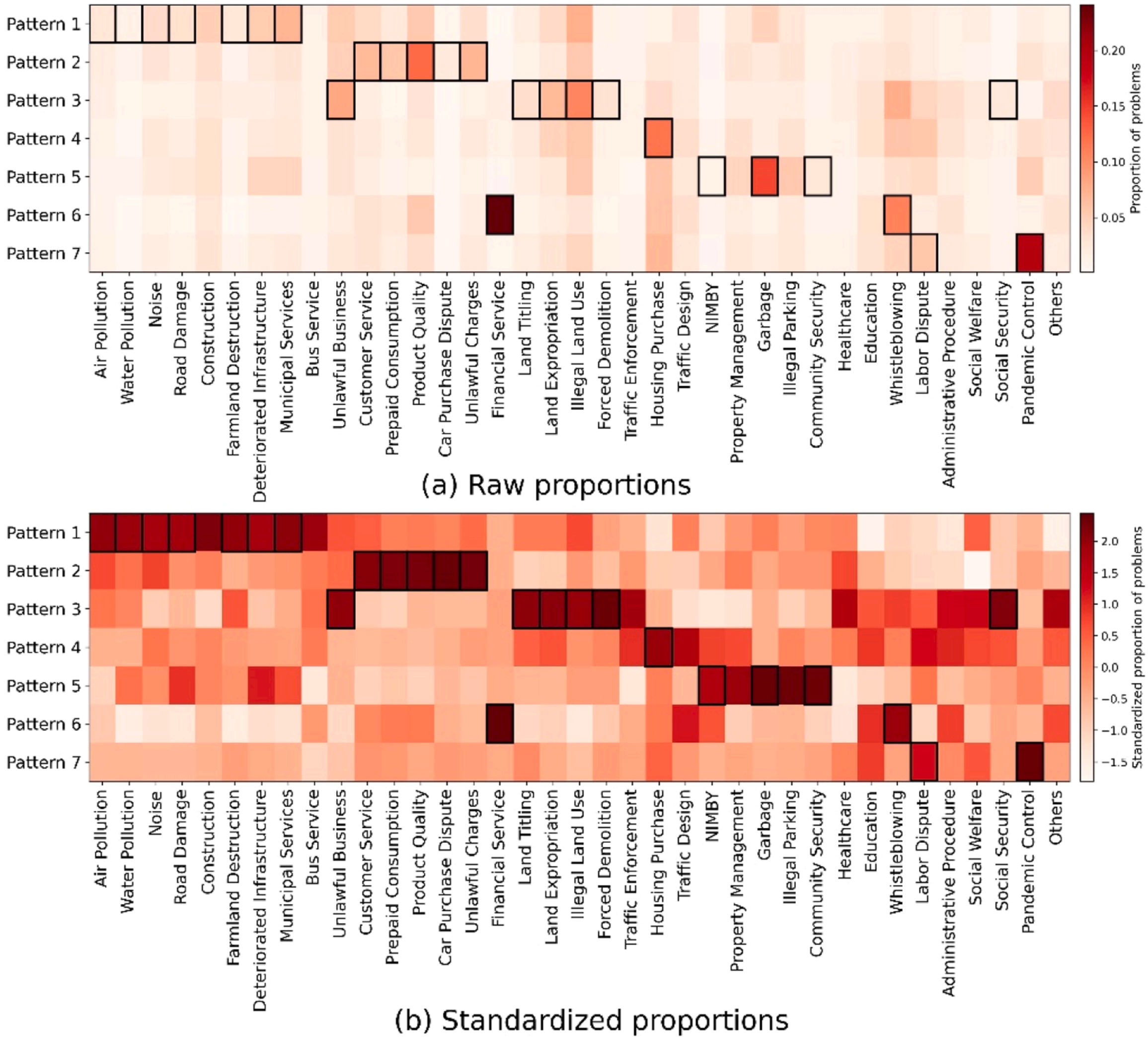


Fig. 10. Seven problem patterns and associated representative problems: (a) raw proportions and (b) normalized proportions.

Whether a county undergoes pattern persistence or pattern transition depends on whether its dominant pattern changes during the observation period. Extending the observation period to ten years can detect gradual shifts undetectable in five years. Conversely, a shorter observation period of two or three years may miss changes that become evident over longer spans. For example, focusing solely on 2020–2022, counties indexed from 28 to 32 in Fig. 12b may consistently align with Pattern 4, suggesting pattern persistence rather than transition. Thus, the length of time window significantly influences the interpretation of the pattern evolutionary process. Regardless of the observation window, both pattern persistence and pattern transition exhibit a consistent inverse relationship: as membership in one pattern increases, another declines. The key difference is that pattern transition exhibits a faster rate of membership change: persistence entails slow growth of an emerging pattern without significantly eroding the dominant pattern's prevalence, whereas transition involves rapid replacement of the dominant pattern during the observation period.

However, the continuous pattern evolution in certain counties can be abruptly disrupted by the emergence of unexpected patterns within a single year. Fig. 12c displays the counties where pattern persistence was interrupted during the observation period. For example, counties indexed from 39 to 42 experienced the sudden replacement of dominant Pattern 1 with Patterns 5 or 7, while counties indexed from 45 to 49 experienced Pattern 4 being replaced by Patterns 5 or 7. Fig. 12d highlights the counties experiencing sudden disruptions during the pattern transitions. For instance, counties indexed from 55 to 70 experienced sudden shifts to Pattern 7 when transitioning from Pattern 3 to Pattern 4. Notably, in both types of evolutionary trajectories, the interrupting patterns did not show signs of emergence in earlier years.

Except for Patterns 5 and 7, most patterns demonstrated inherent predictability and did not disrupt evolutionary trajectories. Further analysis of urban problem distributions revealed that counties experiencing abrupt shifts in patterns had disproportionately high levels of pandemic control or garbage problems, especially in 2022, which significantly reshaped the problem patterns. Excluding these two problems restored previous trajectories: after recalculating the membership values for the affected counties in Fig. 12c and

