Exploring the convergence patterns of PM2.5 in Chinese cities

Yan Wang[1] · Yuan Gong[2] · Caiquan Bai[1] · Hong Yan[3] · Xing Yi[1]

Abstract
Economic development and ongoing urbanization are usually accompanied by severe haze pollution. Revealing the spatial and temporal evolution of haze pollution can provide a powerful tool for formulating sustainable development policies. Previous studies mostly discuss the differences in the level of PM2.5 among regions, but have paid little attention to the change rules of such differences and their clustering patterns over long periods. Therefore, from the perspective of club convergence, this study employs the log t regression test and club clustering algorithm proposed by Phillips and Sul (Econometrica 75(6):1771–1855, 2007. 10.1111/j.1468-0262.2007.00811.x) to empirically examine the convergence characteristics of PM2.5 concentrations in Chinese cities from 1998 to 2016. This study found that there was no evidence of full panel convergence, but supported one divergent group and eleven convergence clubs with large differences in mean PM2.5 concentrations and growth rates. The geographical distribution of these clubs showed significant spatial dependence. In addition, certain meteorological and socio-economic factors predominantly determined the convergence club for each city.

✉ Xing Yi
  yixing0111@mail.sdu.edu.cn

  Yan Wang
  yan_wang@mail.sdu.edu.cn

  Yuan Gong
  gongyuan@ruc.edu.cn

  Caiquan Bai
  baicaiquan@sdu.edu.cn

  Hong Yan
  yanhong@163.sufe.edu.cn

[1] The Center for Economic Research, Shandong University, Ji'nan 250100, Shandong, People's Republic of China

[2] School of Environment & Natural Resources, Renmin University of China, Beijing 100872, People's Republic of China

[3] School of International Relations and Public Affairs, Fudan University, Shanghai 200433, People's Republic of China

1 Introduction

Environmental degradation has become an ongoing concern across China, especially haze events (Ministry of Ecology & Environment of the People's Republic of China, 2019). According to the *2019 World Air Quality Report* released by IQ Air, 48 Chinese cities featured among the top 100 most polluted cities in the world, which was frequently due to fine particulate matter with an aerodynamic diameter ≤ 2.5 μm (PM2.5).[1] Long-term exposure to high PM2.5 concentrations is a potential health hazard (Guan et al., 2016). Additionally, air quality deterioration results in huge medical costs and negatively impacts tourism, which seriously restricts regional economic development (Becken et al., 2017; Hao et al., 2018; Yang et al., 2019). It has been estimated that the economic cost of environmental degradation has reached approximately 1–8% of the GDP in China (World Bank, 2007; Heck & Hirschberg, 2011). Thus, alleviating haze pollution has become one of the top priorities for sustainable economic development. To formulate targeted policies for haze control, it is necessary to investigate the differences in PM2.5 pollution and its long-term trends among regions.

Current studies mainly focus on the variations of PM2.5 concentrations over time, the heterogeneity among regions, and their influencing factors. Cities, provinces, and urban agglomerations with severe PM2.5 pollutions have been comprehensively explored, such as Beijing, Hefei, and Bohai Rim (Mi et al., 2019; Wang & Fang, 2016; Yan et al., 2018; Zhang et al., 2020; Zou et al., 2016). Specifically, in terms of spatial variations, PM2.5 concentrations have illustrated both heterogeneity and autocorrelation at both the national and regional levels (He et al., 2019; Yang et al., 2018). The highest PM2.5 pollution areas are mainly located in the northern plains and central regions, especially the Beijing–Tianjin–Hebei urban agglomeration (Fan et al., 2020; Ma et al., 2016). Meanwhile, PM2.5 concentrations in urban areas are significantly higher than those in rural areas (Lin et al., 2020). Moreover, PM2.5 concentrations are frequently positively correlated among neighboring areas, indicating spatial clustering characteristics in the form of "high-high" and "low-low" agglomeration centers (Shen et al., 2019).

In terms of temporal trends, PM2.5 concentration demonstrates seasonal, weekly, and diurnal variations. Pollution is much worse in winter than in summer (Fan et al., 2020; Ma et al., 2016); January and December are the worst months of the year for PM2.5 pollution, while July and August are the best (Song et al., 2017). Simultaneously, due to vehicular traffic, the concentration of PM2.5 is the highest on Monday, followed by the weekend, and is the lowest on Wednesday (Huang et al., 2015). During the day, the highest PM2.5 concentrations occur from 8 to 12 am, and the lowest occur from 3 to 5 pm (Shen et al., 2019).

