



Yong Sun <sup>1,2</sup>, Min Ji <sup>2</sup>, Fengxiang Jin <sup>1</sup> and Huimeng Wang <sup>1,3,\*</sup>

- <sup>1</sup> College of Surveying and Geo-informatics, Shandong Jianzhu University, Jinan 250101, China; ttsunyong@163.com (Y.S.); fxjin@sdjzu.edu.cn (F.J.)
- <sup>2</sup> College of Geodesy and Geomatics, Shandong University of Science and Technology, Qingdao 266590, China; jamesjimin@126.com
- <sup>3</sup> State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Science and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China
- \* Correspondence: wanghm@lreis.ac.cn

Abstract: As air users, the public is also participants in air pollution control and important evaluators of environmental protection. Therefore, understanding the public perception and response to air pollution is an essential part of improving air governance. This study proposed an analytical framework for public response to air pollution based on online complaint data and sentiment analysis. In the proposed framework, the emotional dictionary of air pollution was firstly constructed using microblog data and complaint data. Secondly, the emotional dictionary of air pollution and the sentiment analysis method were used to calculate public complaints' emotional intensity. Besides, the spatial and temporal characteristics of air pollution complaint data and public emotional intensity, the complaints content, and their correlation with PM2.5 (particulate matters smaller than 2.5 micrometers) and PM10 were analyzed using address matching, spatial analysis, and word cloud analysis. Finally, the proposed framework was applied to 13,469 air pollution complaint data in Shandong Province from 2012 to 2018. The obtained results indicated that: the public was mainly complaining about the exhaust gas emissions from enterprises and factories. Spatially, the geographical center of complaint data was located in the inland industrial urban agglomeration of Shandong Province. Correlatively, air pollution complaints' negative emotional intensity was significantly negatively correlated with PM2.5 (-0.73). Moreover, the number of public complaints about air pollution and the intensity of negative emotions also decreased with improved air quality in Shandong Province in recent years.

**Keywords:** air pollution; public complaints; sentiment analysis; address matching; spatiotemporal analysis; correlation

#### 1. Introduction

Air pollution is a major environmental and health problem affecting everyone, since it is associated with a negative impact on human physical and mental health, and social sustainable economic growth [1–3]. Therefore, it has become the subject of widespread attention from the academic community. Several previous studies have been conducted in environmental science, geography, health, and biology. They have discussed the temporal and spatial changes and trends of air pollution, influencing factors, dynamic monitoring, and its impact on public physical, mental health, and quality of life [3–6].

Thus, it is necessary to monitor the level of pollutants and reduce pollutant emissions to solve the problem of air pollution. Besides, understanding the public's subjective attitude, concern, and behavioral response to air pollution is very important for improving air quality, because the public is also a participant in air pollution control and an essential evaluator of environmental protection performance [6–8]. The public perception of air quality could find the potential pollution sources that the monitoring stations cannot



Citation: Sun, Y.; Ji, M.; Jin, F.; Wang, H. Public Responses to Air Pollution in Shandong Province Using the Online Complaint Data. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 126. https://doi.org/ 10.3390/ijgi10030126

Academic Editors: Wolfgang Kainz and Yeran Sun

Received: 13 January 2021 Accepted: 23 February 2021 Published: 1 March 2021

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



**Copyright:** © 2021 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 (https:// creativecommons.org/licenses/by/ 4.0/). identify. It can also reflect the effectiveness of air pollution management to a certain extent. That is because the stations monitoring air quality are local and limited, while the public perception of air pollution is ubiquitous. Moreover, the development of the Internet and the enhancement of people's awareness of environmental protection have resulted in the public becoming more and more concerned about the dynamic effect of air pollution and its impact on daily life and health [7,8]. The national government and the public have also taken some measures to control air pollution. For example, in 2015 and 2018, China repeatedly revised and improved the law of the people's Republic of China on the prevention and control of atmospheric pollution, which guides the public to perceive and control air pollution [9]. Many scholars have also developed air pollution monitoring equipment and actively constructed and improved the urban green ecosystem [10]. The public has again gradually turned to clean energy and green consumption to reduce emissions [10,11], such as buying new energy vehicles, and taking buses and subways.

Some studies used questionnaires and social media to analyze public perceptions of air pollution. The research on public perception of air pollution based on a questionnaire survey involves setting up relevant survey questions according to different research purposes, sample public attitudes, and opinions in the form of questionnaire (online or offline) in a large area or small range. The obtained questionnaire survey results are then analyzed to understand the public's understanding and concerns about air pollution [7,12,13]. For example, Dons et al. (2018) used an online questionnaire survey involving 7622 adults to analyze people's concern about air pollution and its health effects [12]. On the other hand, Pu et al., (2019) used 9744 qualified questionnaires as samples, covering 31 provinces in China, assessed the public's awareness of the risk of air pollution using the psychometric paradigm method [7]. Besides, the study by David et al., (2020) used 198 residents (in the San Joaquin Valley, California), and they found that people who are more worried about air quality tend to pay more attention to and check air quality [13]. The research of public perception of air pollution based on social media platforms used the data obtained from Twitter, Facebook, or Sina Weibo to assist in monitoring the dynamic changes and trends of air quality, and to analyze the public response to the changes in air quality [5,14–16]. For example, Jiang et al. (2015) investigated the Spatio-temporal trend of social media information with geographic positioning function, and effectively monitored the dynamic changes of air quality in big cities [14]. Wang et al., (2017) proposed a method to infer urban air quality in China based on Sina Weibo (China's largest social media platform) and other meteorological data, which can be used for urban air quality monitoring and evaluation in cities with few or no monitoring sites [5]. Furthermore, Gurajala et al., (2019) explored the content of Twitter data (tweets) to understand the public's response to changes in air quality over time and their understanding of air quality issues [15]. On the other hand, Hswen et al., (2019) captured public tweets containing air pollution terms from the Twitter platform. They used correlations to analyze the relationship between the number of public tweets with positive, negative, or neutral sentiments regarding air quality and PM2.5 changes [16].

Questionnaire surveys and social media platforms have certain advantages in understanding the public perception of air pollution. For example, questionnaires can be designed to collect information according to the research problems of air pollution. Social media platforms can provide real-time updates of air pollution information and detect dynamic changes in human behaviors. However, they also have some disadvantages. Questionnaire surveys are limited because they require many human resources, material resources, and time [5]. On the one hand, social media platforms are complex, and it requires one to analyze the positive, negative, neutral, mixed, and other emotion types in the text. In addition, there may be bias, and users may express unreal emotions due to social purposes [17].