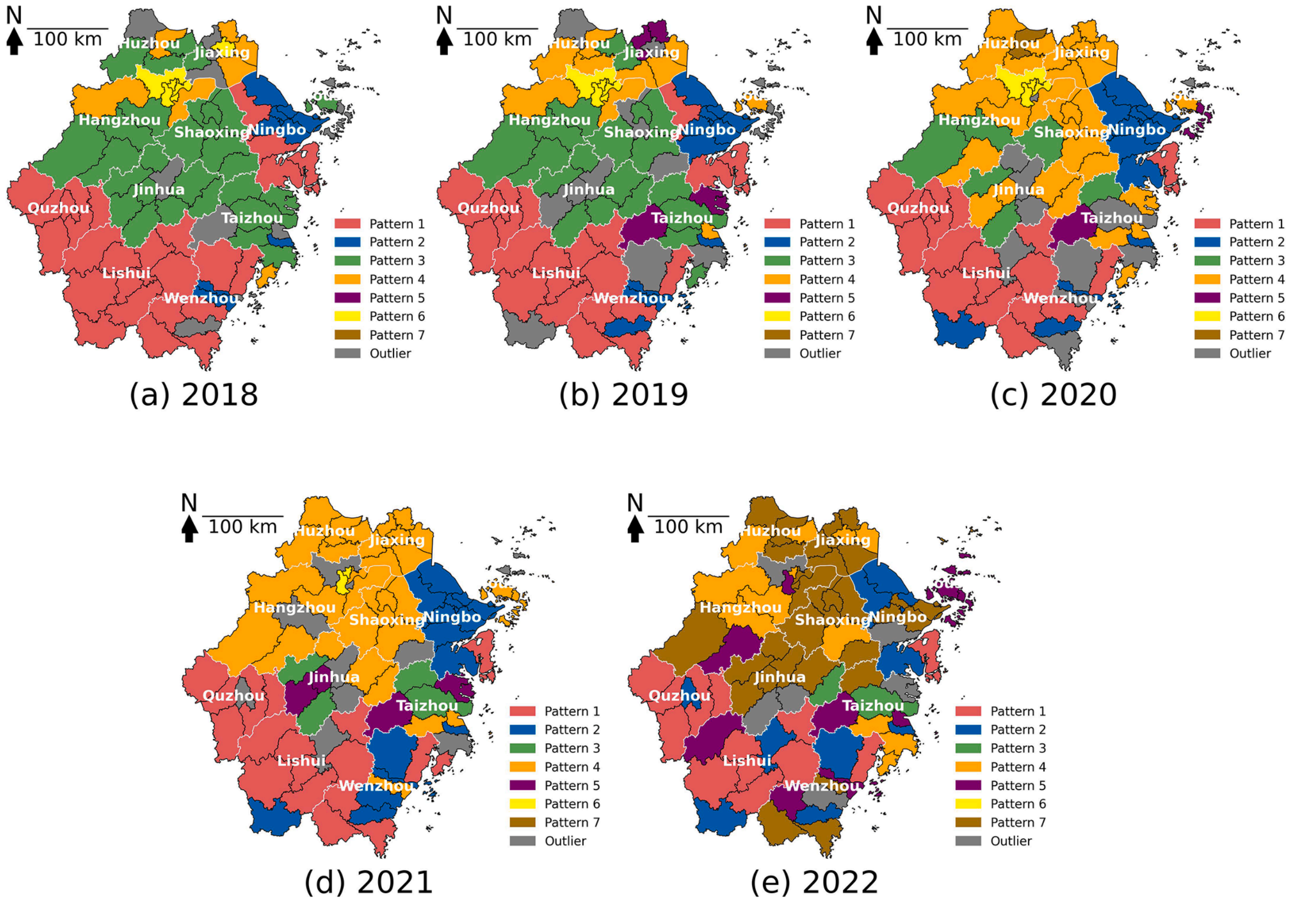


Fig. 11. Spatial distribution of the identified patterns (2018–2022): (a–e) annual results for 2018, 2019, 2020, 2021, and 2022, respectively.

Fig. 12. County-level temporal evolution of pattern membership (2018–2022): (a) pattern persistence, (b) pattern transition, (c) disrupted pattern persistence, (d) disrupted pattern transition, and (e) irregular pattern evolution.

Note: Each stacked bar plot illustrates the membership dynamics of seven problem patterns in a given county from 2018 to 2022. Taking the 38th subplot as an example (highlighted in a red box), the x-axis represents the years (2018–2022), and the y-axis indicates pattern memberships. In 2018, the membership of Pattern 6 (yellow) was nearly 1, suggesting that the county was predominantly characterized by Pattern 6. By 2022, Pattern 4 (orange) had emerged as the dominant pattern, replacing Pattern 6. For visual clarity, the axes are omitted from the other stacked bar plots.

Fig. 12d, we observed renewed pattern persistence or gradual transitions (see Fig. C.1 in Appendix C for details). These results confirm that the 2022 surge in pandemic and garbage-related complaints disrupted otherwise stable evolutions. Besides pattern persistence, transition, and sporadic disruptions, a subset of counties exhibited irregular trajectories (Fig. 12e). For example, counties indexed 79 and 80 showed nonmonotonic trends in Pattern 2, while counties indexed from 81 to 87 displayed disordered pattern evolution, with their problem distributions dominated by more than two patterns.

4.4. Analysis of influencing factors

To assess the impacts of socioeconomic factors in shaping the problem patterns, we employed two multinomial logit models, with results detailed in Fig. 13. Each subplot contrasts causal and associative relationships via two DAGs: the left panel displays a correlation DAG derived from the ordinary model (Eq. (6)), while the right panel illustrates a causation DAG from the DiD model (Eq. (7)). Correlation DAGs exhibit denser connectivity than causation DAGs, underscoring the prevalence of associative over causal relationships. This is consistent with the principle that causation implies correlation, but correlation alone does not confirm causation. Subsequent analyses explore the impacts of each factor on seven problem patterns through observing the two DAGs.

(1) GDP per capita (Fig. 13a)**:** The correlation DAG positions Pattern 1 (pollution and infrastructure problems) and Pattern 3 (land use problems) at the bottom, suggesting their prevalence in economically disadvantaged counties or periods. However, the causation DAG reveals no clear hierarchical structure, even for Pattern 1, where causal links were anticipated. To evaluate the

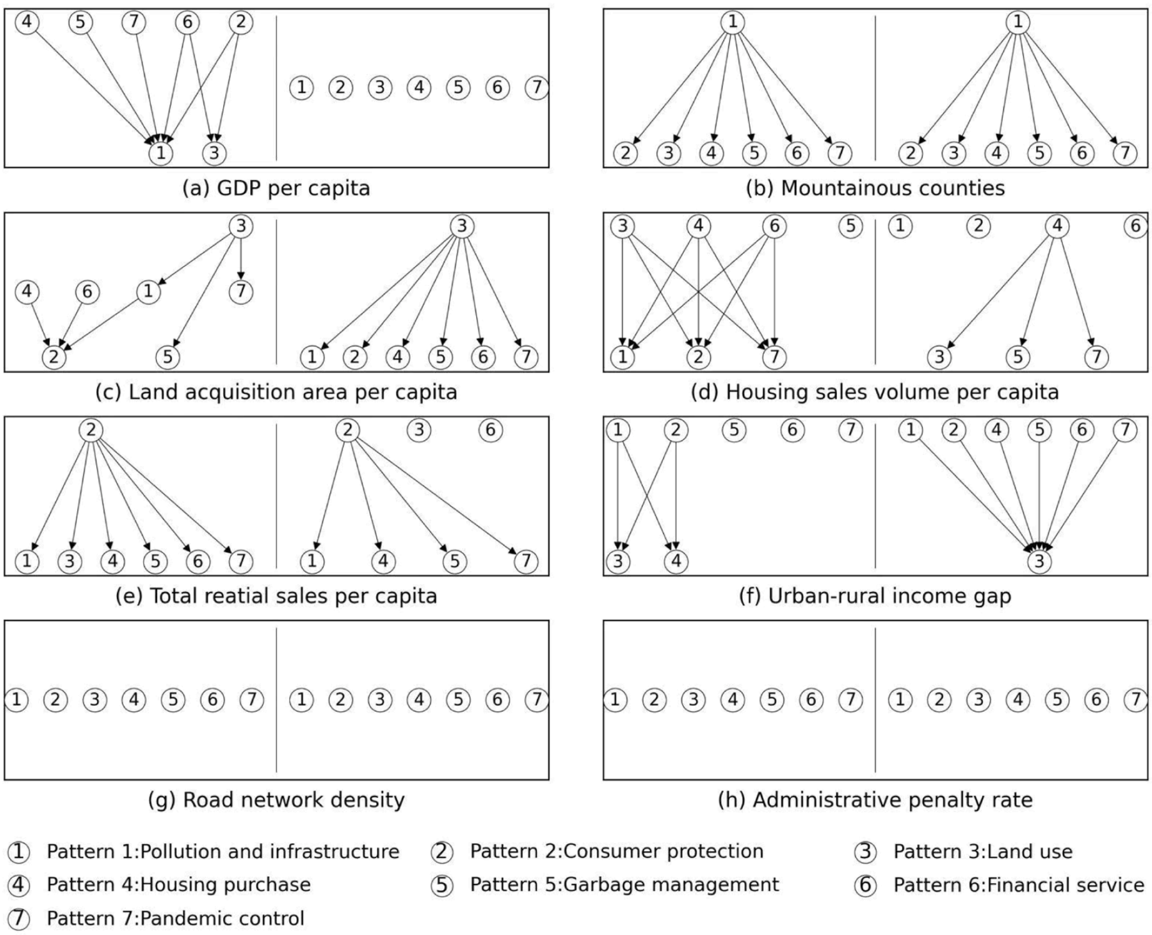


Fig. 13. DAGs reflecting the correlation and causation relationships between problem patterns and socioeconomic factors, including (a) GDP per capita, (b) mountainous counties, (c) land acquisition area per capita, (d) housing sales volume per capita, (e) total retail sales per capita, (f) urban-rural income gap, (g) road network density, and (h) administrative penalty rate.