In addition to spatial and temporal variations, scholars have investigated the factors that influence the regional PM2.5 pollution. Studies have shown that there are mainly two types of factors. The first are meteorological factors, such as temperature, air pressure, and wind speed (He et al., 2019; Tai et al., 2010); the second are socio-economic factors, such as land use, economic development, population growth, and urbanization (Han et al., 2018; Lin et al., 2020; Xu, Chen, et al., 2019; Xu, Miao, et al., 2019).

Although the above studies have performed comprehensive investigations, there are still some inadequacies. First, most relevant studies compared haze pollution among

---

[1] https://www.iqair.com/newsroom/air pollution far deadlier coronavirus new data ranks health threat cities worst best.

regions by directly comparing the level differences in PM2.5 concentrations during a set period. However, comparisons in this way are not only susceptible to occasional extreme meteorological conditions (Luo et al., 2019), but may also ignore change trends, making it difficult to determine convergence. For example, although the PM2.5 concentration of Longnan was always higher than that of Zhangzhou from 1998 to 2016, the continuous improvement of air quality in Longnan and the continuous increase in PM2.5 pollution in Zhangzhou led to their gradual convergence. However, if only their mean PM2.5 concentrations during this period are compared, it is difficult to find such narrowing gap. In contrast, a dynamic, long-term convergence analysis would depict it. Second, the current studies were mostly based on monitoring data from fixed observation points in cities (Song et al., 2017; Shen et al., 2019; Fan et al., 2020; Jiang et al., 2020). However, because air quality monitoring stations were not fully covered in China until 2013, the available data are temporally limited, which prevented any long-term or dynamic calculation, including convergence analysis. Notably, long-term convergence analysis of PM2.5 concentrations, especially the spatial distribution of convergence clubs, can provide a new perspective for understanding the current haze pollution. Third, previous research on club convergence usually artificially divided samples into several subgroups according to specific information, such as geographical locations or institutions; however, these artificial divisions may cause biased results. This is because this approach may divide samples that should belong to the same convergence club into different subgroups or fail to detect potential convergence clubs (Bai et al., 2021). Fortunately, this bias can be avoided by using the data-driven algorithm developed by Phillips and Sul (2007). Recent studies have applied this algorithm to various issues, such as carbon emissions (Apergis & Payne, 2017) and energy efficiency (Parker & Liddle, 2017). However, there are no studies that use this method to rigorously investigate the convergence of PM2.5 concentrations so far. Fourthly, although analyses of PM2.5 pollution in typical Chinese cities, provinces, or urban agglomerations are common, few studies have involved all cities over a long period, which may lead to bias due to sample size and data fineness, especially when exploring the factors that influence PM2.5 pollution.

For these reasons, this study conducted convergence analysis to explore the dynamic evolution of PM2.5 concentrations for 342 prefecture-level cities in China from 1998 to 2016. To this end, we first conducted the log t regression test proposed by Phillips and Sul (2007), followed by a club clustering algorithm, to empirically investigate the overall convergence and club convergence of the PM2.5 concentrations. Then, we analyzed the spatial distribution patterns and driving forces of the convergence clubs using the spatial autocorrelation test and Ordered Probit regression.

This study contributes to the literature in the following ways. First, we first explored the convergence patterns and its distribution characteristics of PM2.5 concentrations for Chinese cities, describing the trends in the differences in PM2.5 pollution among cities from a dynamic perspective. Second, by applying the nonlinear regression-based test and club clustering algorithm, the obtained convergence patterns were more reliable and robust. This is because this method allows for inter-city heterogeneity; it does not set specific assumptions about the stationarity properties of the data and avoids possible biased results that may be caused by artificial division (Du, 2017). Third, we found that PM2.5 concentrations in Chinese cities showed club convergence characteristics and were the first to investigate the factors characterizing the observed clustering patterns. Fourthly, compared with the ground-observed data in previous studies, satellite-derived data from the Socio-economic Data and Applications Center (SEDAC) are generally more accurate and have a longer time span (Van Donkelaar, 2015). Thus, it allowed us to explore the evolution

of PM2.5 concentrations in Chinese cities over a longer period, which produced a more complete picture of how haze pollution evolved during China's urbanization process.