Complaint data (in the form of text) are the behavior caused by the public's dissatisfaction. It can also be defined as the behavior of the public to protect themselves through complaints. Therefore, the data obtained from public complaints about air pollution can directly reflect the real dissatisfaction (negative emotion) of the public about air pollution. Generally, complaint data can be obtained from the Internet. For example, government departments in China have opened an environmental pollution complaint website to facilitate public complaints to control environmental pollution effectively. The government website also collects related pollution complaints through multiple channels, such as social media networks like complaint telephone, Baidu Tieba, and Sina Weibo. Therefore, the complaint data are relatively comprehensive and has an essential impact on the assessment environment. Some scholars have used the complaint data to research environmental or air pollution to understand the potential pollution problems hidden in the data. Previous studies have reported that these problems are sometimes not found by the monitoring station [18–20]. For example, Gallagher et al. (2014) used customers' complaints about their drinking water quality to analyze severe potential water quality problems [15]. Avaliani et al. (2015) found potential pollution sources according to residents' odor complaint data submitted by residents [19]. Moreover, Pan et al. (2020) used the spatiotemporal analysis of odor complaint data to identify the specific manhole or sewer location that caused the odor [20]. However, the studies mentioned above mainly focus on the content of complaint data from the perspective of statistics and spatiotemporal analysis, and rarely focus on the range of public complaints' subjective feelings. It is essential to understand the subjective emotion of public complaint data because the change of public emotion intensity in complaint data can reflect the trend, the severity of air pollution, and the effectiveness of governance to a certain extent.

The primary research question is on how to understand public sentiment through complaint text data. Sentiment analysis, precisely text sentiment analysis, is the most effective method. The main aim of text sentiment analysis, also known as opinion mining and tendency analysis, is to detect the polarity of emotions (such as negative or positive emotions) or quantify the intensity of emotions from the text [21–23]. Currently, there are three commonly used methods for text sentiment analysis, namely, "dictionary-based", "machine learning", and "deep learning" methods. Machine learning and deep learning methods require manual labeling of large amounts of data, followed by supervised training [22]. The dictionary-based sentiment analysis does not require training, and thus it is the most commonly used and most straightforward method [23]. A previous study reported that the dictionary-based sentiment analysis method refers to calculating sentiment value based on the labeled sentiment dictionary [24]. A well-annotated and perfect sentiment dictionary is an important work of sentiment analysis based on dictionary. There are several requirements of the emotional dictionary: to include positive words and negative words, the number of each type should be enough, and the scope should be wide enough [25,26]. Therefore, it is very necessary to have a good dictionary and then a method of how to analyze emotions.

Therefore, this paper proposed an analysis framework of public response to air pollution based on complaint data and sentiment analysis. In the proposed framework, the air pollution emotion dictionary was first constructed. In addition to the complaint text data obtained from the online platform, about 300,000 microblog text data associated with air pollution were integrated to extract wide enough emotional words related to air pollution to make the constructed dictionary more complete. Secondly, the intensity of public dissatisfaction with air pollution was calculated using the sentiment analysis method, which is used as the public response to air pollution. Moreover, address matching, spatial analysis, and word cloud analysis methods were used to explore the spatial and temporal distribution characteristics of air pollution complaint data and complaint emotional intensity, and the content of concern. At the same time, the correlation between the number of air pollution complaint data and the intensity of complaint emotions and the daily air quality index (AQI), PM2.5, and PM10 levels issued by the Ministry of Environmental Protection of China was also discussed to answer the public's feelings about air pollution. The main question was whether the intensity is associated with air quality. In addition, the paper also discusses the correlation between the number of complaint data and the

public's emotional intensity of air pollution and the daily air quality parameters of PM2.5 and PM10 issued by the Ministry of Environmental Protection of China to answer the question of whether the public's emotional intensity of air pollution is associated with air quality. Finally, the proposed framework was applied to the air pollution complaint data of Shandong Province from 2012 to 2018.

#### 2. Materials and Methods

# 2.1. Data

This paper mainly used three types of data, namely, complaint data related to air pollution, microblog data, and AQI data. The complaint data and microblog data related to air pollution are used as samples for extracting emotional words to construct the emotional dictionary for air pollution. Complaint data are used primarily for public sentiment analysis of complaints. The AQI data are used to confirm the changing trend of air quality and analyze its correlation with the number of public complaints and the emotional intensity of public complaints. The following are the sources and information of these three types of data.

# 2.1.1. Public Complaint Data

The public complaint data used in this paper were obtained from the Shandong environmental public prosecution network platform. On this website, the public can choose the type of environmental pollution (classification) independently, make the ecological pollution complaints by filling in the accusation's general title, the complaint's location, the pollution status, and other details. After successful submission, the website will automatically generate the number, complaint date, and other information, and update the relevant departments' replies in real-time.

The complaint data published on the Shandong Environmental Complaint platform were collected through multiple channels. In addition to the environmental problems directly reflected by the public through the complaint platform, it also includes collecting negative environmental issues from the public through complaint telephone, Baidu Tieba, and Sina Weibo. Baidu Tieba is the world's largest Chinese-language community, and Sina Weibo is China's largest broadcast-style social media sharing short and real-time information. Therefore, the Shandong environmental complaints platform's complaint data is relatively comprehensive and can reflect the Shandong public's perception of air pollution.

This study obtained the initial data of 13,469 complaints about air pollution from the Shandong Environmental Prosecution Platform dating from 2012 to 2018. Each complaint data includes an ID (automatically generated primary key) number, complaint date, title, content, reply, and other information. The examples of complaint records of air pollution and their corresponding English translations are shown in Table 1.

#### 2.1.2. Microblog Data

Before analyzing public air pollution complaint data, we first had to build an emotional dictionary for air pollution. In addition, it is not enough to only use the complaint data. Therefore, this study used Sina API (application program interface) and a web crawler technology to crawl some microblog messages about "air pollution" to make the number of words in the constructed emotional dictionary large enough and wide enough. That resulted in more than 300,000 pieces of microblog messages being obtained through this collection method.

