Public Participation in Smart-City Governance: A Qualitative Content Analysis of Public Comments in Urban China

Zhe Gao 1,2, Siqin Wang 3,* and Jiang Gu 1,2

1 Hubei Provincial Key Laboratory for Geographical Process Analysis & Simulation, Central China Normal University, Wuhan 430079, China; gaozhe@mail.ccnu.edu.cn (Z.G.); gujiang@mail.ccnu.edu.cn (J.G.)
2 The College of Urban & Environmental Sciences, Central China Normal University, Wuhan 430079, China
3 School of Earth and Environmental Sciences, University of Queensland, St Lucia 4067, Australia
* Correspondence: s.wang6@uq.edu.au

Received: 7 September 2020; Accepted: 15 October 2020; Published: 17 October 2020

Abstract: Public participation is crucial in the process of urban governance in smart-city initiatives to enable urban planners and policy makers to take account of the real public needs. Our study aims to develop an analytical framework using citizen-centred qualitative data to analyse urban problems and identify the areas most needed for urban governance. Taking a Chinese megacity as the study area, we first utilise a web-crawling tool to retrieve public comments from an online comment board and employ the Baidu Application Programming Interfaces and a qualitative content analysis for data reclassification. We then analyse the urban problems reflected by negative comments in terms of their statistical and spatial distribution, and the associative factors to explain their formation. Our findings show that urban problems are predominantly related to construction and housing, and most frequently appear in industry-oriented areas and newly developed economic development zones on the urban fringe, where the reconciling of government-centred governance and private governance by real estate developers and property management companies is most needed. Areas with higher land prices and a higher proportion of aged population tend to have fewer urban problems, while various types of civil facilities affect the prevalence of urban problems differently.

Keywords: urban governance; public participation; public comments; web-crawling data; qualitative content analysis; urban China

1. Introduction

Public participation in the process of urban planning and governance has been unprecedentedly encouraged and promoted in the context of smart-city initiatives [1]. The essentials of building a cyber, digital, and wired smart city are the modern information and communications technology (ICT) infrastructures (i.e., Internet of Things and cloud computing). ICT-based urban systems provoke the new concept of e-governance comprising ICTs to support democratic processes, enhance services delivery and effective resource management for urban administration to citizens [2], and e-participation involving technology-mediated interaction among civil society, formal politics, and the public administration. Most of the studies on e-governance and e-participation have been developed in the Western context, largely relating to advanced technologies, frameworks, models, and methods to support decision-making in terms of urban performance, governance, and policy [3–6]. However, such studies in the Eastern context are underexplored and need to be extended to developing countries with a large population base, such as China, for a holistic understanding of smart-city initiatives.
The increasing prevalence of mobile devices and emergence of cloud technologies generate a large volume of quantitative and qualitative information created by different agents (i.e., governments, organisations, and individuals) and shared through social media and online platforms. Internet-centred big data is increasingly attractive as a data source for urban analysis to provide new insights for evaluating public participation in urban governance [7]. In the current scholarship, most researchers focus on macro-level policy- and decision-making based on analysing collective-level big data produced by civil collective action in urban affairs such as transportation, emergency management, health care, or environment management [8–10]. However, citizen-centred big data at the individual level that provide temporal–spatial insights into citizen behaviour and individual opinions gain relatively scarce attention, covered merely by a few recent studies in the Chinese literature [11,12]. There is a pressing need in the current research agenda to incorporate opinions and comments of citizens in urban governance to support decision-making based on the real public needs [13].

To fulfil the aforementioned knowledge gaps, our study contributes to an emerging research agenda of e-governance facilitating the formation of smart cities, by constructing an analytical framework to retrieve qualitative big data to reflect urban problems and to examine the spatial distribution and associative factors of urban problems to guide us through citizen-directed participatory practices in urban governance in a Chinese context. Taking Wuhan City, the largest capital city in Central China, as our study area, we aim to address the following research questions: (1) What is the spatial distribution of urban problems? (2) What is the spatial relationship between urban problems and surrounding areas? and (3) To what extent are urban problems explained by factors related to the demographic and socioeconomic profiles of communities at the micro level and civil facility provision at the macro level? To answer these questions, we commence with retrieving public comments via an online comment board by a web-crawling tool and employ the Baidu Application Programming Interfaces and a qualitative content analysis for data reclassification. We then analyse the urban problems reflected by negative comments in terms of their statistical and spatial distribution, and the associative factors to explain their formation. By doing so, we identify the areas in most need of urban governance and provide the political implications of public participation and e-governance in the smart-city initiatives, which are fundamentally rooted in the political–economic institutions characterised in the Chinese context.

The rest of this paper is organised in the following way. The next section reviews e-governance and e-participation in smart cities, and performs content analysis by using big data and the rationale of urban governance in China. Then an analytical framework for retrieving and processing data is introduced, together with the description of data sources and methodologies. This is followed by results covering the statistical and spatial distribution of urban problems identified by the negative public comments, the spatial relationship between urban problems and their surrounding areas, and the associative factors that explain urban problems. The final section presents a number of conclusions and discussion of our analytical framework to analyse a wider range of qualitative data.