The remainder of this paper proceeds as follows. In Sect. 2, the methods and data sources are introduced in detail. Then, in Sect. 3, the empirical results of the PM2.5 convergence tests for the full panel and within groups are described. In Sects. 4 and 5, the reasons for the observed PM2.5 convergence club classification are explored, including their geographical distribution patterns and the influencing factors contributing to the patterns. The last section concludes the paper and proposes targeted policies.

2 Method and data

2.1 Method

The methods of log t regression test and a club clustering algorithm proposed by Phillips and Sul (2007) were used to explore the convergence of PM2.5 concentrations in Chinese cities. In general, the log t regression method tests for convergence by checking whether the cross-sectional variance of the idiosyncratic part of the interest variable eventually converges to zero. Compared with classical convergence tests, this method has numerous advantages. First, because it also works under the possibility of transitional heterogeneity or even transitional divergence, long-term common fluctuation will always be captured even in the absence of cointegration (Panopoulou & Pantelidis, 2009). Second, this methodology does not set specific assumptions about the smoothness trends or stochastic non-stationarity of the data (Du, 2017). Third, as mentioned above, by repeatedly conducting the log t regression test for nearly all the combinations of cities in a specific order, its identification of club convergence not only avoids potentially biased results caused by artificial division, but also can detect the subgroup with the most obvious convergence characteristics.

2.1.1 log t regression test

According to the convergence test method proposed by Phillips and Sul (2007), the PM2.5 concentration of city i during year t, as denoted by x_{it}, can be decomposed as follows:

$$x_{it} = \delta_{it}\mu_t \tag{1}$$

where μ_t represents the common trend section of the data that describes some stochastic or deterministic trending behaviors, and δ_{it} measures the evolution of idiosyncratic elements over time. In Eq. (1), a long-run equilibrium in x_{it} is equivalent to the convergent form of δ_{it}, namely $\lim_{t\to\infty}\delta_{it} = \delta$ for any i. To eliminate the common factor μ_t, x_{it} is scaled down to the h_{it}.

$$h_{it} = \frac{x_{it}}{\frac{1}{N}\sum_{i=1}^{N}x_{it}} = \frac{\delta_{it}}{\frac{1}{N}\sum_{i=1}^{N}\delta_{it}} \tag{2}$$

Thus, if δ_{it} converges to δ, the cross-sectional variance of h_{it}, that is $\eta_t = \frac{1}{N}\left(h_{it} - 1\right)^2$, will eventually converge to zero. Using this property, a log t regression can be constructed as follows:

$$\log\left(\frac{\eta_1}{\eta_t}\right) - 2\log\left[\log(t+1)\right] = \widehat{a} + \widehat{b}\log(t) + \widehat{\varepsilon}_t \tag{3}$$

Under some regularity assumptions, the robust t statistic for the coefficient $\widehat{b}$ is constructed as follows:

$$t_{\widehat{b}} = \frac{\widehat{b} - b}{s_b} \tag{4}$$

where s_b is the long-run heteroskedasticity and autocorrelation consistent standard error. For the one-sided test at the 5% level, the null hypothesis of convergence for the whole sample should be accepted if $t_{\widehat{b}} > -1.65$; otherwise, it is rejected.

2.1.2 Club clustering algorithm

Phillips and Sul (2007) proposed a data-driven algorithm for identifying clusters of samples with homogeneous convergence behaviors. First, a core group that consists of the cities with the highest t statistic is determined. Adding one city at a time to the core group, the log t regression test is repeatedly performed to decide which cities could be included in the core group. This updated core group is the first convergent subgroup. Then, the log t regression test is performed among the remaining cities screened above. If the convergence cannot be rejected, this subgroup forms the second convergence club; otherwise, the steps above are repeated to detect smaller convergent subgroups.[2]

Later, Schnurbus et al. (2017) improved the original algorithm in terms of club merging. To produce the minimum number of convergence clubs, the joint log t test is performed for the adjacent pairs generated above, and they are merged if the convergence hypothesis holds. This step is iterated until the resulting clubs can no longer be merged.