#### 2.1.3. AQI Data

The AQI data was obtained from China's Ministry of Environmental Protection in order to verify the correlation between public complaints' number and emotional intensity and air quality. China began to monitor PM2.5 on a large scale and released air pollution monitoring data around December 2013. Therefore, we downloaded the AQI daily mean data from a period ranging from 1 January 2014 to 31 December 2018.

| Number       | Date      | Title  | Content   | Reply  |
|--------------|-----------|--|---|--|
| 201211031525 | 2012/11/3 | Heze chemical plant<br>seriously pollutes<br>the environment   | Recently, there is a strong smell of<br>chemical industrial waste gas near<br>Guangzhou road and Huanghe East<br>Road in Heze City from 19 o'clock<br>every night to 7 o'clock the next<br>morning, which seriously threatens<br>the physical and mental health<br>of residents.                                | [Heze Environmental Protection<br>Bureau] After receiving the problems<br>reported by the masses, it<br>investigated the chemical companies<br>in the development zone one by one,<br>and found that Shandong Dongyao<br>Pharmaceutical Co., Ltd. had many<br>environmental problems such as<br>odor of exhaust gas, and issued a<br>notice of production suspension<br>and rectification.   |
| 201708033338 | 2017/8/3  | Exhaust gas is secretly<br>discharged, smoke and<br>dust pollution is serious,<br>evading inspection | I am a villager where this factory is<br>located. Every day, the factory has<br>serious odors, secretly venting gases,<br>and does not have any<br>environmental protection equipment.<br>And the dust smell is serious. I hope<br>that the provincial leaders will pay<br>attention to our living environment. | [Zibo Yiyuan Environmental<br>Protection Bureau] there is<br>unorganized emission of smoke and<br>dust in Zibo Deyuan metal materials<br>Co., Ltd. The environmental<br>supervisors of our bureau required<br>the enterprise to: first, build new air<br>pollutant treatment facilities<br>matching the production process;<br>second, during the rectification<br>period of the production workshop,<br>take measures to limit production or<br>stop production to ensure the<br>emission up to the standard. |

Table 1. Sample data of public complaints.

AQI describes how clean or polluted the air is and how it affects health. It is worth noting that the AQI value begins to characterize different degrees of air pollution when the value is greater than 100. In addition to AQI, the data also includes monitoring six parameters: CO, NO<sub>2</sub>, SO<sub>2</sub>, O<sub>3</sub>, PM2.5, and PM10. The parameters CO, NO<sub>2</sub>, SO<sub>2</sub>, and O<sub>3</sub>, are associated with the concentrations of different gases. PM2.5 and PM10 are associated with airborne particles, where the diameter of PM2.5 is smaller than 2.5  $\mu$ m, and the diameter of PM10 is shorter than 10  $\mu$ m. Air pollution occurs when the PM2.5 level is more excellent 75  $\mu$ g/m<sup>3</sup>, or the PM10 level is more excellent 150  $\mu$ g/m<sup>3</sup>.

In this study, we mainly used the days in which the AQI value exceeded the standard to analyze the change of air quality, and used PM2.5 and PM10 levels to analyze the correlation with public complaints' number and emotional intensity of public complaints, because it is easier for the public to perceive PM2.5 and PM10 with a certain concentration when compared with invisible gases such as CO, NO<sub>2</sub>, SO<sub>2</sub>, and O<sub>3</sub>. For example, the formation of haze is mainly associated with PM2.5 [16].

#### 2.2. Methods

This study proposed an analysis framework of public response to air pollution based on online complaint data and sentiment analysis. The research method route used is shown in Figure 1.

Firstly, the air pollution emotion dictionary was constructed using microblog data and complaint data, The emotion intensity of complaint text was calculated using the text emotion analysis method. Secondly, the address matching method was used to locate each complaint datum's geographic location. Then, we used spatial analysis methods such as kernel density, standard deviation ellipse, hot spot analysis to analyze the spatial aggregation characteristics and temporal variation characteristics of public complaint data and their emotional intensity. Moreover, the air quality data, PM2.5, and PM10 were used to verify the actual trend of air pollution. Finally, we used correlation analysis of the statistical method to study the relationship between the number and emotional intensity of public complaints about air pollution, and PM2.5 and PM10 levels.



**Figure 1.** Flow chart of the research methods for public response to air pollution based on online complaint data and sentiment analysis.

#### 2.2.1. Building an Emotional Dictionary for Air Pollution

The emotion dictionary of air pollution was constructed according to the extracted microblog data and complaint data associated with "air pollution", combined with basic emotion dictionaries such as the emotion dictionary of Dalian University of Technology and modifier dictionary (including degree adverb dictionary and negative word dictionary). The emotion dictionary of Dalian University of Technology is organized and annotated by researchers of the Information Retrieval Research Office under Professor Lin Hongfei [27]. The dictionary describes a Chinese word or phrase from different perspectives, including word part of speech, emotion type, emotion intensity, and polarity. It can provide an efficient and reliable method for sentiment analysis and orientation analysis of Chinese text.

Firstly, Hanlp was used for Chinese word segmentation and part of speech tagging to extract verbs, nouns, adjectives (adj), network terms (nw), adverbs (adv), and idioms as candidate emotional words. The extracted words were then combined with the Dalian University of Technology Emotion Dictionary for deduplication. The labeling criteria was then formulated as shown in Table 2, according to the polarity classification of air pollution emotional words reported in a previous study [28]. The artificial selection of emotional words enabled the calculation of emotional words' intensity using multi person tagging. The following judgment and analysis were then made according to the results of each person's annotation: If  $\sum_{i=1}^{n} \sum_{j=1}^{n} |S_i^t - S_j^t| \le 2n$ , then  $S^t = \frac{1}{n} \left(\sum_{i=1}^{n} S_i^t\right)$ , where n is the number of tags,  $S_i^t$  indicates the result of the *i*-th person marking the *t*-th word, and  $S^t$  indicates the

final emotional strength of the *t*-th word; If  $\sum_{i=1}^{n} \sum_{j=1}^{n} |S_i^t - S_J^t| > 2n$ , then label again. After the labeling was completed, we integrated the constructed air pollution emotional dictionary with the Dalian University of Technology emotional dictionary. We added the modifier dictionary to obtain the final air pollution emotional dictionary.

Table 2. Emotion intensity labeling criteria.

| Polarity | Intensity                 | Description                | Example           |
|----------|---------------------------|----------------------------|-------------------|
|          | 9                         | Cannot survive             | Killed, Cancer    |
| -        | 7                         | Danger to the human body   | Dizziness, Nausea |
| Negative | 5                         | Affect normal life         | Pungent, Choking  |
| reguire  | 3 Has an unpleasant smell | Bad smell, Stench          |                   |
|          | 1                         | Can feel, but not affected | Sewage, Soot      |

#### 2.2.2. Sentiment Analysis

This study analyzed the sentiment of the public complaints about air pollution by subscribing and segmenting the complaint text to obtain the sentiment intensity value of the public complaints based on the air pollution sentiment vocabulary that we had constructed. The following steps are also included:

# 1) Text preprocessing

The complaint text is divided into sentences, words, and stop words from realizing the preprocessing of text data.

2) Emotional word matching and modifier matching

The words obtained from text preprocessing were matched with the emotion words and modifier in the air pollution emotion dictionary constructed in Section 2.2.1. We could get the emotional words and modifiers in the sentence and mark their position in the sentence.