2. Literature Review

2.1. E-Governance and E-Participation in Smart Cities

Smart cities encompassing emerging technologies and urban sustainability have become an increasingly popular concept advocated by governments in countries worldwide [14]. A smart city is defined as a public–private urban system of governments, industries, organisations, and citizens centred around sustainable economic growth and quality of life through participatory urban governance [15]. The essentials of building a cyber, digital, wired, informational or knowledge-based smart city are the modern information and communications technology (ICT) infrastructures (i.e., Internet of Things and cloud computing), which not only change the way people live and work but also foster innovation and knowledge creation [16]. What accompanies a rising smart city is the smart governance or e-governance. The concept of e-governance has had a major evolution in the last 15 years, changing from an element of political and administrative authority to a comprehensive political tool in which governments are able
to manage the process of urban development, addressing social, environmental, and economic issues through an integrated and planned long-term strategic vision [17]. Combining the different aspects that characterise the smart city (e.g., socioeconomic factors, environment, population, and living style), the city becomes able to know, to understand and to govern problems through sophisticated ICTs [17]. ICT-based e-governance relies heavily on a wealth of temporal–spatial big data about the urban environment collected through ubiquitous digital devices and platforms, and contributed by various agents in urban systems (i.e., governments, industries, non-profit organisations, private sectors, and citizens). As such, e-governance inevitably involves citizen participation and collaboration among stakeholders, provoking the concept of e-participation in the implementation of smart-city initiatives. [6,18].

Public participation in urban governance originally referred to a governing arrangement where one or more agencies directly engaged non-state stakeholders in a collective decision-making process in order to achieve a consensus-oriented and deliberative goal of collaborative governance [19]. Enabled by ICTs, e-participation or e-governance participation involves technology-mediated interaction among civil society, formal politics, and the public administration. E-participation has been implemented in different countries as part of e-governance for policy-relevant issues such as communicating information on legal issues and negotiating the benefits of different agents involved in the urban planning process [20]. The popularity of digital devices (i.e., smart phones and tablets) facilitates the prevalence of e-participation, through which citizens interact with people living in the same neighbourhood or having similar interests to form online communities and express their opinions on urban environment and policies. Engaging in an online environment can be especially valuable for people with disabilities, as restricted physical mobility or other life circumstances might make it more difficult for them to meet others offline. Such online public comments and options, mostly anonymous, can be easily supported and followed by other Internet users and are normally characterised by a highly sentimental tendency. It is worth noting that e-participation inevitably introduces some concerns [2,6,17–20], including privacy issues, legitimacy problems, information transparency, the dissemination of disoriented ideas, and the under-representativeness of participants (e.g., limited to Internet or smart phone users). However, governments and the administrators of Internet companies have given increasing attention and a considerable amount of effort to eliminate privacy and legitimacy issues, and regulate the online environment for e-participation; the large volume of user-based data can greatly offset the possible stratification of data sampling and maximise data representativeness [20]. Despite these concerns, e-participation has become an important approach and initiative in the past decade to facilitate public participation [21]. Studies on e-governance and e-participation are relatively more popular in the Western context and largely divided into two streams—technological and managerial [21]. The former exploits advanced technologies to enhance e-governance and e-participation of citizens [2–4], while the later aims to design conceptual frameworks, models, and methods to support decision-making in terms of urban performance, governance, and policy on the basis of experience [6,16,22]. However, such studies beyond the Western context in non-English literature are very under-explored and need to be extended to developing countries with large population bases, such as China, for a holistic understanding of smart-city initiatives.

2.2. Content Analysis in the Era of Big Data

Big data collection and analysis play an indispensable role in smart-city initiatives to enhance the reliability and efficiency of e-governance and e-participation, and, in turn, to strengthen the intelligence of ICT-based urban systems [23]. The increasing prevalence of mobile devices and the emergence of clouding technologies have contributed to the so-called ‘data revolution’ [24]. As a large volume of quantitative and qualitative information created by different agents (i.e., governments, organisations, and individuals) and shared through social media and online platforms becomes relatively easy to capture and analyse in real time, Internet-centred big data is increasingly attractive as a data source for urban analysis to provide new insights for evaluating public participation in urban governance [7]. Different from traditional approaches stimulating public participation in policy
decisions through suggestions and indications that follow a typical top-down decision-making process, ICTs, and especially social media and online platforms, are used to widen the number of participants in the public debate and to give a voice to individuals who are not usually willing or able to participate in person in public debates, such as disabled people. Online forums, comment boards, communicative platforms developed by governments at different levels engage citizens through a mix of online initiatives that collect their points of view, opinions, comments, criticisms, and suggestions on urban environment. In the current scholarship using Internet-centred big data, most researchers focus on macro-level policy- and decision-making based on analysing collective-level big data produced by collective civil action in urban affairs such as transportation, emergency management, health care, or environmental management [8–10]. However, citizen-centred big data at the individual level that provides temporal–spatial insights into citizen behaviour and individual opinion has significant potential value but gains relatively scarce attention. Although a few studies along the aforementioned direction have emerged recently [11,12], they are published in the Chinese literature with merely limited access by international readers. Our study aims to extend this research realm with a more robust research design and outcomes.

Analysing cities with citizen-centred big data brings different challenges, including approaches and tools for data retrieving and mining, and methods used to analyse such data. Citizen-centred big data to reflect public opinions are usually text-based and geotagged, retrieved from social media and online platforms by advanced technologies of web data extraction, such as web crawling or harvesting. The unstructured nature of such qualitative data makes it difficult to conduct meaningful research, and thus requires different analytical methodologies compared to those used to deal with traditional census or survey data. Many methods have been applied in studies of online messages and comments, for example, textual analysis, triangulation analysis, discourse analysis, narrative analysis, and content analysis [25]. Among these, content analysis is one of the most commonly used methods to deal with qualitative big data [26]. Qualitative content analysis (QCA) was first used as an analytic technique to examine a wide range of textual materials, including communications, narrative responses, survey questions, interviews, focus groups, observations, political speeches, advertisements, and printed media such as magazine articles, books, or manuals [27]. The purpose of QCA is to classify written or oral materials into identified categories of similar meanings. These categories need be mutually exclusive and exhaustive, and no data should fall between two categories or be placed in more than one category [28]. In this case, researchers need to determine how best to categorise data if they lack a single meaning or interpretation. QCA not only enables researchers to focus on the themes and topics of the categories to investigate the distribution, meaning, and relations across categories, but also allows a great amount of textual data to be reduced to numbers and frequencies that are suitable for statistical analysis [29]. However, QCA has also been criticised as being labour-intensive and time-consuming, and inappropriate for open explorative research. Thanks to the development of advanced computing interfaces such as application programming interfaces (APIs), which allow the interactions between multiple software intermediaries, the reliability and efficiency of QCA has been much improved because APIs can be entirely customised and specific to a certain purpose of data manipulation. For example, a wide range of APIs have been developed to provide extension mechanisms for data conversion and categorisation, retrieving locational information, and identifying sentiment tendency [30–33]. As such, our study primarily develops an analytical framework embedding QCA and APIs to analyse citizen-centred qualitative big data.