2.1.3 The ordered probit model

The obtained convergence clubs are successively assigned positive integer values of 1 to J, where J is the total number of convergence clubs, according to rule that the higher the PM2.5 concentration mean is among the club members, the higher the assignment will be. In this case, it is assumed that:

$$P\left(y_i = j | X_i\right) = P(r_{j-1} < X_i'\beta + \varepsilon_i < r_j | X_i) = \Phi\left(r_j - X_i'\beta\right) - \Phi\left(r_{j-1} - X_i'\beta\right) \tag{5}$$

where y_i is the convergence club type to which the city i belongs, the X_i vectors are the factors that may affect the PM2.5 concentration, r_j is the cut point, and ε_i is the disturbance term. Using the assumption that ε_i is $N(0, 1)$, the sample likelihood function and maximum likelihood estimation value can be obtained.

---

[2] For a detailed introduction of club clustering algorithm, please see Phillips and Sul (2007).

2.2 Data sources and processing

The PM2.5 concentration data in this study were derived from the global annual average PM2.5 grids database obtained by the SEDAC at Columbia University.[3] These data were obtained from the aerosol optical depth instruments of several satellites, were gridded at a $0.01° \times 0.01°$ resolution, and had dust and sea-salt removed. By matching this data to maps of China, this study determined the annual average PM2.5 concentrations of 342 Chinese prefecture-level cities from 1998 to 2016, which is a longer data set than those of other studies (Ye et al., 2018; Zhang et al., 2016).

Gridded satellite-derived data are generally more accurate and reliable tha2017n ground-observed data from urban monitoring stations, which are used in most studies (Song et al., ; Shen et al., 2019; Fan et al., 2020; Jiang et al., 2020). As non-point source data, gridded data can better capture spatial variability from a regional perspective (Lin et al., 2018). Moreover, because global monitoring can be performed from a fixed satellite, possible biases attributed to differences in the designs or operations among regional monitoring stations are avoided (Van Donkelaar, 2015). In contrast, owing to local weather, factory emissions, or population agglomeration, ground-based PM2.5 observations exhibit fluctuations that may lead to deviation during the retrieval and simulation by relatively sparse monitoring (Chen et al., 2006).

In terms of the influences on convergence clubs, this study considered that PM2.5 pollution is affected by both meteorological and socio-economic factors (He et al., 2019; Xu et al., 2020; Yun et al., 2019) to select ten possible variables, referring to the methods of Tai et al. (2010) and Xu et al. (2020). In addition, because the explained variable is the convergence club type, which is cross-sectional data, the ten explanatory variables were selected as the mean values from 2006 to 2016.[4]Detailed descriptions of the data are provided in Table 1[5].

To remove the influence of price, the actual completed amount of real estate development investment was adjusted to the real value according to its corresponding city's fixed asset investment price index for each year. Data for the above variables were obtained from the National Meteorological Science Data Center (http://data.cma.cn/) and the China City Statistics Yearbook (2006–2016).

3 Convergence test

3.1 Results of the log t test for the full panel

We first conducted a log t regression test to detect the overall convergence of PM2.5 concentrations for 342 Chinese cities.[6]The results are shown in Table 2.

---

[3] Center for International Earth Science Information Network—CIESIN—Columbia University. 2021. Annual PM2.5 Concentrations for Countries and Urban Areas, 1998–2016. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/rja8-8h89. Accessed DAY MONTH YEAR.

[4] The values of temperature and relative humidity are the annual average value from 1981 to 2010.

[5] The key terms to search including pollution, haze, PM10, PM2.5, and so forth

[6] To remove initial unstable observations, samples were selected starting at $t = [rT], [rT] + 1, \ldots, T$, where $r > 0$. Based on Monte Carlo simulation results, $r = 0.3$ was recommended for small or moderate time periods ($T \leq 50$).

Table 1 Descriptions of the influencing factors

Name	Variable	Definition	Unit
Temperature	tempe	annual mean temperature	degrees Celsius
Humidity	humid	average annual relative humidity	%
Industrial structure	industry	the percentage of the added value of secondary industry in the gross domestic product	%
Population density	popula	the number of permanent residents per unit land area	per/km^2
Land use	conratio	the percentage of constructed land in an urban area	%
Public transport utilization rate	pubtrans	annual per capita public bus passenger volume	person-time/person
Environment regulation	regula	the percentage of the words in the sentences related to environmental protection in the whole government work report	%
Energy consumption	elec	the annual electricity consumption of the whole society in the municipal district	10 billion KWH
Research and development intensity	rd	the percentage of scientific research, technical service, and geological survey employees to total employees	%
Real estate investment and construction	estate	the per capita actual completed amount of real estate development investment	1000 yuan/person

Table 2 Results of the log *t* test of full panel convergence

Variable	Coefficient	Standard error	*t* statistic	Observations
log(*t*)	− 0.9422	0.0115	− 81.9438	6498

Table 2 shows that the coefficient for the entire sample set test was negative, and the corresponding t statistic was -81.9438, which is far less than the critical value of -1.65. Thus, the null hypothesis of convergence for the full panel should be rejected at the 5% significance level, which indicates that the cross-sectional dispersion of PM2.5 did not tend to shrink over time.