#### 3) Calculation of emotional intensity of public complaints

After matching emotion words and modifiers, the emotional intensity was calculated. Firstly, we used formula (1) to calculate the emotional intensity of public complaints in short sentences; secondly, we used formula (2) to calculate the emotional intensity of public complaints in sub sentences according to the calculation results of short sentences; finally, we used formula (3) to calculate the emotional intensity E(S) of public complaints on the complaint text according to the calculation results of sub sentences.

$$E(P_i) = E(SW) \times E(DW) \times (-1)^n \tag{1}$$

$$E(C_i) = Min(E(P_1), E(P_2), \cdots)$$
<sup>(2)</sup>

$$E(S) = Min(E(C_1), E(C_2), \cdots)$$
(3)

where: *SW* is the emotional word, *DW* is the degree adverb, E(SW) is the emotional intensity of the emotional word, *n* is the number of negative words, and E(DW) is the weight of the degree adverb. Moreover,  $E(P_i)$  is the emotional value of the *i*th phrase,  $E(C_i)$  is the affective value of the *i*th clause, and E(S) is the emotional value of the whole text.

#### 2.2.3. Address Matching

After calculating public complaint emotional intensity, it is necessary to match the address for the geographical location described by each complaint datum because the complaint data are not recorded in the form of geographical coordinates. Previous studies have reported that address matching is the process of establishing the corresponding relationship between the literal description address and its spatial geographic coordinates [29,30].

In this paper, address matching was done for the title, content in the complaint data, and the text in the reply using Chinese word segmentation and part of speech tagging, address extraction, address normalization, longitude and latitude matching, and other methods.

- (1) Hanlp was used for Chinese word segmentation and part of speech tagging;
- (2) Address extraction: we analyzed the words after part of speech tagging, and detected the address parts of speech in terms including organization name and place name. Besides, we stored the words of organization name and place name.
- (3) Address normalization: it is necessary to standardize the extracted address and standardize the address information of each comment data as: (Province, city, district and county, and detailed address] since the public often uses abbreviations to represent address information in expression. It is worth noting that when performing detailed address matching, the factory or company name will be used directly in instances where the factory or company name appears. If there is no factory or company name, match 'town' + 'village'. If none of the above, match' Road name' + 'community'.
- (4) Longitude and latitude matching: This study used the AutoNavi Map API to perform latitude and longitude matching on the address information after address normalization.

#### 2.2.4. Spatial Analysis and Statistical Analysis

This study used the kernel density method to analyze the spatial distribution characteristics of complaint data. The kernel density analysis method was used to calculate the density of the feature in its surrounding neighborhood. The obtained kernel density of the complaint points could show the places where air pollution mainly occurred.

The standard deviation ellipse method was then used to analyze the interannual change trend of the geographical distribution of complaint data. The standard deviation ellipse method is one of the classic techniques used to analyze the spatial distribution's directional characteristics. It creates a standard deviation ellipse to summarize the complaint data's spatial characteristics, that is, the central trend and the directional trend. The size of the ellipse can reflect the concentration of the spatial pattern of complaint data, while the deflection angle (semi-major axis) reflects the dominant direction of the pattern.

The hot spot analysis method was then used to identify the spatial heterogeneity of public emotional intensity at the local level. This method can calculate Getis-Ord Gi\* statistics [31] for each emotional intensity value in the complaint data. We then used the obtained z-score and p-value to identify significant spatial clustering of high and low emotional intensity values, referred to as hot spots and cold spots, respectively [32].

Finally, the Pearson correlation coefficient was used to analyze the correlation coefficients between the number of complaints and public complaint sentiment and PM2.5 and PM10 levels to explore whether the number of public complaints and complaint emotions is associated with air quality changes. We then conducted a significant test to determine the degree of correlation between them.

The Pearson correlation coefficient between two variables is defined as the quotient of the covariance and standard deviation between the two variables. The Pearson correlation coefficient is defined as:

$$\rho_{XY=\frac{cov(X,Y)}{\sigma_X\sigma_Y}} \tag{4}$$

COV(X, Y) is the covariance of variables *X* and *Y*,  $\sigma_X$  is the standard deviation of *X*, and  $\sigma_Y$  is the standard deviation of *Y*.

The closer the correlation coefficient is to 1 or -1, the stronger the correlation, and the closer the correlation coefficient is to 0, the weaker the correlation.

# 3. Results

# 3.1. Spatiotemporal Analysis of Air Pollution Complaint Data

3.1.1. Temporal Characteristics of Air Pollution Complaint Data

The analysis framework proposed in this article indicated that the total number of public complaints associated with air pollution in 17 cities in Shandong Province from 2012 to 2018 was 13,469, with an average of 1924 being generated every year. The data was obtained from the Shandong Environmental Prosecution Network Platform. The interannual and seasonal changes of the total number of complaint data are shown in Figures 2 and 3.



**Figure 2.** The interannual changes in the total number of air pollution-related complaint data and the number of days that PM2.5 and PM10 exceeded the standard in 17 cities in Shandong Province from 2012 to 2018.



**Figure 3.** Seasonal variation of the total amount of air pollution-related complaints data and days of exceeding AQI, PM2.5, and PM10 levels in 17 cities of Shandong Province from 2014 to 2018.

The blue column in Figure 2 shows the number of public complaints, while the curve shows the trend line of annual complaints. Besides, the orange, yellow, and broken green lines show the days of AQI exceeding the standard (AQI > 100), PM2.5 exceeding the standard (PM2.5 > 75  $\mu$ g/m<sup>3</sup>), and PM10 exceeding the standard (PM10 > 150  $\mu$ g/m<sup>3</sup>), respectively, in 17 cities in Shandong Province.

The results obtained in this study indicated that public complaints about air pollution increased year by year from 2012 to 2015, and decreased year by year from 2015 to 2018 (Figure 2). Accordingly, the number of days exceeding AQI, PM2.5, and PM10 decreased year by year from 2014 to 2018, with an average annual decrease of 9.4, 12.8, and 12.5%, respectively.

On the one hand, the increase in the number of complaints may indicate that the public's awareness of environmental protection has increased. It may also suggest that air pollution has increased. On the other hand, the decrease in the number of public complaints can indicate that the air quality in Shandong Province has shown a trend of improvement in recent years. This phenomenon can also be confirmed by the obtained statistical analysis results, which showed a decreasing trend in the number of days where the AQI, PM2.5, and PM10 levels exceeded the standard.