2.3. Rationale of Urban Governance and Public Participation in China

In the Western context, public participation in urban governance is often seen as a reflection of democracy to serve various political rationales, including improving policy decisions, realising democratic values, and stimulating policy implementation [13]. In contrast, public participation in the Chinese context is largely government-led, which has evolved with the progress of institutional changes, fundamentally the government’s deepened market reforms relating to land and housing development since the 1990s [34]. Before 1990, socialist work-unit housing was predominantly
provided and governed by the government, and such state provision and governance of housing was ended in 1998 [35]. As alternatives, newer residential development projects, including gated communities and master-planned estates, have proliferated as some of the dominant residential neighbourhood forms in different regions in favour of the free housing market and the private sector [36]. The growing prevalence of gated communities and master-planned estates reflects a shift from the government-initiated governance towards a neoliberal mode of urban governance, emerging with a new phenomenon of private governance within which public participation plays an increasing role [37,38].

Private governance in the Chinese context is defined as providing public services and administration within communities through organisations that are held by private sectors and through joint governance by residents [39]. On the one hand, through capital accumulation and implemented incentives, state and local governments draw real estate developers and property management companies into developing residential land, providing neighbourhood services and facilities, and maintaining space, landscape, and environment within a community [40]. Residents pay property management companies for the provision of neighbourhood services and facilities, and environmental maintenance. In this sense, private governance through a fee-paying system has also been understood as a market-driven process, and such market provision of public goods and services reduces public expenditure on community governance [41]. On the other hand, gated communities have residents who elect the homeowners’ association to run the community by their own rules, to secure the value of private property, to enforce neighbourhood covenants, and to cope with local governance failure [37]. The use of private compacts in these gated estates recasts residents’ everyday practices and reshapes residents’ engagement with the public realm [42]. Changed governance and deepening housing privatisation in China have received much attention with discussion of both advantages and disadvantages in existing housing studies [41,43].

Many studies on the liberalised housing system agree that private governance has salient goodness, including increasing public participation and awareness of community management [44,45], removing governments from the provision of neighbourhood services, supplementing inadequate public services, and boosting property management industries in urban China. However, private governance has also stimulated public debates [46]. On one hand, although homeowners’ associations have gradually emerged for managing neighbourhoods, they lack autonomy and financial independence, and poorly represent the residents because residents do not have sufficient power to determine the system of neighbourhood governance but rather supervise property management companies. On the other hand, property management companies do not guarantee the efficiency of neighbourhood services provision because service delivery is managed in accordance with property management fees, leading to discrepancies of service and administration across communities because a wealthy community has “a supply of civil goods which apparently are better acquired there than elsewhere” in the urban area [47] (p. 153). Also, because of the fee-paying system, property management companies may tend to exclude the public space beyond the boundary of their own communities, which generates grey space filled with nuisances, conflicts, and urban problems that may be neglected by state governments, as well [41]. In this sense, the appearance of local problems in the gap between government-centralised and private governance is subject to both to the characteristics of land use and urban space at the macro level and the socioeconomic profile of communities at the micro level [48]. Thus, with the aforementioned considerations, our study selects factors relating to the demographic and socioeconomic characteristics of communities at the micro level as well as the public service configuration at the macro level to examine how such factors are associated with urban problems reflected by public comments in urban China.

3. Study Context, Data and Methods

3.1. Study Context

Our study area was Wuhan City, the capital city of Hubei province, located in the central plain of China (Figure 1). Wuhan had a 11.06 million urban residential population and a total area of 8494.41 km²
in 2018, comprising two national-level economic development zones (EDZs) and 13 districts, or a total of 186 jiedaos (sub-districts). The two EDZs are Wuhan EDZ and Donghu EDZ; the 13 districts include seven districts in the main urban area (Jiang’an, Jianghan, Qiaokou, Hanyang, Wuchang, Qingshan, and Hushan) and six new districts located in the outer city (Jiangxia, Huangpi, Xingzhou, Dongxi Hu, and Hannan). Wuhan had a GDP of RMB 1484.73 billion (Chinese Yuan) in 2018, making the city the largest economy in Central China [49]. We used Wuhan for our empirical study for the following reasons. First, as the most important political, economic, transport, educational, and cultural hub in Central China, Wuhan has been defined as the central city at the national level, and thus has profound effects on surrounding provinces. Also, Wuhan represents some common characteristics of hinterland cities in China in relation to its geographic location, industrial structures, and development trajectories. Second, Wuhan has experienced rapid urbanisation and significant economic growth since the early 2000s, alongside the emergence of a series of urban problems caused by industrial transformation, urban regeneration, and land development. Third, Wuhan municipal government has started to establish intelligent urban systems to accelerate the smart-city development, to implement the smart-city initiatives, and to widen the public participation in the process of urban planning and policy making. Such comprehensive systems for urban administration provide rich multidimensional data that are reliable in quality, quantity, and ease of access. Hence, Wuhan serves as a good test case to examine how to use citizen-centred qualitative big data to reflect urban problems and to demonstrate how public participation helps with urban government in the smart-city operation.

Figure 1. Study area.