Note: Each subplot illustrates the impact of a single socioeconomic factor on seven problem patterns using two DAGs: the left DAG depicts associative effects, while the right shows causal effects. For example, the left DAG in Fig. 13a shows that counties with lower GDP per capita are more likely to exhibit Pattern 1, as Node 1 is positioned at the bottom. In contrast, the right DAG indicates no significant causal effect of GDP per capita, with all nodes horizontally aligned.

lagged economic effects, we replaced current-year GDP with one- and two-year lagged values and re-estimated the models. The absence of significant causal relationships across these models implies that short-term economic growth exerted a negligible effect on Pattern 1. Over the past two decades, China's municipal infrastructure development has consistently outpaced depreciation (Liu et al., 2017), leading to sustained improvements in facilities and services. Consequently, Pattern 1′s temporal variation shows limited sensitivity to short-term economic fluctuations, implying that counties undergoing temporary economic downturns may still maintain progress in mitigating pollution and infrastructure deficits.

The observed negative correlation between economic advancement and Pattern 3 likely stems from the study's temporal scope, where rising GDP per capita coincided with declining Pattern 3 memberships. However, an analysis of the alternative period (2010–2018)—characterized by concurrent surges in land expropriation disputes and economic growth in Zhejiang Province (Zhejiang Provincial Bureau of Statistics, 2010–2018)—might instead reveal a positive correlation between Pattern 3 and economic development. This contextual interpretation aligns with the causal DAG's lack of direct links, suggesting their relationship is spatiotemporal rather than intrinsically deterministic.

(2) Mountainous counties (Fig. 13b)**:** In both the correlation and causation DAGs, Pattern 1 ranks the highest, indicating that membership in the "26 Mountainous Counties" significantly increases the likelihood of exhibiting Pattern 1. These counties generally experience low levels of economic development. Notably, the association between mountainous status and Pattern 1 remains significant even after controlling for GDP per capita, suggesting the presence of additional contributing factors. Beyond economic limitations, these counties face challenges such as rugged terrain, poor accessibility, and ecological conservation policies that restrict large-scale land development. Consequently, their infrastructure and environmental quality lag considerably behind provincial averages, despite some improvements observed between 2018 and 2022.

(3) Land acquisition area per capita (Fig. 13c)**:** In both the correlation and causation DAGs, Pattern 3 ranks the highest, indicating that increased land expropriation significantly raises the likelihood of belonging to this pattern. As land requisition and acquisition expand, problems such as land titling, compensation disputes, and forced demolitions naturally intensify. Additionally, extensive land expropriation and housing demolition incentivize residents to construct unauthorized buildings to claim higher compensation or avoid displacement, thereby increasing the proportion of illegal land use complaints. Patterns 2 and 5 rank the lowest in the correlation DAG and show no significant hierarchical relationship with other patterns in the causation DAG, except for Pattern 3. This suggests that although counties categorized under Patterns 2 or 5 had the lowest per capita land expropriation, reducing land expropriation did not increase the likelihood of transitioning to these patterns.

(4) Housing sales volume per capita (Fig. 13d)**:** In the correlation DAG, Patterns 3, 4, and 6 rank higher. The causation DAG, however, reveals that only Pattern 4, which centers on housing purchase problems, shows a significant causal link with housing sales volumes. From 2018 to 2022, most complaints on hotline platforms focused on property transaction disputes, such as prepaid properties and bundled parking spaces, rather than affordability concerns. This suggests that, in the absence of effective regulation, higher transaction volumes correlate with increased complaints rather than a reduction in complaints. In Zhejiang Province, enhancing housing affordability indeed mitigated housing challenges, but also risked exacerbating transaction-related problems, underscoring the necessity for stricter oversight of real estate market. Higher property sales volume reduced the likelihood of a county being classified as Pattern 3, marked by land problems. However, counties in this pattern often showed increased property sales. This is not contradictory, as the counties in Pattern 3 tended to transition to Pattern 4 in subsequent years. During the transition, property sales continued to rise even while counties remained classified as Pattern 3, resulting in higher property sales volumes compared to other counties.

(5) Total retail sales of consumer goods per capita (Fig. 13e)**:** Pattern 2, characterized by consumer protection problems, ranks the highest in both correlation and causation DAGs, indicating that higher retail sales of consumer goods contribute to an increase in such problems. Retail sales reflect the level of market activity. While frequent transactions of goods and services can stimulate local economic growth, they also create complexities that may compromise consumers' interests. In competitive markets, companies may resort to unfair practices, especially when regulatory agencies struggle to keep up with the pace and volume of transactions. For example, manufacturers might reduce quality control to boost sales, and others may cut after-sales services to lower costs. These strategies ultimately worsen consumer protections.

(6) Urban-rural income gap (Fig. 13f)**:** In the correlation DAG, Patterns 1 and 2 rank above Patterns 3 and 4, while in the causation DAG, only Pattern 3 ranks at the bottom, with no significant hierarchical relationships among the remaining patterns. Counties with a larger study's temporal scope, where rising GDP per capita coincided with declining Pattern 3 memberships. However, an analysis of the alternative period (2010–2018)—characterized by concurrent surges in land expropriation disputes and economic growth in Zhejiang Province (Zhejiang Provincial Bureau of Statistics, 2010–2018)—might instead reveal a positive correlation between Pattern 3 and economic development. This contextual interpretation aligns with the causal DAG's lack of direct links, suggesting their relationship is spatiotemporal rather than intrinsically deterministic.urban-rural divide are primarily in southern Zhejiang Province, where Patterns 1 and 2 prevail. However, a decreasing urban-rural income gap does not significantly impact the likelihood of these patterns, suggesting a correlational rather than causal relationship. Instead, the causation DAG indicates that an increasing urban-rural income gap significantly reduces the likelihood of counties belonging to Pattern 3. While the DiD model can reveal causal relationships, it cannot fully address reverse causality. Therefore, the diminishment of Pattern 3 may be directly related to an expanding urban-rural income gap. Counties dominated by Pattern 3 often undergo large-scale land expropriation and housing demolition, transforming a substantial portion of the agricultural population into non-agricultural status while simultaneously increasing agricultural incomes, thereby narrowing the urban-rural income gap (Huang et al., 2024; Yi et al., 2024).

Finally, the variables “road network density” and “administrative penalty rate,” intended to capture the influence of municipal infrastructure and regulatory enforcement, demonstrate statistically insignificant effects on problem patterns in the corresponding DAGs. The lack of significance is likely caused by the measurement limitations in representing infrastructure investments and enforcement capacity. Future investigations should incorporate more precise indicators, such as fixed asset investment and administrative governance costs, for a more comprehensive analysis.