Based on the results, not only were there significant differences in PM2.5 concentrations among Chinese cities, but overall, the differences did not narrow during 1998 to 2016. Therefore, the variation in PM2.5 concentrations across cities was not a contingency, but a long-term feature. According to Eq. (1), such differences could not be explained by common factors, but came from the heterogeneous factors of different cities.

3.2 Results on the club clustering

Although the data did not support full panel convergence, cities with similar economic development patterns or meteorological conditions might converge within the groups. To examine this, the empirical club clustering algorithm was used to further identify the club convergence and divide all the samples into respective clubs.

The results initially divided the 342 cities into 20 convergence clubs and one divergent group. Further, by conducting the iterative club merging algorithm proposed by Schnurbus et al. (2017), the initial 20 clubs were merged into 11 final convergence clubs. The list of cities they include are shown in Table 3.

As can be seen from Table 3, the cities converged to 11 different steady states, indicating that the PM2.5 concentrations across Chinese cities generally diverged, but followed convergence patterns within 11 subgroups.The existence of club convergence means that the differences in PM2.5 concentrations among the cities within the same club gradually shrank or even disappeared over time. In other words, PM2.5 pollution in these cities converged and showed the same trend. For example, the PM2.5 concentration in Handan in 1998 was 21.22 μg/m^3, while that of Cangzhou was 46.25 μg/m^3. Although the initial PM2.5 pollution in Handan was significantly better, it increased faster, narrowing the difference. In 2012, the difference has narrowed to less than 12 μg/m^3 and has remained roughly the same since then. Because of this narrowing gap, Handan and Cangzhou were both divided into the same club.

3.3 Statistical characteristics of the convergence clubs

Although the PM2.5 pollutions of the cities in the same club showed similar trends, the PM2.5 pollutions of the cities belonging to different clubs showed various long-term patterns. In other words, different convergence clubs exhibited significant differences. We first classified and named each convergence club according to their mean PM2.5 concentrations and growth rates and then conducted a comparative analysis. Table 4 shows the descriptive statistics of the 11 convergence clubs, where the mean and annual growth rates are calculated as the average of its member cities.