3.2. Data Collection and Manipulation

The Chinese statistical geographic hierarchy follows the structure of municipal (shi in Chinese), district (qu), sub-district (jiedao), and community (xiao qu), while the jiedao is designed as the basic building block of a city proper and the smallest census unit with the availability of digital boundaries. At the data pre-processing stage, there was a need to aggregate data at a community level to a jiedao level, however the local government of Wuhan City does not provide correspondence data to convert geographic units at different levels. Thus, we constructed a correspondence file for the conversion of spatial units by web crawling the qualitative data from LianJia, one of the most popular real estate websites in China (www.lianjia.com). The correspondence file included the names of communities in the whole municipal area of Wuhan, together with their jiedaos and districts where they are located,
which would be used for data aggregation at the later stage. Next, we collected primary data of public comments from the Wu Han Comment Board (WHCB) (http://liuyan.cjn.cn/) through web crawling. WHCB is an online board where citizens can post their comments, suggestions, and complaints about urban problems and public issues. Utilising Houyi Collector, an open source web-crawling program (http://houyicaiji.com/), we retrieved a total number of 65,295 records of public comments posted on WHCB from 16 November 2017 to 26 June 2019, and matched the name of the jiedaos in these records with the correspondence file we previously created and excluded the unmatched ones. Through data filtering and clearing, the final records we used for the analysis were 12,164. This text-based data set had attributes including title, description, nature, and category. We then re-classified the default categories of records (originally 193 sub-categories) into 22 categories and five broad categories based on the nature of records (Figure 2): (1) safety and security, (2) construction and housing, (3) environment and pollution, (4) infrastructure and configuration, and (5) social service and administration. We further aggregated the data set at the community level to the jiedao level by category, generating the number of negative and non-negative records in each category at the jiedao level. To examine the associative factors of the formation of public comments, we further collected data of land value from the Wuhan land grade and benchmark land price standard [50], data of public service configurations retrieved from Point of Interest [51], and data of aged population above 65 from the 2015 population census [52].

![Figure 2. Data re-categorisation.](image)

### 3.3. Methodology

We developed an analytical framework integrating qualitative content analysis and the Baidu Artificial Intelligence Open Platform (https://ai.baidu.com/), which is embedded by various application programming interfaces (APIs) and functionalities for natural language processing. Through the API of Named Entity Recognition, we commenced extracting the names of communities from the title of records and using the correspondence file generated previously to assign these communities where public issues occurred to the jiedao level. Next, we employed the API of Sentiment Tendency to analyse the sentiment tendency of comment providers based on the description of records, generating a score of sentiment and confidence. Sentiment represented the sentiment tendency of comments as negative, neutral, and positive (coded as 0, 1, and 2, respectively); confidence was a percentage indicating how much a certain public comment belonged to the sentiment tendency defined by sentiment (80% used in our study). The above procedure distinguished 9488 negative records and 2676 non-negative (combining neutral and positive) records, covering 127 jiedaos. Given that the non-negative records only accounted for 22% of the total records and were predominantly related to construction and housing, our ensuing analysis and visualisation focused on the negative comments to reflect urban problems as they involved more diversities and variations across category and space.

We then conducted a descriptive analysis to summarise the number of urban problems and mapped out their spatial distribution at the jiedao level and further examined the spatial relationship between problem-concentrated areas and their surrounding areas by a local index of spatial autocorrelation
(LISA). LISA, in particular, the local Moran’s I, was utilised to capture the spatial distribution of socioeconomic characteristics. LISA has been used widely in the literature to indicate spatial clustering, commonly referred to as hotspot analysis [53]. In GeoDa software, the LISA cluster map is generated to illustrate the distribution of the local Moran’s I statistic across observations [54]. LISA values are colour-coded by type of spatial autocorrelation: the high–high and low–low clusters (positive local spatial autocorrelation) are typically referred to as spatial clusters; while the high–low and low–high clusters (negative local spatial autocorrelation) are spatial outliers. The global Moran’s I statistic indicates the mean of the local Moran’s I statistics, working as an indicator of the spatial autocorrelation across the whole region. In order to obtain more robust results, the number of permutations was set to 999, along with a threshold significance level of $p = 0.05$.

Finally, we employed a pairwise Pearson correlation analysis and then a spatial lag regression (SLR) analysis to examine the association between urban problems and factors relating to the demographic and socioeconomic profile of areas at the micro level and public service provision at the macro level. The bivariate Pearson correlation served as a pre-test of the relationship among the above variables to reduce the effect of the variables’ collinearity on calculation. A Pearson correlation coefficient ranges between $-1$ and $1$; $1$ means total positive correlation, $0$ means no correlation, and $-1$ mean total negative correlation [55]. SLR is a linear spatial autoregressive method with the advantage of capturing spatial dependency in regression analysis, avoiding statistical problems such as unstable parameters and unreliable significance test, as well as providing information on spatial relationships among the independent variables [56]. In GeoDa, SLR sets up the first-order rook’s move contiguity (that is, adjacent edges) and uses the diagnostics from GeoDa to determine the most appropriate weights matrix [54]. The spatial lag model can be specified by [57]:

$$ Y = \alpha W_y + \beta X + \varphi, $$

$$ \varphi = (1 - W_y)^{-1} \epsilon, $$

where $Y$ is the spatially lagged dependent variable (the number of negative comments normalised by the number of population at the jediao level in our study); $\alpha$ is the spatial autoregressive structure of the disturbance $\varphi$; $W_y$ represents the spatial weight matrix of the dependent variable $Y$; $X$ is a matrix containing a set of independent variables (the average land price, the number of aged population and public facilities by type at the jiedao level in our study); $\beta$ is the coefficient of $X$; and $\epsilon$ is the normally distributed error. The overall analytical workflow is presented in Figure 3.