In summary, we categorized the problem patterns into three types based on their relationship to the socioeconomic context: development-driven, trend-driven, and event-driven. First, the socioeconomic factors, such as land acquisition area, housing sales volume, and retail sales volume, play a crucial role in shaping Patterns 2–4. Although these factors are typically viewed as positive indicators of urban and economic development, they may inadvertently undermine residents’ well-being and lead to development-driven problems when governance frameworks fail to balance growth with livability. Second, Pattern 1, disproportionately concentrated in economically disadvantaged counties, does not exhibit significant causal relationships to the short-term change of socioeconomic factors. Despite this, it follows a consistent downward trend over time, suggesting that it is primarily trend-driven, influenced by long-term socioeconomic progression rather than short-term fluctuations. Third, Patterns 5–7 lack significant correlations with the examined factors in Table 2. As detailed in Section 4.2, these are instead tied to discrete incidents or shocks: Pattern 5 to garbage management failures, Pattern 6 to the collapse of P2P platforms in Hangzhou, and Pattern 7 to the stringent Zero-COVID policies in 2022. These event-driven patterns reflect acute crises that sparked widespread public dissatisfaction, dominating urban problem distributions in specific years or counties.

5. Discussions

5.1. Implications

5.1.1. Relationships among representative urban problems

Concerning **RQ1** (“What patterns characterize the urban problem distributions across counties and years?”), we identified seven typical urban problem patterns based on 5×87 urban problem distributions. Each pattern is described by a set of representative problems—issues with the highest relative proportions within a pattern and significant variation across patterns. A key implication is that these representative problems often exhibit certain relationships. For example, the representative problems of Pattern 1 include air pollution, water pollution, and noise, all linked to environmental degradation. We propose four relationship types to interpret these linkages:

(1) Direct causality: Problem A directly causes Problem B, so resolving Problem A can substantially reduce Problem B. For example, in Pattern 6, financial service defaults (e.g., defaults in P2P or bonds) often co-occurred with whistleblowing of authorities. Complaint texts suggest that unresolved financial defaults were frequently followed by citizens’ complaints about administrative negligence and buck-passing. Once financial defaults were resolved, whistleblowing complaints sharply declined.
(2) Indirect causality: Problem A contributes to Problem B through an intermediate mechanism, and addressing Problem A can mitigate Problem B. For example, in Pattern 7, pandemic-control measures such as strict lockdowns and prolonged quarantines disrupted economic activity, leading to layoffs and wage reductions, which in turn triggered labor disputes (Gong et al., 2024).
(3) Shared underlying factors: Problems A and B arise from shared underlying conditions. For example, in Pattern 2, disputes about car purchases, customer service, and product quality frequently coincided, reflecting common weak regulatory oversight and insufficient consumer protections.
(4) Contextual co-occurrence: Problems with distinct causes overlap because of broader socioeconomic conditions. For example, in Pattern 1, air pollution, caused by illegal industrial emissions, often appeared alongside road damage, caused by insufficient infrastructure maintenance. While driven by different factors, both reflect challenges common in economically underdeveloped regions.

Future research should systematically analyze these relational structures, employing network analysis to move beyond descriptive categorization. Such work could deepen the understanding of urban problem patterns and derive the insights about problem resolution.

5.1.2. Pattern trend prediction

Concerning **RQ2** (“How do these problem patterns evolve over time?”), our analysis showed that emerging patterns generally increase in prevalence, whereas dominant patterns tend to decline, with occasional deviations. This suggests that the temporal trends

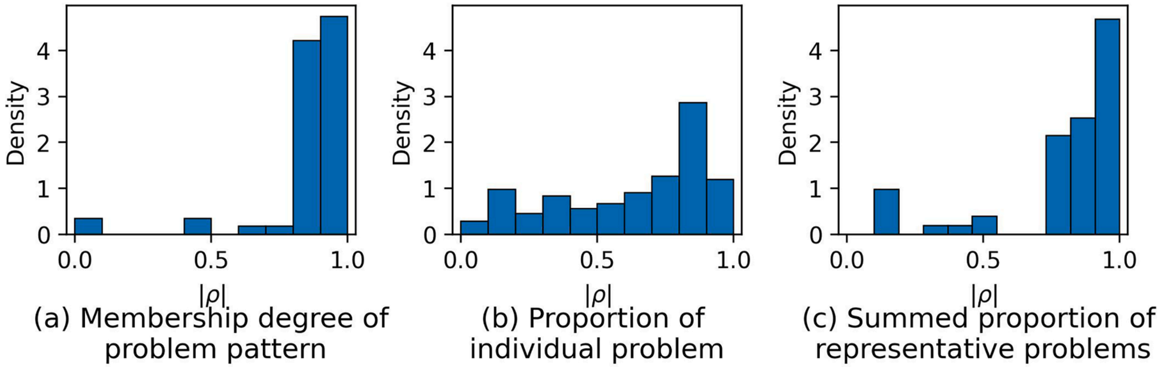


Fig. 14. The frequency density plot of $|\rho|$ between (a) time and membership degree of problem pattern, (b) time and proportion of individual problem, and (c) time and summed proportion of representative problems.

of a county's dominant or emerging patterns are relatively stable, which can support short-term forecasting of pattern dynamics. In contrast, predicting the trajectories of a county's individual urban problems is more challenging. To examine this difference, we calculated the Spearman rank correlation coefficients (ρ) between time and membership degrees of dominant or emerging patterns, and between time and proportions of representative problems, focusing on counties showing pattern persistence or gradual persistence. Because both strongly positive and strongly negative ρ indicate consistent monotonic trends, we report absolute values ($|\rho|$).

As shown in Fig. 14a, $|\rho|$ values for dominant and emerging patterns cluster near 1, indicating strong monotonic trends over time. In contrast, the $|\rho|$ distribution for individual problems is broader and shifts toward lower values (Fig. 14b), reflecting greater volatility of temporal trends. However, aggregating the proportions of representative problems in each pattern (Fig. 14c) again yields $|\rho|$ distributions concentrated near 1, implying that when individual fluctuations are combined, the collective pattern dynamics become more stable. As each problem pattern is characterized by its representative problems, this supports the interpretation that problem patterns smooth out idiosyncratic changes in single problems, producing relatively persistent trajectories.

These findings confirm that while the trajectories of individual problems can fluctuate markedly, the temporal evolution of dominant and emerging patterns remains comparatively stable. Practically, this stability indicates that historical dominant and emerging patterns can inform short-term forecasting of urban problem structures. Such forecasts may help governments, especially in budget-driven planning systems like China's, allocate governance resources more effectively. Nevertheless, contingency measures remain essential, as unforeseen events may disrupt expected trajectories.

5.1.3. Implication for urban governance strategies

Regarding **RQ3** ("What socioeconomic factors drive the evolution of these patterns?"), we found that Patterns 2–4 exhibit significant short-term causal links with socioeconomic factors, such as retail sales volume. We therefore classify them as development-driven. Pattern 1 showed no significant short-term causal links with the examined factors but appeared less frequently in economically advanced counties and continued to decline, suggesting it is shaped by long-term structural dynamics and can be considered trend-driven. Patterns 5–7 did not exhibit significant correlations with the examined socioeconomic factors. However, analysis of their representative complaints indicates that they are linked to abrupt, large-scale shocks, such as the collapse of P2P platforms and the restrictions under Zero-COVID policy. We thus identify them as event-driven.

These distinct driving mechanisms require differentiated strategies for urban governance. Development-driven patterns, stemming from tensions between rapid socioeconomic growth and well-being, require stronger regulation of fast-growing sectors and the design of policy tools that reduce negative externalities associated with socioeconomic activities. Trend-driven patterns call for long-term, system-oriented interventions, including top-down regional planning, targeted fiscal investment, and infrastructure upgrades to address persistent disparities. Event-driven patterns underscore the necessity of robust crisis management systems, featuring transparent communication, rapid response capacity, and efficient recovery mechanisms.