This study further analyzed the number of complaints and the days AQI, PM2.5, and PM10 levels exceeded the standard in spring (March, April, and May), summer (June, July, and August), autumn (September, October, and November), and winter (December, January, and February) to analyze the difference of the number of public complaints about air pollution and the number of days the AQI data exceeded the standard in different seasons.

In Figure 3, the blue column represents the total number of complaints in 2014–2018 according to four seasons, while the orange, yellow, and broken green lines represent the number of days that the AQI exceeded the standard (AQI > 100), PM2.5 exceeded the standard (PM2.5 >  $75\mu g/m^3$ ), and PM10 exceeded the standard (PM10 >  $150\mu g/m^3$ ), respectively, in 17 cities in Shandong Province. Figure 3 shows that the number of days in which the AQI, PM2.5, and PM10 levels exceeded the standard were the least in summer and the most in winter in the four seasons. However, there were relatively more public complaints in the summer and the least in the winter.

This paper can attribute the above result to two reasons. Firstly, it may be that the high temperature in the summer makes the public more sensitive and more irritable. For example, a higher temperature is associated with faster sports and a stronger pungent smell, making it easier for the public to complain. In winter, the cold and low temperature makes the public less sensitive, and the public travel activities are reduced. Thereby the number of public complaints is also relatively reduced. Secondly, air quality monitoring stations' spatial distribution is limited [30], while the public is everywhere. Therefore, it is understandable that the air pollution phenomenon is seen or smelled deviates from the monitoring content of air quality monitoring stations.

#### 3.1.2. Spatial Characteristics of Air Pollution Complaint Data

The spatial distribution map (Figure 4a) and kernel density map (Figure 4b) of "air pollution" complaint data in Shandong Province from 2012 to 2018 were obtained using the address matching method described in Section 2.2.3 and the kernel density analysis method described in Section 2.2.4 (Figure 4).

Figure 4a shows that all 17 cities in Shandong Province had complaints about air pollution, while Figure 4b shows that the high-density areas of Shandong public complaints about air pollution are mainly concentrated in Laiwu, Zibo, and Weifang. It is worth noting that Laiwu was the city with the highest density of complaints. This shows that the public in Laiwu area has the most positive response to air pollution, and it also indicates that there may be serious air pollution problems in Laiwu area.

This study used the standard deviation ellipse method to analyze the changing trend of the average geographic center and orientation of the complaint point data in Shandong Province year by year to explore whether the distribution of public air pollution complaints data from 2012 to 2018 had evolutionary characteristics. The obtained results are shown in Figure 5.



**Figure 4.** Distribution of air pollution complaint data in Shandong Province from 2012 to 2018 (**a**) and kernel density map of complaint location (**b**).



(a) Standard deviation ellipse of Shandong Province

(b) Statistics of ellipse center offset in 17 cities

**Figure 5.** Standard deviation ellipse analysis of Shandong Province (**a**) and statistics of ellipse center offset in 17 cities (**b**) of air pollution complaint data from 2012 to 2018.

The different colored ellipses in Figure 5a represent the complaint data's standard deviation ellipses in different years. Figure 5b shows that the interannual offset of 17 cities' ellipse centers is within 14.5 km. The obtained results indicated that no noticeable evolution difference in the spatial distribution of Shandong Province public complaints about air pollution from 2012 to 2018, whether the average geographical center (the blue dot in the figure) or the direction trend (the blue line). Although there are some differences in the evolution differences in the evolution of 17 cities in different years, the ellipse center has not changed much.

In recent years, Shandong Province's public complaints about air pollution have mainly been distributed along the southwest–northeast direction. The complaints are primarily concentrated in Shandong's central cities such as Jinan, Laiwu, Zibo, Weifang, Jining, and other cities. Furthermore, the complaints are relatively less distributed in coastal cities such as Qingdao, Weihai, Yantai, Dongying, Binzhou, Rizhao, and other cities.

These findings may be attributed to these cities' development direction, where Jinan, Laiwu, Zibo, Weifang, Jining, and other cluster cities mainly focus on industrial projects [33]. These cities have developed industries, and thus they naturally cannot avoid air pollution. Moreover, most of the industries in the cities mainly focus on producing chemical products

by refineries, which is also the birthplace of air pollution. Therefore, the cities with industries as the key development projects have suffered industrial development and air quality problems despite the industries driving up economic development. Besides, there was no great difference in the center and direction of public complaints in recent years, which may be due to some inherent air pollution problems associated with industrial development. However, the specific problems and reasons should be further studied with the help of refined data. From another perspective, the results also show that the public has a stronger sense of complaint in cities with industrial development.

# 3.2. Spatiotemporal Analysis of Public Complaint Sentiment

# 3.2.1. Statistical Analysis of Public Complaint Emotion

Figure 6 shows the obtained annual average of public complaint sentiment intensity of Shandong Province as a whole and in the 17 different cities based on the complaint data results mentioned above and the sentiment analysis method for air pollution complaint data proposed in this study. Negative numbers indicate that the public's emotion is negative, that is, the intensity of the public's negative sentiment. Moreover, the smaller the number, the smaller the negative emotion and the greener the color.

| -         | 2012  | 2013  | 2014  | 2015  | 2016  | 2017  | 2018  | _     |
|-----------|-------|-------|-------|-------|-------|-------|-------|-------|
| Shandong  | -6.54 | -6.74 | -6.63 | -6.62 | -6.55 | -6.22 | -6.16 |       |
| Dongying  | -7.49 | -7    | -6.97 | -7.11 | -6.93 | -6.41 | -5.47 |       |
| Linyi     | -6.4  | -6.9  | -6.86 | -6.21 | -6.33 | -6.11 | -5.96 |       |
| Weihai    | -6.6  | -6.33 | -6.47 | -7.04 | -6.51 | -6.43 | -6.51 |       |
| Dezhou    | -6.55 | -6.28 | -6.69 | -6.63 | -6.76 | -6.08 | -6.07 | -5.47 |
| Rizhao    | -6.39 | -6.57 | -6.75 | -6.66 | -6.62 | -6.1  | -5.57 |       |
| Zaozhuang | -6.6  | -6.42 | -6.63 | -6.71 | -6.84 | -6.31 | -5.89 |       |
| Taian     | -6.78 | -6.77 | -6.98 | -6.86 | -6.5  | -6.36 | -6.51 |       |
| Jinan     | -6.3  | -6.53 | -6.65 | -6.59 | -6.1  | -5.81 | -6.19 |       |
| Jining    | -6.66 | -7.49 | -6.88 | -6.71 | -6.51 | -6.56 | -6.35 |       |
| Zibo      | -6.19 | -6.84 | -6.59 | -6.35 | -6.61 | -6.14 | -5.89 |       |
| Binzhou   | -6.41 | -6.8  | -6.92 | -6.82 | -6.57 | -6.12 | -6.35 |       |
| Weifang   | -6.31 | -6.78 | -6.4  | -6.45 | -6.62 | -6.22 | -6.3  |       |
| Yantai    | -6.77 | -6.9  | -6.5  | -6.43 | -6.59 | -6.47 | -6.3  |       |
| Liaocheng | -5.77 | -6.86 | -6.56 | -6.41 | -6.36 | -5.83 | -6.3  | -7.49 |
| Laiwu     | -6.67 | -6.62 | -6.07 | -6.29 | -5.86 | -6.28 | -6.76 |       |
| Heze      | -6.71 | -6.77 | -6.55 | -6.89 | -7.11 | -6.26 | -5.86 | 1     |
| Qingdao   | -6.41 | -6.78 | -6.44 | -6.68 | -6.54 | -6.32 | -6.43 |       |