![Figure 3. Analytical workflow.](image-url)
4. Results

4.1. Statistical and Spatial Distribution of Urban Problems

There were 124 jiedaos (97.6% of the jiedaos with public comments) observed with urban problems reflected by negative comments. It was interesting to notice that the top 10 jiedaos with the highest number of urban problems appeared in industry-oriented and newly developed areas beyond the 3rd transport circle (Figure 4), spreading out to the remote region on the urban fringe. The top 10 jiedaos included several economic development zones (EDZs) (e.g., Pan Long Cheng, Yang Luo, Miao Shan) and industrial parks (e.g., Wuhan Iron and Steel Company (WISC) and Yaojiashan). Such areas have relatively large in- and out-flows of migrants and temporary residents, and thus were observed to have more problems in a pressing need for urban governance compared to the well-developed areas in and around the inner city. We further analysed the proportion of urban problems by category in Figure 5 with the X axis showing 124 jiedaos ordered by a descending number of urban problems. In most of the jiedaos, construction and housing accounted for the predominant proportion of urban problems, followed by infrastructure and configuration, and social service and administration. Safety and security issues tended to appear in a high proportion in jiedaos with a relatively smaller number of urban problems (towards the right of Figure 5).

We further mapped the spatial distribution of urban problems by category (Figure 6). Overall, the red-spots of urban problems in each category all tended to be sporadically distributed outside of the 3rd transport circle, in the newly developed region and remote districts including Qingshan, Caidian, Jiangxia, Hanyang, and Xingzhou Districts. When urban problems were broken down into each category, the most obvious prevalence of safety and security issues, and environmental pollution issues was observed in old industrial jiedaos in Qingshan District (e.g., Changqian, worker villages, WISC industrial park) and some of the EDZs (e.g., Yangluo, Yaojiashan, and Panlongcheng). Furthermore, the spatial pattern of infrastructure and configuration was similar to that of social service and administration, and red-spots of both categories were distributed outside of the 3rd transport circle, mostly in newly developed EDZs (e.g., Yangluo, Panlongcheng, Yaojiashan, and Miaoshan) as well as in the WISC industrial park. However, different from the aforementioned four categories of urban
problems distributed more towards remote districts on the urban fringe, construction and housing issues spread out in the inner-city districts with a large number of historical jiedaos (e.g., Jiang’an, Jianghan, and Qingshan District), where urban regeneration and gentrification occurred.

4.2. Spatial Relationship of Urban Problems

The result of LISA showed the spatial relationships of urban problems and their surrounding areas, where our main interests focused on the clusters of urban problems (high–high red-spots in...
Figure 7) indicating the concentrations of adjacent jiedaos with a high prevalence of urban problems. Globally, the magnitude of global Moran’s I for all types of urban problems was small, indicating that the distribution of urban problems was dispersed in the whole study area. However, several local clusters were observed around three EDZs (e.g., Miaoshan, Yangluo, Yaojiashan), indicating a local concentration of urban problems across each EDZ and its surrounding areas. When urban problems were divided into each category, with the exception of safety and security, which was most obviously clustered in Wangji in Xingzhou District, all other categories of urban problems were highly concentrated in several EDZs. For instance, Yaojiashan EDZ was a high concentration of construction and housing, environment and pollution, and social service and administration; Miaoshan EDZ was a high concentration of construction and housing, and infrastructure and configuration; WISC industrial park was a high cluster of construction and housing, environment and pollution, and infrastructure and configuration. In sum, the patterns of these local clusters indicated that urban problems were highly prevalent in old industrial areas and newly developed EDZs, where urban problems may spill over and have chain effects on surrounding areas to form local clusters most in need of urban governance endorsed by governments or related sectors.

Figure 7. Local clusters of urban problems by category.

4.3. Associative Factors of Urban Problems to Guide Urban Governance

We first conducted a pairwise correlation analysis among all measures in the later regression model to test data collinearity, including eight independent variables (civil facilities contained five categories) and six dependent variables representing all urban problems and each category of urban problems (Table 1). The result showed that both land price and age above 65 were significantly \((p < 0.01)\) correlated to urban problems as a whole, Category 2 and Category 4. In addition, five types of civil facilities were also significantly \((p < 0.01)\) correlated to urban problems as a whole, Category 2 and Category 4. The extent that land price, age above 65, and civil facilities affected urban problems needed further exploration by the ensuing SLR model.
Table 1. Coefficients of the pairwise Pearson correlation between urban problems and selected associative factors.

<table>
<thead>
<tr>
<th>No.</th>
<th>Variable</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Land price</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Age above 65</td>
<td>0.195</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Civil facilities</td>
<td>0.163</td>
<td>-0.159</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Traffic</td>
<td>-0.002</td>
<td>-0.143</td>
<td>0.936</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Commercial</td>
<td>0.173</td>
<td>-0.162</td>
<td>0.999</td>
<td>0.924</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Public</td>
<td>0.098</td>
<td>-0.047</td>
<td>0.925</td>
<td>0.950</td>
<td>0.913</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Educational</td>
<td>0.043</td>
<td>-0.259</td>
<td>0.583</td>
<td>0.608</td>
<td>0.567</td>
<td>0.603</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Medical</td>
<td>0.307</td>
<td>-0.022</td>
<td>0.910</td>
<td>0.834</td>
<td>0.907</td>
<td>0.848</td>
<td>0.514</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Urban problems</td>
<td>-0.298</td>
<td>-0.323</td>
<td>0.372</td>
<td>0.478</td>
<td>0.361</td>
<td>0.416</td>
<td>0.258</td>
<td>0.279</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Category1</td>
<td>-0.017</td>
<td>-0.024</td>
<td>-0.036</td>
<td>-0.046</td>
<td>-0.037</td>
<td>0.022</td>
<td>-0.026</td>
<td>0.006</td>
<td>0.232</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Category2</td>
<td>-0.283</td>
<td>-0.320</td>
<td>0.392</td>
<td>0.509</td>
<td>0.380</td>
<td>0.439</td>
<td>0.280</td>
<td>0.302</td>
<td>0.965</td>
<td>0.120</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Category3</td>
<td>-0.143</td>
<td>-0.171</td>
<td>-0.038</td>
<td>-0.030</td>
<td>-0.037</td>
<td>-0.026</td>
<td>-0.096</td>
<td>-0.040</td>
<td>0.586</td>
<td>0.512</td>
<td>0.440</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Category4</td>
<td>-0.288</td>
<td>-0.255</td>
<td>0.411</td>
<td>0.497</td>
<td>0.402</td>
<td>0.443</td>
<td>0.298</td>
<td>0.300</td>
<td>0.852</td>
<td>0.239</td>
<td>0.717</td>
<td>0.488</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Category5</td>
<td>-0.163</td>
<td>-0.218</td>
<td>-0.054</td>
<td>-0.030</td>
<td>-0.054</td>
<td>-0.035</td>
<td>-0.097</td>
<td>-0.067</td>
<td>0.664</td>
<td>0.360</td>
<td>0.515</td>
<td>0.845</td>
<td>0.610</td>
<td>1</td>
</tr>
</tbody>
</table>