Overall, systematically integrating citizen complaint data into governance processes can complement conventional enforcement-focused approaches. While traditional strategies may achieve short-term compliance, they often fail to resolve underlying structural issues. Leveraging these data enables earlier detection of emerging problem patterns, supports interventions tailored to local contexts, and fosters durable solutions to complex urban challenges.

5.2. Limitation and further study

This study has several limitations. First, while urban problem patterns capture structural characteristics of problems, they do not necessarily reflect governance costs. Problem patterns fail to indicate problem severity or the fiscal burden of mitigation, meaning counties classified under the same pattern may face widely different costs. Addressing this requires considering both the absolute complaint volumes and proportional distribution. However, data collection remains challenging. Hotline complaints provide a useful proxy, but variations in 12,345-hotline penetration across regions and years may bias spatiotemporal analyses. Urban areas typically generate more complaints due to higher engagement, while rural areas may underreport. As the Zhejiang 12,345-hotline system continues to mature and gain wider adoption, complaint data increasingly reflect actual problem severity. Future research should integrate problem patterns with case volumes to assess governance costs more accurately and to examine the spatiotemporal severity of urban problems.

Second, authority responses to complaints offer valuable insights into the formation and resolution of urban problems. However, due to space constraints, this study did not examine these responses. Future research should systematically summarize and categorize them as key indicators of governance capacity, and assess their impact on urban problem patterns. This would offer a deeper understanding of governance effectiveness and inform strategies for enhancing government performance.

6. Conclusion

This study demonstrates the value of large-scale public complaint data for understanding urban problems by analyzing their patterns, evolutions, and drivers. Using 3.5 million complaint records from 87 county-level units in Zhejiang Province (2018–2022), we constructed 435 urban problem distributions and grouped them into seven typical patterns using a cluster ensemble approach: (1) pollution and municipal infrastructure, (2) consumer protection, (3) land use, (4) housing purchase, (5) garbage management, (6) financial services, and (7) pandemic control. Each problem pattern was characterized by a set of interrelated representative problems. Fuzzy membership analysis revealed the nuanced temporal dynamics of these patterns: emerging patterns gained prominence while previously dominant ones declined, though the pace of change varied across counties. These results suggest that short-term forecasting of problem patterns is feasible.

We further employed multinomial logit models to examine the socioeconomic drivers of problem patterns. Increases in retail sales, land acquisition area, and housing sales volume significantly raised the relative likelihood of a county falling into Patterns 2 (consumer protection), 3 (land use), and 4 (housing purchase), respectively. Pattern 1 (pollution and municipal infrastructure) showed no short-term causal link to the socioeconomic factors but was aligned with long-term economic development trends. Patterns 5 (garbage problems), 6 (financial service), and 7 (pandemic control) were not significantly associated with the examined factors. Instead, these patterns were driven by major events or shocks.

This study makes three key contributions. First, it introduces a data-driven perspective that complements traditional concept-based and case-focused analyses, thereby enhancing the objectivity and generalizability of urban problem research. Second, it develops a systematic methodological framework for complaint data-driven urban problem analysis, covering distribution calculation, pattern identification, evolution tracking, and factor examination. Third, it highlights three county-level urban problem types—development-driven, trend-driven, and event-driven—that require differentiated governance strategies. Although derived from Zhejiang data, these findings offer theoretical insights into constructing more generalized patterns of urban problems.

CRediT authorship contribution statement

Xiangfu Kong: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Bo Dong:** Writing – original draft, Project administration, Investigation, Conceptualization.

Declaration of competing interest

The authors 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 work was supported by the Young Scientists Fund of the National Natural Science Foundation of China (grant number 52402435).

Appendix A

Table A.1
Classification scheme of urban problems.

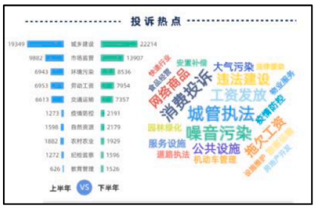
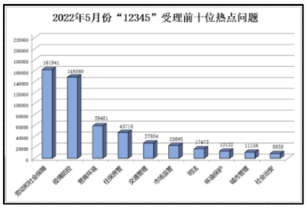
Id	Urban problem	Meaning
1	Air pollution	Pollution of the atmosphere from factory emissions, dust particles, and waste burning.
2	Water pollution	Pollution caused by the discharge of wastewater.
3	Noise pollution	Disruptive or unwanted sound from sources such as construction, traffic, and manufacturing.
4	Construction problems	Environmental problems resulting from construction activities, including road blockages and unsafe buildings.
5	Road damage	Deterioration of road surfaces, posing hazards to drivers, pedestrians, and cyclists.
6	Deteriorated municipal infrastructure	Damaged or outdated municipal facilities, such as streetlights, toilet, parks, sewages, drainage systems, and manhole covers.
7	Poor municipal service	Substandard municipal services, including water and power outages, inadequate landscaping, and road flooding.
8	Poor public bus service	Inefficient bus service, characterized by unreliable schedules, overcrowding, infrequent service, and limited accessibility.
9	Farmland destruction	Deliberate destruction of farmland through sand excavation or soil digging.
10	Unlawful business practices	Business operations that violate laws or regulations, including unlicensed businesses, fraudulent invoicing, and street vendor issues.
11	Land titling problems	Problems related to the legal recognition of land ownership, including unclear land boundaries and lack of formal documentation.
12	Land expropriation problems	Issues caused by land expropriation, such as unfair compensation, involuntary resettlement, and loss of livelihood.
13	Illegal land use	Unauthorized land use, including illegal construction, zoning violations, land grabbing, and unapproved subdivision.
14	Forced demolition	Government-mandated demolitions of illegal structures.
15	Excess traffic enforcement	Overly strict monitoring and penalizing of traffic violations, including excessive fines and over-policing.
16	Inefficient traffic design	Poor road layouts, inadequate signaling, traffic bottlenecks, and lack of traffic signs.
17	Illegal parking	Parking in restricted areas such as driveways, sidewalks, bike lanes, and fire escape routes.
18	Community security risks	Hazardous practices, including illegal scooter parking, storage of flammable materials, and substandard wiring and electrical equipment.
19	Garbage management issues	Littering, overflowing trash bins, and insufficient waste disposal facilities.
20	Property management issues	Poor service quality, water leakage, insufficient parking spaces, elevator malfunctions, and inadequate security measures.
21	NIMBY (Not in my backyard)	Resident opposition to the construction of facilities perceived to have negative effects, such as waste disposal sites.
22	Housing purchase issues	Disputes related to housing transactions, including prepayment properties, bundled parking spaces, and property delivery quality.
23	Car purchase dispute	Disputes related to vehicle purchases, such as bundled sales, forced loans, poor after-sales services, and unclear pricing.
24	Prepaid consumption issues	Consumers who pay for services in advance face difficulties with refunds, unclear terms, and discrepancies in service quality, particularly in retail, fitness, and personal care sectors.
25	Unlawful charging	Unauthorized or inflated charges to consumers, especially in mobile services, taxi, healthcare, and hospitality.
26	Poor customer service	Businesses failing to meet customer expectations in sectors such as express delivery, hotels, banking, driving schools, restaurants, and telecommunications.
27	Poor product quality	Consumers receiving defective or substandard products in online or offline markets.
28	Administrative procedure issues	Problems related to government administration and public service delivery, such as processing documents, permits, and registration.
29	Social welfare issues	Issues in accessing social benefits such as minimum livelihood guarantee (dibao) and subsidies, including eligibility issues, complex documentation, and delays in distribution.
30	Social security issues	Issues related to pension, medical, unemployment, work injury, and maternity insurances and housing funds.
31	Labor dispute	Issues such as wage arrears, employee exploitation, forced resignations, inadequate workplace safety, and unclear contract terms.
32	Healthcare problems	Dissatisfaction with healthcare services, including long waiting times, high medical costs, and strained doctor-patient relations.
33	Education problems	Concerns over education services, such as poor quality, overloaded curriculum, admission qualifications, and cancellations of holidays.
34	Whistleblowing	Complaints by whistleblowers regarding corruption, election manipulation, public official misconduct, failure to file cases, mismanagement of funds, and lack of transparency in budgeting.
35	Financial service issues	Problems in financial services leading to customer losses, particularly in banking, lending, insurance, and other related sectors.
36	Pandemic control problems	Disruptions caused by strict pandemic control measures, including extended lockdowns, quarantine protocols, vaccination campaigns, mask mandates, mass testing, and health code systems.
37	Other problems	Unclassified and minor issues.