**Figure 6.** Analysis on the interannual change of public complaint emotional intensity of air pollution in Shandong Province from 2012 to 2018.

The obtained results indicated that the public's negative emotional intensity on air pollution in Shandong Province weakened as a whole from 2012 to 2018, especially in 2017 and 2018 (Figure 6). According to Figure 2, the AQI, PM2.5, and PM10 of Shandong Province showed a downward trend year by year from 2014 to 2018. Therefore, the weakening of public complaints' negative emotion intensity can be attributed to the improvement of air quality. Italso reflects that the Shandong provincial government has made some achievements in air pollution control in recent years.

#### 3.2.2. Spatial Pattern of Public Complaint Emotion

This study used the hot spot analysis method to analyze the emotional intensity of public complaints. The obtained results indicated that the public's complaint emotion intensity was negative (negative emotion). Thus, we took the absolute value of complaint emotion intensity as hot spot analysis to facilitate understanding. It is worth noting that the larger the absolute value, the stronger the negative intensity. The obtained results are



shown in Figure 7. Red indicates a high value (high negative emotion intensity) aggregation area, while blue indicates a low value (low negative emotion intensity) aggregation area.

Figure 7. Hot spot analysis of public complaints about air pollution in Shandong Province from 2012 to 2018.

The results obtained in this study indicate that although the number of complaints in Jinan, Laiwu, Liaocheng, and Weifang was relatively high (Figure 4b), the intensity of negative public emotion towards air pollution was relatively low (Figure 7). Strong negative emotions towards air pollution were observed in public in Jining, Heze, and Dongying cities, especially in Dongying. These results indicate that there are differences between the emotional intensity and the number of complaints. Therefore, a large number of complaints does not necessarily mean that the emotional intensity of public complaints is high.

#### 3.3. Correlation Analysis

To further explore whether the number and emotional intensity of public complaints were associated with the change of air quality, we analyzed the Pearson correlation coefficient and its significance test between the number and emotional intensity of public complaints, and PM2.5 and PM10 in 17 cities of Shandong Province from 2014 to 2018.

The obtained results indicated that the public's complaint emotion was subjective, different, and complicated due to the noise and other interference factors in the complaint data. Therefore, this paper selects the maximum value of public complaints' emotional intensity, under the same air quality parameter of PM2.5 or PM10 level, for statistical correlation analysis. The results are shown in Table 3.

From Table 3, we can see that the correlation between the emotional intensity of public complaints and the concentration of PM2.5 was -0.733, while that of PM10 was -0.606. Besides, the correlation between the number of public complaints and the concentration of PM2.5 was -0.718, while that of PM10 was -0.735.

These results show that, firstly, the intensity of public complaint emotion was inversely proportional to the concentration of PM2.5 and PM10. That is, the intensity of public negative complaint emotion was stronger with the deterioration of air quality. Secondly, the number of public complaints was inversely proportional to the concentrations of PM2.5

and PM10, which indicates that the public was more likely to complain about potential air pollution when the air index was good. However, the number of public complaints was relatively reduced when the air quality was worse. That can be attributed to the public keenly observing air pollution-related phenomena when the air index is good. Thus, they are often worried that this phenomenon might damage the air quality. Therefore, they actively complain to alert the government departments, who may not know about it, thereby preventing the slightest failure. However, when the air index exceeds the standard, it indicates that the monitoring station has already detected the existing pollution problem. Moreover, the government and the public are also aware of the air pollution problem in the area, so the number of public complaints decreases.

tional intensity.
Correlations
Emotion Score Number of Complaints

Table 3. Correlation analysis of PM2.5 and PM10 with the number of public complaints and emo-

|       |                     | <b>Emotion Score</b> | Number of Complaints |
|-------|---------------------|----------------------|----------------------|
|       | Pearson correlation | -0.733 **            | -0.718 **            |
| PM2.5 | Sig. 2-tailed       | 0.000                | 0.000                |
|       | N                   | 301                  | 301                  |
|       | Pearson correlation | -0.606 **            | -0.735 **            |
| PM10  | Sig. 2-tailed       | 0.000                | 0.000                |
|       | Ň                   | 419                  | 419                  |

\*\* Pearson correlation coefficient was significant at p < 0.01.

Moreover, the negative emotional intensity of public complaints is more correlated with PM2.5 when compared with PM10. This shows that PM2.5 has a more significant emotional impact on the public. The results shown in Figure 2 indicate that the number of days that PM2.5 exceeded the standard in Shandong Province between 2014 and 2018 were more than the number of days that PM10 exceeded the standard. That indicates that PM2.5 pollution is more severe than PM10 pollution, and the public's emotional intensity for PM2.5 is also more negative.

#### 4. Discussion

The results obtained in this study indicate that the number of public complaints about air pollution and the intensity of negative emotions, which were -0.7 and -0.73, respectively, were significantly negatively correlated with PM2.5. It indicates that as the concentration of PM2.5 increases, the intensity of public complaints' negative emotions also increases, and the negative emotions are strong. However, the number of public complaints decreases. It is easy to understand why the public's negative emotions become stronger when the air quality deteriorates. However, it is hard to know why the number of complaints was inversely proportional to the concentration of PM2.5. That also reflects the deviation between the public's subjective response to air pollution and actual air pollution indicators [34].