Note: Bold: $p < 0.01$; Category 1: Safety and security; Category 2: Construction and housing; Category 3: Environment and pollution; Category 4: Infrastructure and configuration; Category 5: Social service and administration.
The model performance of SLR was reasonably good, with the R-square ranging from 0.292 to 0.439, and the joint significance of independent variables was above the level of 99%, indicating that independent variables were significantly ($p < 0.01$) associated with the dependent variable. When it broke down to each independent variable, land price, age above 65, and civil facilities were significantly ($p < 0.05$) associated with urban problems as a whole (the first column in Table 2). More specifically, jiedaos with higher land prices tended to have fewer urban problems, indicating that socioeconomically advanced communities in such jiedaos may have had better urban governance and administration, thus were observed with fewer problems. Furthermore, jiedaos with a higher proportion of aged population tended to have fewer urban problems, possibly due to that aged population having less access to the Internet and thus fewer chances to lodge complaints and report problems. However, it was surprising to observe that jiedaos with more provisions of civil facilities tended to have more urban problems, possibly due to facilities that brought nuisance, noise and disturbance to local communities. For example, traffic facilities (e.g., train and bus stations, and transit hubs) were observed to be positively ($\beta = 0.687$, $p < 0.05$) relevant to the prevalence of urban problems.

When urban problems were broken down into each category, both land price and age above 65 were negatively (significant at $p < 0.05$) associated with each category of urban problems except safety and security (Category 1), while civil facilities were only positively (significant at $p < 0.01$) associated with Categories 2 and 4—in other words, urban problems related to construction and housing, and infrastructure and configuration. Meanwhile, it was noteworthy that different types of facilities were observed to be associated with different categories of urban problems. For example, fewer commercial facilities (e.g., supermarkets and shopping centres) were associated with more problems in construction and housing, and safety and security; fewer educational facilities (e.g., schools) were associated with more problems in social service and administration, while more traffic facilities were consistently associated with more problems in most categories, in particular, in infrastructure and configuration ($\beta = 1.061$), and construction and housing ($\beta = 0.827$). These observations indicated that different types of civil facilities had various impacts on urban problems; jiedaos, with deficits in the provision or lack of civil facilities, may have provoked the need for urban governance in infrastructure and configuration, and social service and administration, whereas some civil facilities, especially traffic facilities that increased human mobility and brought in annoyance, nuisance, and disruptions to local residents, had imposed adverse effects on local neighbourhoods.

### Table 2. Spatial lag regression (SLR) coefficients and model performance.

<table>
<thead>
<tr>
<th>Variable</th>
<th>All</th>
<th>Category1</th>
<th>Category2</th>
<th>Category3</th>
<th>Category4</th>
<th>Category5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Land price $^A$</td>
<td>$-0.196^{**}$</td>
<td>$-0.137$</td>
<td>$-0.154^{*}$</td>
<td>$-0.36^{***}$</td>
<td>$-0.271^{**}$</td>
<td>$-0.273^{**}$</td>
</tr>
<tr>
<td>Age above 65</td>
<td>$-0.287^{***}$</td>
<td>$-0.189$</td>
<td>$-0.29^{***}$</td>
<td>$-0.437^{***}$</td>
<td>$-0.187^{*}$</td>
<td>$-0.341^{***}$</td>
</tr>
<tr>
<td>Civil facilities</td>
<td>0.393 $^{***}$</td>
<td>0.015</td>
<td>0.407 $^{***}$</td>
<td>0.100</td>
<td>0.461 $^{***}$</td>
<td>0.114</td>
</tr>
<tr>
<td>Traffic</td>
<td>0.687 $^{**}$</td>
<td>$-0.033$</td>
<td>0.827 $^{***}$</td>
<td>0.411 $^{***}$</td>
<td>1.061 $^{**}$</td>
<td>0.471 $^{***}$</td>
</tr>
<tr>
<td>Commercial</td>
<td>$-0.508^{*}$</td>
<td>$-0.461^{**}$</td>
<td>$-0.649^{***}$</td>
<td>$-0.453^{***}$</td>
<td>$-0.281$</td>
<td>$-0.225$</td>
</tr>
<tr>
<td>Public</td>
<td>0.234</td>
<td>0.721 $^{***}$</td>
<td>0.218</td>
<td>0.515 $^{***}$</td>
<td>$-0.238$</td>
<td>0.257 $^{**}$</td>
</tr>
<tr>
<td>Educational</td>
<td>$-0.122$</td>
<td>$-0.237$</td>
<td>$-0.146$</td>
<td>$-0.311^{***}$</td>
<td>0.007</td>
<td>$-0.279^{**}$</td>
</tr>
<tr>
<td>Medical</td>
<td>0.085</td>
<td>0.145</td>
<td>0.135</td>
<td>0.129</td>
<td>$-0.086$</td>
<td>$-0.011$</td>
</tr>
<tr>
<td>Constant $^B$</td>
<td>204.98 $^{***}$</td>
<td>197.9 $^{***}$</td>
<td>240.06 $^{***}$</td>
<td>219.92 $^{***}$</td>
<td>211.65 $^{***}$</td>
<td>263.56 $^{***}$</td>
</tr>
</tbody>
</table>