Appendix B

Several cities, including Ningbo, Lishui, Hangzhou, and Taizhou, released aggregated complaint counts or proportions of common urban issues for specific periods, providing valuable references for validating the representativeness of our dataset. Table B.1 presents the data sources, snapshots, and corresponding English translations to support the analysis.

Table B.1
Support data sources to analyze the representativeness of our subset.

<table>
<tr><th>City (Time)</th><th>Data sources and Snapshot</th><th>Corresponding English translation</th></tr>
<tr><td>Ningbo (2021)</td><td>https://mp.weixin.qq.com/s/FVM-wBZcJ4Ha1uLs4kbx0Q

2021年，宁波市民在社会生活中最关心哪些民生诉求，向12345热线反映的热点都有哪些？
近日，记者从宁波市12345政务服务热线中心获悉，2021年12345热线共收到市民来电215.79万件，同比上升30.15%；实现100%按期受理，群众满意度97.63%。
分析12345热线大数据，2021年全市群众关于城乡建设、市场监管、交通运输领域的来电诉求相对集中，占比位列前三。
其中，消费纠纷、物业服务、电商交易类问题高发频发，分别占比10.61%、6.48%和4.97%。
<table><tr><th colspan="3">2021 年 12345 热线办理热点统计表</th></tr><tr><td>名次</td><td>领域</td><td>占比</td></tr><tr><td>1</td><td>城乡建设</td><td>27.06%</td></tr><tr><td>2</td><td>市场监管</td><td>22.96%</td></tr><tr><td>3</td><td>交通运输</td><td>8.82%</td></tr><tr><td>4</td><td>政法</td><td>8.68%</td></tr><tr><td>5</td><td>生态环境</td><td>6.41%</td></tr></table></td><td>According to the Ningbo Municipal 12345 Government Service Hotline Center, the hotline platform received 2.16 million calls in 2021, marking a 30.15% year-on-year increase. All cases were processed on time, with a public satisfaction rate of 97.63%. Analysis of the hotline’s data indicates that the most frequently reported concerns were related to urban and rural development, market regulation, and transportation.
<table><tr><th colspan="3">2021 12345 Hotline Top Issues</th></tr><tr><td>Rank</td><td>Problem</td><td>Proportion</td></tr><tr><td>1</td><td>Urban and rural development</td><td>27.06%</td></tr><tr><td>2</td><td>Market regulation</td><td>22.96%</td></tr><tr><td>3</td><td>Transportation</td><td>8.82%</td></tr><tr><td>4</td><td>Political and legal affairs</td><td>8.68%</td></tr><tr><td>5</td><td>Ecological environment</td><td>6.41%</td></tr></table></td></tr>
<tr><td>Lishui (2021)</td><td>https://mp.weixin.qq.com/s/jnpntGhlZyJbn3UDM6CxKw

</td><td>Hotline Top Issues
<table><tr><th>Problem</th><th>Volume (January–June)</th><th>Volume (July–December</th></tr><tr><td>Urban and rural development</td><td>19349</td><td>22214</td></tr><tr><td>Market regulation</td><td>9882</td><td>13907</td></tr><tr><td>Pollution</td><td>6943</td><td>8536</td></tr><tr><td>Labor dispute</td><td>6953</td><td>7954</td></tr><tr><td>Transportation</td><td>6613</td><td>7357</td></tr><tr><td>Pandemic control</td><td>1273</td><td>2191</td></tr><tr><td>Natural resources</td><td>1598</td><td>2179</td></tr><tr><td>Rural affairs</td><td>1882</td><td>1929</td></tr><tr><td>Whistleblowing</td><td>1272</td><td>1596</td></tr><tr><td>Education</td><td>626</td><td>1526</td></tr></table></td></tr>
<tr><td>Hangzhou (May 2022)</td><td>https://www.hzxf12345.gov.cn/12345/content/2022-06/07/8a7fbb82810f2a1b01813d618df633f8.html

</td><td>Hotline Top Issues in May 2022
<table><tr><th>Problem</th><th>Volume</th></tr><tr><td>Labor and social welfare</td><td>161941</td></tr><tr><td>Pandemic control</td><td>148560</td></tr><tr><td>Business environment</td><td>59481</td></tr><tr><td>Housing problems</td><td>43718</td></tr><tr><td>Traffic regulation</td><td>27934</td></tr><tr><td>Market regulation</td><td>23695</td></tr><tr><td>Judicial issues</td><td>17472</td></tr><tr><td>Environmental protection</td><td>13132</td></tr><tr><td>Urban management</td><td>11136</td></tr><tr><td>Social order</td><td>8858</td></tr></table></td></tr>
<tr><td>Taizhou</td><td>https://mp.weixin.qq.com/s/EsNxFCpQXQFm</td><td>Based on the content of citizen calls, the top six problems were</td></tr>
</table>

(continued on next page)

Table B.1 (*continued*)

(November 2022)	uCI1SrTnCQ 	pandemic control (40,313 calls), labor and social welfare (15,661), market regulation (9,591), judicial issues (5,092), housing management (4,936), and traffic management (4,098), accounting for 37.89%, 14.72%, 9.01%, 4.79%, 4.64%, and 3.85% of all cases, respectively.