Table 3 List of cities in each convergence club

Club	Membership
Club 1 [17]	Cangzhou, Handan, Hengshui, Langfang, Xingtai, Puyang, Changchun, Dezhou, Jinan, Jining, Laiwu, Liaocheng, Linyi, Tai 'an, Zaozhuang, Zibo, Tianjin
Club 2 [64]	Bengbu, Bozhou, Chaohu, Chuzhou, Fuyang, Hefei, Huaibei, Huainan, Lu 'an, Maanshan, Tongling, Wuhu, Suzhou, Beijing, Baoding, Shijiazhuang, Tangshan, Anyang, Hebi, Jiaozuo, Kaifeng, Luohe, Pingdingshan, Shangqiu, Xinxiang, Xuchang, Zhengzhou, Zhoukou, Zhumadian, Harbin, Suihua, Jingzhou, Qianjiang, Tianmen, Wuhan, Xiantao, Xiaogan, Liaoyuan, Siping, Songyuan, Changzhou, Huaian, Lianyungang, Nanjing, Nantong, Suzhou, Taizhou, Wuxi, Suqian, Xuzhou, Yancheng, Yangzhou, Zhenjiang, Anshan, Liaoyang, Panjin, Shenyang, Tieling, Dongying, Qingdao, Rizhao, Weifang, Shanghai, Jiaxing
Club 3 [61]	Anqing, Chizhou, Xuancheng, Foshan, Guangzhou, Zhaoqing, Beihai, Chongzuo, Fangchenggang, Guigang, Guilin, Laibin, Nanning, Qinzhou, Wuzhou, Yulin, Qinhuangdao, Jiyuan, Luoyang, Nanyang, Xinyang, Daqing, Qitaihe, Ezhou, Huanggang, Huangshi, Jingmen, Suizhou, Xianning, Xiangfan, Changde, Hengyang, Loudi, Shaoyang, Xiangtan, Yiyang, Yongzhou, Yueyang, Changsha, Zhuzhou, Baicheng, Jilin, Jian, Jingdezhen, Jiujiang, Nanchang, Pingxiang, Xinyu, Yichun, Yingtan, Benxi, Dalian, Fushun, Fuxin, Jinzhou, Yantai, Yuncheng, Weinan, Chengdu, Ziyang, Huzhou
Club 4 [43]	Huangshan, Dongguan, Jiangmen, Maoming, Qingyuan, Shaoguan, Shenzhen, Yunfu, Zhongshan, Zhuhai, Hechi, Hezhou, Liuzhou, Guiyang, Qiandongnan Miao and Dong autonomous prefecture, Tongren region, Sanmenxia, Yichang, Chenzhou, Huaihua, Fuzhou, Shangrao, Huludao, Weihai, Jincheng, Changzhi, Xi 'an, Xianyang, Dazhou, Deyang, Guangan, Luzhou, Meishan, Nanchong, Neijiang, Suining, Yibin, Zigong, Hangzhou, Jinhua, Ningbo, Shaoxing, Chongqing
Club 5 [24]	Xiamen, Heyuan, Huizhou, Yangjiang, Zhanjiang, Baise, Qiannan Buyi and Miao autonomous prefecture, Zunyi, Zhangjiajie, Tonghua, Ganzhou, Chaoyang, Dandong, Jinzhong, Linfen, Taiyuan, Yangquan, Tongchuan, Leshan, Mianyang, Xishuangbanna Dai autonomous prefecture, Quzhou, Zhoushan
Club 6 [71]	Fuzhou, Longyan, Nanping, Ningde, Putian, Quanzhou, Sanming, Zhangzhou, Baiyin, Dingxi, Lanzhou, Linxia Hui autonomous prefecture, Longnan, Pingliang, Qingyang, Tianshui, Chaozhou, Jieyang, Meizhou, Shantou, Shanwei, Bijie region, Liupanshui, Buyi and Miao autonomous prefecture, Haikou, Chengde, Zhangjiakou, Hegang, Jixi, Jiamusi, Mudanjiang, Tsitsihar, Shuangyashan, Yichun, Enshi Tujia and Miao autonomous prefecture, Shennongjia forest region, Shiyan, Baishan, Yanbian Korean autonomous prefecture, Hohhot, Tongliao, Guyuan, Wuzhou, Yinchuan, Zongwei, Haidong region, Xining, Datong, Lvliang, Shuozhou, Xinzhou, Ankang, Baoji, Hanzhong, Shangluo, Yan 'an, Yulin, Bazhong, Guangyuan, Yaan, Shihezi, Dehong Dai and Jingpo autonomous prefecture, Honghe Hani and Yi autonomous prefecture, Kunming, Lincang, Pu 'er, Wenshan Zhuang and Miao autonomous prefecture, Yuxi, Lishui, Taizhou, Wenzhou
Club 7 [16]	Wuwei, Hainan, Sanya, Heihe, Chifeng, Erdos, Wuhai, Ulanqab, Hinggan League, Shizuishan, Panzhihua, Urumchi, Baoshan, Chuxiong Yi autonomous prefecture, Qujing, Zhaotong
Club 8 [13]	Gannan Tibetan autonomous prefecture, Jinchang, Zhangye, Baotou, Liangshan Yi autonomous prefecture, Aksu region, Bortala Mongolian autonomous prefecture, Changji Hui autonomous prefecture, Kashgar region, Karamay, Tacheng region, Yili Kazak autonomous prefecture, Dali Bai autonomous prefecture
Club 9 [8]	Jiayuguan, Xilin Gol League, Haibei Tibetan autonomous prefecture, Hainan Tibetan autonomous prefecture, Huangnan Tibetan autonomous prefecture, Turpan region, Lijiang, Nujiang Lisu autonomous prefecture
Club 10 [11]	Daxinganling region, Bayannur, Hulunbeier, Haxi Mongolian and Tibetan autonomous prefecture, Aba Tibetan and Qiang autonomous prefecture, Xigaze region, Altay region, Bayingguo Leng Mongolian autonomous prefecture, Hetian region, Kizl Sukirgiz autonomous prefecture, Diqing Tibetan autonomous prefecture