We believe that this deviation is, on the one hand, determined by the public's subjectivity and is associated with the public's psychology. For example, when the air index is good, the public will easily complain, out of precautionary psychology, about some phenomena they observe that is associated with air pollution, even if the phenomena do not exceed the standard. They will actively complain because they are worried that this phenomenon may not be monitored. However, when the air index exceeds the standard, both government departments and the public know the air pollution problem, and thus the number of public complaints is relatively reduced. On the other hand, it may also be associated with the location of the air quality monitoring station. It is worth noting that the monitoring range of air quality monitoring stations is limited [35], while the public is everywhere. Therefore, some things that the public sees and smells may not be detected by the monitoring stations [36]. This study aimed at determining what the public sees and smells before they choose to complain. That was possible because the problems associated with air pollution are mainly reflected in the complaint data. Therefore, we used word cloud to analyze further the title and content of air pollution complaint data in Shandong Province. A word cloud is a visual display of the "Keywords" with a high frequency in the text. The words with a high frequency will be presented in a larger font, while the words with a low frequency will be presented in a smaller font [37]. The word cloud image filters out a large amount of low-frequency and low-quality text information, so that the viewer can appreciate the theme of the text as soon as they scan the text [38].

The obtained results are shown in Figure 8. At the same time, the frequency of the main words is also analyzed. Some useless but high-frequency words such as "De, Le, Zai, yes, we, you, and he" are filtered in order to analyze the results better.



Figure 8. Word cloud analysis of the title (a) and content (b) of Shandong public complaint data.

Vocabulary such as "emission", "production", "enterprise", and "exhaust gas" indicate that public complaints about air pollution mainly reflect the polluting gas emissions of some enterprises and factories. At the same time, the public's perception of air pollution is primarily reflected in visual and in terms of smell such as "emission", "dust", "exhaust gas", "odor", "pungent", and other words. Enterprises and factories are indeed prone to produce large amounts of gas, which the public can see and smell. Besides, the terms "villager" and "village" also reflect that most of the complaint's location may be in the suburbs. That is because of the limited monitoring scope because most air quality monitoring stations are located in urban areas. Therefore, it is easy to overlook the potential sources of pollution in the villages.

### 5. Conclusions

This study's main aim was to use air pollution complaint data to analyze the public's response and perception of air quality. The analytic framework for the public to respond to air pollution was proposed based on online complaint data and sentiment analysis. The framework was applied to 13,469 air pollution-related online complaint data from 2012 to 2018 in Shandong Province. The main conclusions of this study include:

- The public's perception of air pollution is mainly reflected in the sense of vision and smell, and the content of complaints focuses on the emission problems of enterprises and factories.
- (2) The number and emotional intensity of public air pollution complaint data, which were -0.7 and -0.73, respectively, were negatively correlated with PM2.5. It means that the number of complaints is large when the air quality is good, but the public's negative emotion is not strong. Moreover, the number of public complaints is reduced when the air quality deteriorates, and the negative emotion is stronger.
- (3) The correlation between public emotional intensity and PM2.5 was higher in Shandong Province than in PM10. Furthermore, the analysis of PM2.5 and PM10 over standard data confirmed that PM2.5 pollution is more severe than PM10 pollution in Shandong Province.

(4) The air quality of Shandong Province improved significantly between 2014 and 2018, and the number of public complaints and the emotional intensity of negative complaints also showed a downward trend.

In summary, although the response of the public to air pollution, such as the number of complaints, deviates from the actual air pollution monitoring data to a certain extent, it also has a very significant correlation. The emotional intensity of public complaints can directly reflect the changing trend of air pollution. Besides, the intensity of public emotion in the complaint data was inversely proportional to the air quality. It indicates that the worse the air quality, the stronger the negative emotion of public complaint.

The quantitative analysis of public response to air pollution in this study cannot only help us find the changing trend of air pollution and potential pollution problems, but also test the effectiveness of government air pollution control to a certain extent. It has important supervision value for improving air quality. However, future research should integrate various data, in-depth analysis of the pollution source problems reflected in the complaint data, and the factors affecting the public's complaint emotion.

**Author Contributions:** Conceptualization, Yong Sun; Data curation, Yong Sun and Min Ji; Formal analysis, Fengxiang Jin and Huimeng Wang; Investigation, Yong Sun and Fengxiang Jin; Methodology, Yong Sun and Huimeng Wang; Project administration, Min Ji and Fengxiang Jin; Resources, Min Ji; Supervision, Fengxiang Jin; Writing—original draft, Yong Sun and Huimeng Wang; Writing review and editing, Huimeng Wang. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the Major Scientific and Technological Innovation Projects in Shandong Province (2019JZZY020103) and a grant from State Key Laboratory of Resources and Environmental Information System.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

**Data Availability Statement:** The Air Quality Index (AQI) data was obtained from China's Ministry of Environmental Protection at (http://106.37.208.233:20035/, accessed on 25 February 2021).

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

#### References

- 1. Darçın, M. Association between air quality and quality of life. *Environ. Sci. Pollut. Res.* 2014, 21, 1954–1959. [CrossRef]
- Zhan, D.; Kwan, M.-P.; Zhang, W.; Yu, X.F.; Meng, B.; Liu, Q. The driving factors of air quality index in China. J. Clean. Prod. 2018, 197, 1342–1351. [CrossRef]
- Carducci, A.; Donzelli, G.; Cioni, L. Air pollution: A study of citizen's attitudes and behaviors using different information sources. *Epidemiol. Biostat. Public Health* 2017, 14, 1–9.
- 4. Sun, Y.; Moshfeghi, Y.; Liu, Z. Exploiting crowdsourced geographic information and GIS for assessment of air pollution exposure during active travel. *J. Transp. Health* **2017**, *6*, 93–104. [CrossRef]
- 5. Wang, Y.; Fu, X.; Jiang, W.; Wang, T.; Tsou, M.; Yue, X. Inferring urban air quality based on social media. *Comput. Environ. Urban Syst.* **2017**, *66*, 110–116. [CrossRef]
- 6. Zhou, Z.; Guo, X.; Wu, H.; Yu, J. Evaluating air quality in China based on daily data: Application of integer data envelopment analysis. *J. Clean. Prod.* **2018**, *198*, 304–311. [CrossRef]
- Pu, S.; Shao, Z.; Fang, M.; Yang, L.; Liu, R. Spatial distribution of the public's risk perception for air pollution: A nationwide study in China. *Sci. Total Environ.* 2019, 655, 454–462. [CrossRef] [PubMed]
- Song, Y.; Zhou, A.; Zhang, M. Exploring the effect of subjective air pollution on happiness in China. *Environ. Sci. Pollut. Res.* 2020, 27, 43299–43311. [CrossRef]
- 9. Li, W.; Fengying, Z.; Eva, P. Taking Action on Air Pollution Control in the Beijing-Tianjin-Hebei (BTH) Region: Progress, Challenges and Opportunities. *Int. J. Environ. Res. Public Health* **2018**, *15*, 306.
- 10. Izabela, S.; Paciorek, M.; Skotak, K. The Analysis of the Effectiveness of Implementing Emission Reduction Measures in Improving Air Quality and Health of the Residents of a Selected Area of the Lower Silesian Voivodship. *Energies* **2020**, *13*, 4001.
- 11. Yang, R.; Wei, J.; Lu, L. Air pollution and green consumption of consumers in China's urban areas: A norm activation perspective. *Hum. Ecol. Risk Assess.* **2019**, *26*, 1988–2010. [CrossRef]
- 12. Dons, E.; Laeremans, M.; Anaya-Boig, E. Concern over health effects of air pollution is associated to NO2 in seven European cities. *Air Qual. Atmos. Health* **2018**, *11*, 591–599. [CrossRef]