Appendix C

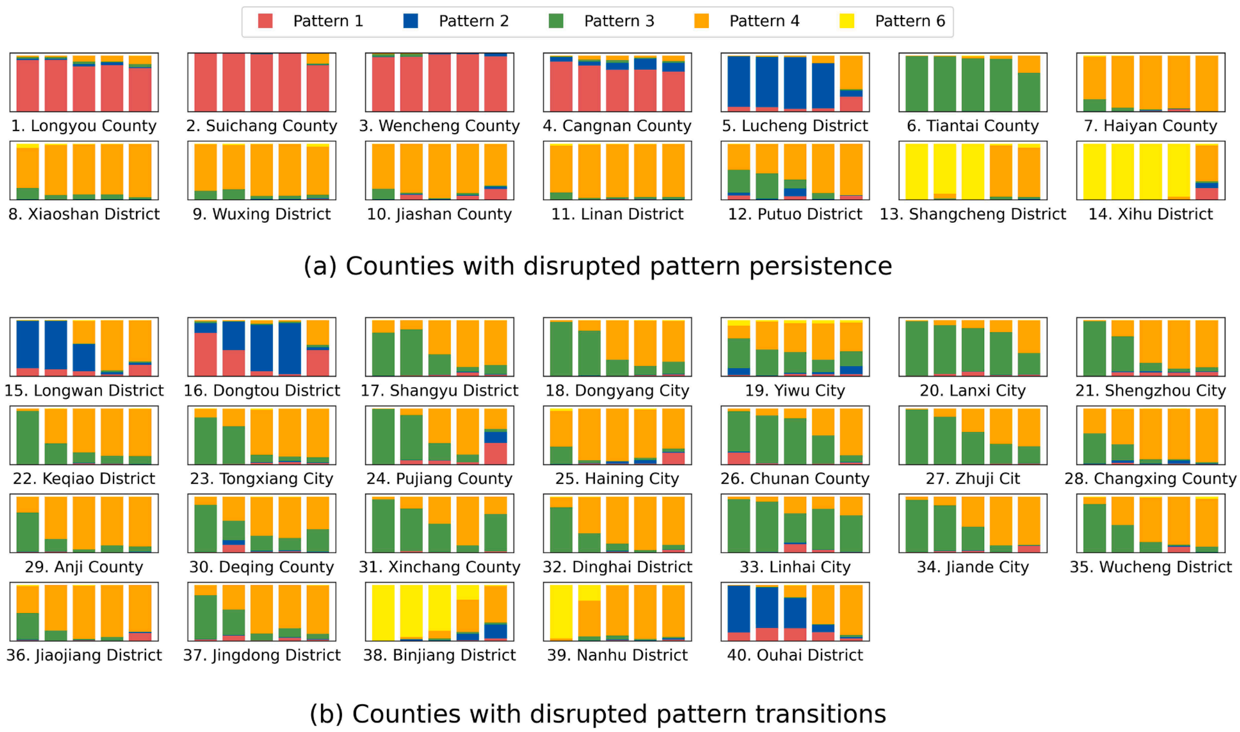


Fig. C.1. County-level temporal evolution of pattern membership (2018–2022): (a) disrupted pattern persistence and (b) disrupted pattern transitions.

Note: The urban problems garbage management and pandemic control are removed from the urban problem distributions.

Data availability

Data will be made available on request.

References

Bastani, K., Namavari, H., & Shaffer, J. (2019). Latent dirichlet allocation (LDA) for topic modeling of the CFPB consumer complaints. *Expert Systems with Applications, 127*, 256–271. https://doi.org/10.1016/j.eswa.2019.03.001

Boongoen, T., & Iam-On, N. (2018). Cluster ensembles: A survey of approaches with recent extensions and applications. *Computer Science Review, 28*, 1–25. https://doi.org/10.1016/j.cosrev.2018.01.003

Brewer, B. (2007). Citizen or customer? Complaints handling in the public sector. *International Review of Administrative Sciences, 4*(73), 549–556. https://doi.org/10.1177/0020852307083457

Chatfield, A. T., & Reddick, C. G. (2018). Customer agility and responsiveness through big data analytics for public value creation: A case study of Houston 311 on-demand services. *Government Information Quarterly, 35*(2), 336–347. https://doi.org/10.1016/j.giq.2017.11.002

Cui, E., Tao, R., Warner, T. J., & Yang, D. L. (2015). How do land takings affect political trust in rural China? *Political Studies, 63*(1), 91–109. https://doi.org/10.1111/1467-9248.12151

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)* (pp. 4171–4186). https://doi.org/10.18653/v1/N19-1423

Engin, Z., Van Dijk, J., Lan, T., Longley, P. A., Treleaven, P., Batty, M., & Penn, A. (2020). Data-driven urban management: Mapping the landscape. *Journal of Urban Management, 9*(2), 140–150. https://doi.org/10.1016/j.jum.2019.12.001

Feng, X., Wang, C., & Wang, J. (2023). Understanding how the expression of online citizen petitions influences the government responses in China: An empirical study with automatic text analytics. *Information Processing & Management, 60*(3), Article 103330. https://doi.org/10.1016/j.ipm.2023.103330

Fern, X. Z., & Brodley, C. E. (2004). Solving cluster ensemble problems by bipartite graph partitioning. In *Twenty-First International Conference on Machine Learning* (p. 36). https://doi.org/10.1145/1015330.1015414

Gelbrich, K., & Roschk, H. (2011). A meta-analysis of organizational complaint handling and customer responses. *Journal of Service Research, 14*(1), 24–43. https://doi.org/10.1177/1094670510387914

Gong, D., Shang, Z., Su, Y., Yan, A., & Zhang, Q. (2024). Economic impacts of China's zero-COVID policies. *China Economic Review, 83*, Article 102101. https://doi.org/10.1016/j.chieco.2023.102101

Guo, M., Lin, Y., Shyu, R.-J., & Huang, J. (2023). Characterizing environmental pollution with civil complaints and social media data: A case of the Greater Taipei Area. *Journal of Environmental Management, 348*, Article 119310. https://doi.org/10.1016/j.jenvman.2023.119310

HaCohen-Kerner, Y., Dilmon, R., Hone, M., & Ben-Basan, M. A. (2019). Automatic classification of complaint letters according to service provider categories. *Information Processing & Management, 56*(6), Article 102102. https://doi.org/10.1016/j.ipm.2019.102102

He, J., Zhang, W., & Yang, M. (2024). The spatial and temporal characteristics of urban public safety under the residents' complaints: Evidence from 12345 data in Beijing, China. *Journal of Urban Management, 13*(2), 217–231. https://doi.org/10.1016/j.jum.2024.01.003

Hu, E. J., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., Wang, L., & Chen, W. (2022). LoRA: Low-rank adaptation of large language models. In *The Tenth International Conference on Learning Representations*. https://openreview.net/forum?id=nZeVKeeFYf9.

Huang, L., & Pontell, H. N. (2023). Crime and crisis in China's P2P online lending market: A comparative analysis of fraud. *Crime, Law and Social Change, 79*(4), 369–393. https://doi.org/10.1007/s10611-022-10053-y

Huang, W., Luo, M., Ta, Y., & Wang, B. (2024). Land expropriation, household behaviors, and health outcomes: Evidence from China. *Journal of Development Economics, 171*, Article 103358. https://doi.org/10.1016/j.jdeveco.2024.103358

Jiao, J., Jin, Y., & Yang, R. (2024). An approach to exploring the spatial distribution and influencing factors of urban problems based on land use types. *Sustainable Cities and Society, 104*, Article 105321. https://doi.org/10.1016/j.scs.2024.105321

Jiao, Y., Li, C., Yao, Z., Weng, C., Lian, A., & Dong, R. (2024). How can online citizen complaints provide solutions to refine environmental management: A spatio-temporal perspective. *Information Processing & Management, 61*(2), Article 103611. https://doi.org/10.1016/j.ipm.2023.103611

Lin, H., Fang, P., Zhou, L., & Xu, L. (2019). A relational view of self-protection amongst China's food safety crises. *Canadian Journal of Development Studies /Revue Canadienne D'Études Du Développement, 40*(1), 131–142. https://doi.org/10.1080/02255189.2019.1525530

Liu, J., Long, R., Chen, H., Wu, M., Ma, W., & Li, Q. (2024). Topic-sentiment analysis of citizen environmental complaints in China: Using a Stacking-BERT model. *Journal of Environmental Management, 371*, Article 123112. https://doi.org/10.1016/j.jenvman.2024.123112

Liu, Q., Wang, S., Zhang, W., Li, J., Zhao, Y., & Li, W. (2017). China's municipal public infrastructure: Estimating construction levels and investment efficiency using the entropy method and a DEA model. *Habitat International, 64*, 59–70. https://doi.org/10.1016/j.habitatint.2017.04.010

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., & Stoyanov, V. (2019). RoBERTa: A robustly optimized BERT pretraining approach. https://openreview.net/forum?id=SyxS0T4tvS.