| | F statistics | 3.984 | 1.056 | 4.587 | 2.923 | 7.769 | 3.370 |
| | Model Sig | 0.001 | 0.011 | 0.0 | 0.01 | 0 | 0.004 |
| | R-squared | 0.386 | 0.292 | 0.421 | 0.395 | 0.439 | 0.342 |
| | Number | 124 | 44 | 119 | 73 | 76 | 76 |

Note: *: $p < 0.1$; **: $p < 0.05$; ***: $p < 0.01$; $^A$: The coefficients of independent variables are standardised, ranging from 1 to 1; $^B$: The constants here are unstandardised and standardised constants are 0; Category 1: Safety and security; Category 2: Construction and housing; Category 3: Environment and pollution; Category 4: Infrastructure and configuration; Category 5: Social service and administration.
5. Discussion and Conclusions

Our study proposes an analytical framework to retrieve citizen-centred data from an online comment board and, through a qualitative content analysis of public comments, to reflect urban problems. On the basis of this we examine the spatial distribution of urban problems and their relation to surrounding areas and associative factors in order to ultimately imply how public participation helps with urban governance. The methodological workflow in our study integrates data web crawling and the usage of APIs for location and sentiment analysis, which can be applied to retrieve non-English qualitative big data in other contexts. By analysing urban problems reflected by negative public comments, our findings reveal that urban problems most frequently appear in industry-oriented and newly developed areas beyond the 3rd transport circle, spreading out to the remote region on the urban fringe. In most jiedaos, construction and housing are the dominant type of urban problem, most obviously observed in the inner-city districts with a large number of historical jiedaos where urban regeneration and gentrification occurred. The local clusters of urban problems by category, with the exception of safety and security, are obviously observed in old industrial areas and newly developed EDZs, and such problems spill over to adjacent areas to form a larger clustering spot most in need of urban governance offered by governments or related sectors. Moreover, jiedaos with higher land prices and a higher proportion of aged population tend to have fewer urban problems, while various types of civil facilities affect the prevalence of urban problems differently.

The spatial distribution of urban problems and the explanation for their formations lie in the institutional and political-economic mechanisms, characterised in the Chinese context. Prior to China’s reform and opening up in 1978, Chinese cities had a slow progression of socioeconomic development with weak capability of urban governance. With the acceleration of neoliberalism, and the implementation of the land system and housing reform since 1978, the reconstruction of traditional work-unit housing (danweifang in Chinese) and the privatisation of housing have stimulated large investment in real estate and infrastructure construction [46]. Consequently, urban governance has been transformed from being under the full supervision of governments to being under the dispersed supervision of multiple parties (e.g., real estate developers and private property management companies) together with the control of governments [43]. Private governance endorsed by real estate developers and property management companies plays an increasingly important role in providing service and administration in post-reform, commercially built commodity housing (shangpingfang in Chinese), mostly gated communities that are privately managed and designed according to the tastes of the growing urban middle class [46]. Such gated communities with high land values and housing prices are usually well located around city centres with easy access to civil facilities and cityscape, or newly developed on the urban fringe with a high-class neighbourhood environment, good service provision and administration, and well-designed public space (e.g., clear streets, sufficient green space and parking lots, and visually present landscape). Residents living in high-class communities pay high corporate and administrative fees for good life quality and well-organised community service provided by real estate developers or property management companies who take over the managerial role from local governments and provide governance at the community level. Conversely, communities with low land value and housing prices are usually old communities built in the early stage of China’s land system reform or traditional work-unit housing mostly located in and around the inner city. Residents living in these low-valued neighbourhoods have limited access to services provided by real estate developers or property management companies, concurrently lacking sufficient governance and public services given by local governments, and thus associated with a high prevalence of urban problems.

The dilemma of government-centred and private governance is accompanied by the deepening process of industrial transformation and upgrading, and such dilemmatic phenomena have been widely observed in old industrial areas in Chinese cities [58]. The particular example is WISC industrial park, with the highest number of urban problems in construction and housing, followed by environmental pollution, infrastructure and configuration. As a large state-owned company, WISC provided both working and living space as well as public services and administration to employees in the old era.
before the economic resolution of the free market in China in 1978 [59]. After the economic resolution and industrial transformation in 1978, WISC changed its company structure, no longer playing a managerial role in providing social services and administration in the jiedao, where providers of urban governance were replaced by government and private sectors [60]. If such replacements are not initiated efficiently, deficits of governance will appear, leading to increasing urban problems and complaints from local residents. This governance-deficit phenomenon is also observed in another three EDZs as the red-spots of urban problems, including Miaoshan, Yangluo, Yaojiashan EDZs. These zones have been transformed from old industrial districts to more mixed land-use areas for residence and recreation in the past two decades of industrial transformation, but they lack the efficient provision of public service and amenities.