- David, V.; Mariaelena, G.; Paul, B.; Hamed, G.; Ricardo, C. Perceptions about air quality of individuals who work outdoors in the San Joaquin Valley, California. *Atmos. Pollut. Res.* 2020, *11*, 825–830.
- Wang, Y.; Tsou, M.; Fu, X. Using Social Media to Detect Outdoor Air Pollution and Monitor Air Quality Index (AQI): A Geo-Targeted Spatiotemporal Analysis Framework with Sina Weibo (Chinese Twitter). *PLoS ONE* 2015, 10, e0141185.
- Gurajala, S.; Dhaniyala, S.; Matthews, J. Understanding Public Response to Air Quality Using Tweet Analysis. Soc. Media + Soc. 2019, 5, 205630511986765. [CrossRef]
- Hswen, Y.; Qin, Q.; John, S.; Brownstein, J.; Hawkins, B. Feasibility of using social media to monitor outdoor air pollution in London, England. *Prev. Med.* 2019, 121, 86–93. [CrossRef] [PubMed]
- 17. Huang, Y.; Fei, T.; Kwan, M.-P.; Kang, Y. GIS-Based Emotional Computing: A Review of Quantitative Approaches to Measure the Emotion Layer of Human–Environment Relationships. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 551. [CrossRef]
- 18. Gallagher, D.L.; Dietrich, A.M. Satistical approaches for analyzing customer complaint data to assess aesthetic episodes in drinking water. *J. Water Supply Res. Technol.-AQUA* **2014**, *63*, 358–367. [CrossRef]
- 19. Avaliani, S.L.; Balter, B.M.; Balter, D.B. Air pollution source identification from odor complaint data. *Air Qual. Atmos. Health* **2015**, *9*, 179–192. [CrossRef]
- Pan, G.; Wang, B.; Guo, S.; Zhang, W. Statistical analysis of sewer odour based on 10-year complaint data. *Water Sci. Technol.* 2020, 81, 1221–1230. [CrossRef] [PubMed]
- 21. Sailunaz, K.; Alhajj, R. Emotion and sentiment analysis from Twitter text. J. Comput. Sci. 2019, 36, 101003. [CrossRef]
- 22. Wang, Z.; Ho, S.B.; Cambria, E. A review of emotion sensing: Categorization models and algorithms. *Multimed. Tools Appl.* **2020**, 79, 35553–35582. [CrossRef]
- 23. Chen, S.; Lv, X.; Gou, J. Personalized Recommendation Model: An Online Comment Sentiment Based Analysis. *Int. J. Comput. Commun. Control* 2020, *15*, 1004. [CrossRef]
- 24. Feng, S.; Wang, D.; Yu, G.; Gao, W.; Wong, K.F. Extracting common emotions from blogs based on fine-grained sentiment clustering. *Knowl. Inf. Syst.* 2011, 27, 281–302. [CrossRef]
- 25. Despoina, A.; Paraskevi, F.; Sotiris, I. A survey of Twitter research: Data model, graph structure, sentiment analysis and attacks—ScienceDirect. *Expert Syst. Appl.* 2020, *164*, 114006.
- 26. Choi, H.-J.; Park, C.H. Emerging topic detection in twitter stream based on high utility pattern mining. *Expert Syst. Appl.* **2019**, 115, 27–36. [CrossRef]
- Yang, L.; Lin, H. Construction and Application of Chinese Emotional Corpus. In Proceedings of the 13th Chinese Conference on Chinese Lexical Semantics, Wuhan, China, 6–8 July 2012; pp. 122–133.
- Ali, Y.; Ameneh, G.; Zaiane, O. Current State of Text Sentiment Analysis from Opinion to Emotion Mining. ACM Comput. Surv. 2017, 50, 1–33.
- 29. Goldberg, W.D.; Wilson, J.P.; Knoblock, C.A. From text to geographic coordinates: The current state of geocoding. J. Urban Reg. Inf. Syst. Assoc. 2007, 19, 33–46.
- Lin, Y.; Kang, M.J.; Wu, Y.Y.; Du, Q.Y.; Liu, T. A deep learning architecture for semantic address matching. *Int. J. Geogr. Inf. Sci.* 2020, 34, 559–576. [CrossRef]
- 31. Getis, A.; Ord, J.K. The Analysis of Spatial Association by Use of Distance Statistics. Geogr. Anal. 1992, 24, 189–206. [CrossRef]
- 32. Freedman, D.; Pisani, R.; Purves, R. Statistics; Norton & Company: New York, NY, USA, 1998; p. 148.
- 33. Li, J.; Sun, Z. Urban Function Orientation Based on Spatiotemporal Differences and Driving Factors of Urban Construction Land. *J. Urban Plan. Dev.* **2020**, *146*, 05020011. [CrossRef]
- 34. Liu, L.; He, P.; Zhang, B.; Bi, J. Red and Green: Public Perception and Air Quality Information in Urban China. *Environment* **2012**, 54, 44–49. [CrossRef]
- 35. Xian, J.; Sun, D.; Xu, W.; Han, W.; Yang, S. Urban air pollution monitoring using scanning Lidar. *Environ. Pollut.* **2019**, 258, 113696. [CrossRef] [PubMed]
- 36. Ryu, Y.H.; Min, S.K. What matters in public perception and awareness of air quality? Quantitative assessment using internet search volume data. *Environ. Res. Lett.* **2020**, *15*, 0940b4. [CrossRef]
- 37. Cui, W.; Wu, Y.; Liu, S.; Wei, F.; Zhou, M.; Qu, H. Context-Preserving, Dynamic Word Cloud Visualization. *IEEE Comput. Graph. Appl.* **2010**, *30*, 42–53. [CrossRef] [PubMed]
- Heimerl, F.; Lohmann, S.; Lange, S.; Ertl, T. Word Cloud Explorer: Text Analytics Based on Word Clouds. In Proceedings of the 2014 47th Hawaii International Conference on System Sciences, Waikoloa, HI, USA, 6–9 January 2014.