Ma, B., Zhang, N., Liu, G., Li, L., & Yuan, H. (2016). Semantic search for public opinions on urban affairs: A probabilistic topic modeling-based approach. *Information Processing & Management, 52*(3), 430–445. https://doi.org/10.1016/j.ipm.2015.10.004

Münzel, T., Sørensen, M., & Daiber, A. (2021). Transportation noise pollution and cardiovascular disease. *Nature Reviews Cardiology, 18*(9), 619–636. https://doi.org/10.1038/s41569-021-00532-5

Nobile, F., Forastiere, A., Michelozzi, P., Forastiere, F., & Stafoggia, M. (2023). Long-term exposure to air pollution and incidence of mental disorders. A large longitudinal cohort study of adults within an urban area. *Environment International, 181*, Article 108302. https://doi.org/10.1016/j.envint.2023.108302

Orsingher, C., Valentini, S., & De Angelis, M. (2010). A meta-analysis of satisfaction with complaint handling in services. *Journal of the Academy of Marketing Science, 38*(2), 169–186. https://doi.org/10.1007/s11747-009-0155-z

Pan, J., & Chen, K. (2018). Concealing corruption: How Chinese officials distort upward reporting of online grievances. *American Political Science Review, 112*(3), 602–620. https://doi.org/10.1017/S0003055418000205

Pei, X., Tandon, A., Alldrick, A., Giorgi, L., Huang, W., & Yang, R. (2011). The China melamine milk scandal and its implications for food safety regulation. *Food Policy, 36*(3), 412–420. https://doi.org/10.1016/j.foodpol.2011.03.008

Peng, S., Liu, R., Sun, Y., Zhang, J., & Mao, Y. (2024). Spatial characteristics and influencing factors of people's livelihood issues based on urban nnline governance platforms: A case of Chengdu, China. *Journal of Urban Planning and Development, 150*(4), Article 05024033. https://doi.org/10.1061/JUPDDM.UPENG-4960

Peng, X., Li, Y., Si, Y., Xu, L., Liu, X., Li, D., & Liu, Y. (2022). A social sensing approach for everyday urban problem-handling with the 12345-complaint hotline data. *Computers, Environment and Urban Systems, 94*, 101790. https://doi.org/10.1016/j.compenvurbsys.2022.101790

Sha, W. (2023). The political impacts of land expropriation in China. *Journal of Development Economics, 160*, Article 102985. https://doi.org/10.1016/j.jdeveco.2022.102985

Sun, X. (2015). Selective enforcement of land regulations: Why large-scale violators succeed. *The China Journal, 74*, 66–90. https://doi.org/10.1086/681938

Sun, Y., Jin, F., Zheng, Y., Ji, M., & Wang, H. (2021). A new indicator to assess public perception of air pollution based on complaint data. *Applied Sciences, 11*(4), 1894. https://doi.org/10.3390/app11041894

Wang, F., Wang, M., & Yin, H. (2022). Can campaign-style enforcement work: When and how? Evidence from straw burning control in China. *Governance, 35*(2), 545–564. https://doi.org/10.1111/gove.12571

Wang, H. H., Cheng, E. W., Chen, X., & Liang, H. (2025). How institutionalized feedback works: Online citizen complaints and local government responsiveness in China. *Governance, 38*(2), Article e12907. https://doi.org/10.1111/gove.12907

Wang, X., Zhu, Y., Zeng, H., Cheng, Q., Zhao, X., Xu, H., & Zhou, T. (2022). Spatialized analysis of air pollution complaints in Beijing using the BERT+CRF model. *Atmosphere, 13*(7), 1023. https://doi.org/10.3390/atmos13071023

Wei, J., Bosma, M., Zhao, V., Guu, K., Yu, A. W., Lester, B., Du, N., Dai, A. M., & Le, Q. V. (2021). Finetuned language models are zero-shot learners. *International Conference on Learning Representations*. https://openreview.net/forum?id=gEZrGCozdqR.

Wu, X., Yang, D. L., & Chen, L. (2017). The politics of quality-of-life issues: Food safety and political trust in China. *Journal of Contemporary China, 26*(106), 601–615. https://doi.org/10.1080/10670564.2017.1274827

Xie, T., & Yuan, Y. (2023). Go with the wind: Spatial impacts of environmental regulations on economic activities in China. *Journal of Development Economics, 164*, Article 103139. https://doi.org/10.1016/j.jdeveco.2023.103139

Xie, Z., Weng, W., Pan, Y., Du, Z., Li, X., & Duan, Y. (2023). Public opinion changing patterns under the double-hazard scenario of natural disaster and public health event. *Information Processing & Management, 60*(3), 103287. https://doi.org/10.1016/j.ipm.2023.103287

Xu, X., Nie, S., Ding, H., & Hou, F. F. (2018). Environmental pollution and kidney diseases. *Nature Reviews Nephrology, 14*(5), 313–324. https://doi.org/10.1038/nrneph.2018.11

Yang, A., Yang, B., & Zhang, B. (2024). *Qwen2.5 technical report*. https://huggingface.co/papers/2412.15115.

Yang, Z., Dai, Z., Yang, Y., Carbonell, J., Salakhutdinov, R. R., & Le, Q. V. (2019). XLNet: Generalized autoregressive pretraining for language understanding. *Advances in neural information processing systems* (p. 32). https://papers.neurips.cc/paper_files/paper/2019/hash/dc6a7e655d7e5840e66733e9ee67cc69-Abstract.html.

Yi, D., Wu, D., Gan, L., & Clark, W. A. V. (2024). Does urban housing demolition increase wealth inequality in China? *Cities, 150*, Article 105053. https://doi.org/10.1016/j.cities.2024.105053

Zhai, Y. (2023). The politics of COVID-19: The political logic of China's zero-COVID policy. *Journal of Contemporary Asia,, 53*(5), 869–886. https://doi.org/10.1080/00472336.2023.2194322

Zhang, J., Geng, Q., & Jin, J. (2025). EKLI-Attention: An integrated attention mechanism for classifying citizen requests in government−citizen interactions. *Information Processing & Management, 62*(6), Article 104237. https://doi.org/10.1016/j.ipm.2025.104237

Zhang, Y., Liu, N., Li, Y., Long, Y., Baumgartner, J., Adamkiewicz, G., Bhalla, K., Rodriguez, J., & Gemmell, E. (2023). Neighborhood infrastructure-related risk factors and non-communicable diseases: A systematic meta-review. *Environmental Health, 22*(1), 2. https://doi.org/10.1186/s12940-022-00955-8

Zhang, Z., Lin, X., & Shan, S. (2023). Big data-assisted urban governance: An intelligent real-time monitoring and early warning system for public opinion in government hotline. *Future Generation Computer Systems, 144*, 90–104. https://doi.org/10.1016/j.future.2023.03.004

Zhao, H., & Liu, T. (2024). China's social credit system and the family: Punishment and collective resistance. *Economy and Society, 54*(1), 46–69. https://doi.org/10.1080/03085147.2024.2422187

Zhao, X., Jiang, M., & Zhang, W. (2022). The impact of environmental pollution and economic growth on public health: Evidence from China. *Frontiers in Public Health, 10*, 861157. https://doi.org/10.3389/fpubh.2022.861157

Zhao, X., Jin, L., & Sun, S. B. (2022). Gone with the land": Effects of land expropriation on health and subjective well-being in rural China. *Health & Place, 73*, Article 102614. https://doi.org/10.1016/j.healthplace.2021.102614

Zhao, X., & Xie, Y. (2022). The effect of land expropriation on local political trust in China. *Land Use Policy, 114*, Article 105966. https://doi.org/10.1016/j.landusepol.2021.105966

Zhejiang Provincial Bureau of Statistics. (2010). *Zhejiang statistical Yearbook*.

Madyatmadja, E. D., Sianipar, C. P. M., Wijaya, C., & Sembiring, D. J. M. (2023). Classifying crowdsourced citizen complaints through data mining: Accuracy testing of k-nearest neighbors, random forest, support vector machine, and AdaBoost. *Informatics, 10*(4), 84. https://doi.org/10.3390/informatics10040084