The provision of civil facilities needs to be better assessed by type, catering to the needs of local residents, and different types of civil facilities and their supply–demand balance/imbalance will trigger urban problems differently. Surprisingly, this finding runs counter to the empirical observation in existing literature that the better provision of public facilities results in fewer urban problems [61]. The possible explanation is that the diversity of public facilities brings convenience but simultaneously causes some issues for local residents. In reality, citizens seem to have conflicting attitudes towards the large-scale civil facilities. In particular, commercial and educational facilities are expected by residents to be built close to communities for the convenience and easy access, while the lack of such facilities, as seen in our findings, increases the demand for better urban governance in infrastructure and configuration, and social service and administration. The large-scale commercial and entertainment facilities, however, as well as traffic facilities also bring in annoyances, nuisances, and disruptions to local residents and impose adverse effects on local neighbourhoods due to the increasing human mobility, noise, parking shortage, and traffic congestion. Moreover, power stations, garbage dump grounds, waste transit sites and stations, and other ageing facilities have also been mentioned in public comments as facilities ‘kept away from my backyard’, and which have adverse impacts on the health, visibility, and environment of surrounding areas and indirectly decrease property prices.

Our findings are based on the actual voices of citizens reflecting their real problems and needs, and thus provide evidences for political implications that can be considered to improve the efficiency of future public participation and urban governance. First, state and municipal governments are advised to establish policies, such as national planning codes, to standardise the provision of public goods and services at the community level and to further guide through private governance. Second, real estate developers and property management companies are encouraged to participate in the process of policy making and to nominate state agencies to regulate private provision of community services and administration, in terms of detailed responsibilities and obligations as well as the spatial boundaries of administration. The main purpose of introducing private governance should be for responding to residential demands, rather than stimulating the housing market and generating local revenues from residential land leasing. Third, online comment boards and forums are effective platforms where citizens can express their demands before decision-making, and supervise and give feedback after decision-making. The development of such online tools can be extended to interactive maps and cloud sharing empowered by advanced technologies such as machine learning and artificial intelligence, in order to enable and broaden public participation throughout the whole process of decision-making. Finally, public participation needs be promoted through multiple approaches, from online platforms to offline surveys, to involve the aged population who may have limited access to the Internet and be somehow neglected in the current theme of public participation [12,62]. Accounting for the demands of aged groups and ageing communities needs to be prioritised in urban governance for a better social justice and equity.

There are a number of limitations in this study where future research can expand. First, the citizen-centred data retrieved from the online comment board lack the demographic and socioeconomic information of comment providers and the areas where they reside (e.g. house rental, population density, Internet accessibility, educational and occupational levels), which can be part of the
associative factors to explain urban problems. Future work can be extended along this line by involving individual-level data or aggregated areal data from multiple sources (e.g., real estate websites or survey data). Second, the collection of citizen-centred data as a stratified sampling approach inevitably brings in some biases because the data are limited to comment providers within a certain age group with accessibility to Internet or smart phones. The existence of such data biases will be potentially diminished with the progress of smart-city initiatives in China by the increasing prevalence of mobile devices, the more intelligent designation of end-user functions, and the growing dissemination of e-participation in the future to cover a wider range of the population. Moreover, such data biases can be calibrated by involving multi-sourced data such as census data and survey data to validate the robustness of data analysis. Third, the methodological approach proposed in our study integrates web-crawling tools and the use of APIs, which can be applied to non-Chinese qualitative data retrieving and analysis. The accuracy and application of this approach can be improved by advanced technologies such as deep learning and AI optimisation.

The importance of our study lies in the construction of the analytic framework using public comment qualitative data and the provision of the empirical study in a Chinese megacity to investigate e-governance and e-participation. Our study enriches the diversity of studies on urban governance, shifting from a government-centred to a citizen-centred perspective, and from a top-down to a bottom-up approach. The analytical outcomes provide valuable evidence to Chinese local governments and authorities for policy reforms and implementations. Our discussion about participatory governance and citizen-centred governance is on the frontier of research on urban governance, extending the Western-dominant literature with Chinese experiences. In sum, our study, as an initial attempt at retrieving non-English qualitative data to analyse urban problems, explores a promising research direction of integrating citizen-centred big data into urban governance and public participation. The analytical framework and approaches used in our study can be applied to future research in different linguistic and geographic contexts, along with collective efforts that can be put into the cross-contextual studies.

Author Contributions: Conceptualization, Z.G. and J.G.; Data curation, Z.G.; Formal analysis, S.W. and Z.G.; Funding acquisition, Z.G.; Investigation, Z.G.; Methodology, Z.G., S.W. and J.G.; Project administration, Z.G.; Resources, Z.G.; Software, Z.G. and S.W.; Supervision, Z.G.; Validation, Z.G., S.W. and J.G.; Visualization, S.W.; Writing—original draft, Z.G. and S.W.; Writing—review and editing, S.W., Z.G. and J.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (42001188 and 41601166), Hubei Provincial Natural Science Foundation of China (2020CFB350), and Fundamental Research Funds for the Central Universities (CCNU20QQN031).

Acknowledgments: We would thank the efforts from Yuxiao Li, Wenjun Jiang and Tingting Qiang, the undergraduates from the College of Urban & Environmental Sciences, Central China Normal University, for data collection and preparation.

Conflicts of Interest: The authors declare no conflict of interest.

References


22. Abella, A.; Ortiz-de-Urbina-Criado, M.; De Pablos-Heredero, C. Information reuse in smart cities ecosystems. Prof. Inf. 2015, 24, 838–844. [CrossRef]


24. Kitchin, R. Big Data, new epistemologies and paradigm shifts. Big Data Soc. 2014, 1, 2053951714528481. [CrossRef]


29. Schreier, M. Qualitative Content Analysis in Practice; Sage Publisher: London, UK, 2012.


40. Shen, J.; Wu, F. The suburb as a space of capital accumulation: The development of new towns in Shanghai, China. Antipode 2017, 49, 761–780. [CrossRef]
47. He, S. Homeowner associations and neighborhood governance in Guangzhou, China. Urban Stud. 2005, 42, 1859–1878. [CrossRef]


**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).