Table 3 (continued)

Club	Membership
Club 11 [11]	Alxa League, Guoluo Tibetan autonomous prefecture, Yushu Tibetan autonomous prefecture, Ganzi Tibetan autonomous prefecture, Ali region, Qamdo region, Lhasa, Nyingchi, Nagqu region, Shannan region, Hami region
Divergent Group [3]	Jiuquan, Binzhou, Heze

The values in square brackets are the number of club members

According to the Chinese Ambient Air Quality Standards (GB 3095–2012) as well as the relative values of the 11 clubs, we took 30 μg/m^3, 15 μg/m^3, 10 μg/m^3, 7 μg/m^3, and 5 μg/m^3 as the critical points to divide them into six different concentration levels, namely high, relatively high, medium, relatively low, low, and very low. Similarly, we took 4%, 3%, 2%, and 1.5% as the critical points to divide the clubs into five different growth rate levels, namely high, relatively high, medium, relatively low, and low. Consequently, Club 1 to Club 11 were successively named as shown in the column (6) of Table 4, and the column (7) is the abbreviation. Actually, the critical points for naming clubs were ultimately unimportant because the following analyses were based on the relative values of the clubs, rather than the absolute values.

There were some findings from comparing. First, in terms of the differences among these clubs, the difference between the two with the highest and lowest mean concentrations was approximately 60 μg/m^3, and the highest growth rate was more than three times that of the lowest. This indicates that the common trends display very distinctive patterns among convergence clubs. The significant differences between each convergent steady state also suggest that there are some key roles in characterizing the PM2.5 convergence. In addition, all convergence clubs displayed an upward trend.

Second, the number of cities in the four clubs with high concentration levels accounted for more than 50%, which means that more than half of the cities in China were converging to clubs with high PM2.5 concentration from 1998 to 2016. Therefore, PM2.5 pollution remains a concern.

Finally, considering the relationship between the mean and annual growth rate, although the mean PM2.5 concentrations decreased successively from Club H-H to Club VL-H, the growth rates showed a trend of increase and then decrease. Specifically, the annual growth rates of the top two clubs with the highest PM2.5 concentrations, Club H-H and Club H-RH, were both above 3%; thus, the PM2.5 pollutions of the cities in these two clubs have rapidly worsened from 1998 to 2016. As the climate did not change much during that short period, this rapid increase was likely attributed to social and economic factors. Moreover, the annual growth rates of the two clubs with the lowest PM2.5 concentrations, Club L-RH and Club VL-H, were also above 3%, indicating that low PM2.5 concentrations are not always accompanied by low growth rates. Although their current PM2.5 concentrations were favorable, their high growth rates may result in a worsening air quality over time, if effective measures are not taken.

Table 4 Statistical characteristics of the convergence clubs

Club	Obs	Mean	Average annual growth rate (%)	SD	Name	Abbr
Club 1 [17]	323	59.58	4.16	15.12	high PM2.5 concentrations and high growth rates club	Club H-H
Club 2 [64]	1216	49.05	3.04	12.16	high PM2.5 concentrations and relatively high growth rates club	Club H-RH
Club 3 [61]	1159	35.66	2.31	9.52	high PM2.5 concentrations and medium growth rates club	Club H-M
Club 4 [43]	817	31.53	1.38	7.21	high PM2.5 concentrations and low growth rates club	Club H-L
Club 5 [24]	456	26.17	1.59	5.30	relatively high PM2.5 concentrations and relatively low growth rates club	Club RH-RL
Club 6 [71]	1349	20.45	1.36	4.59	relatively high PM2.5 concentrations and low growth rates club	Club RH-L
Club 7 [16]	304	13.01	1.94	2.99	medium PM2.5 concentrations and relatively low growth rates club	Club M-RL
Club 8 [13]	247	9.64	2.33	2.13	relatively low PM2.5 concentrations and medium growth rates club	Club RL-M
Club 9 [8]	152	8.09	1.56	1.55	relatively low PM2.5 concentrations and relatively low growth rates club	Club RL-RL
Club 10 [11]	209	6.24	3.43	2.05	low PM2.5 concentrations and relatively high growth rates club	Club L-RH
Club 11 [11]	209	3.15	4.50	1.24	very low PM2.5 concentrations and high growth rates club	Club VL-H

The values in square brackets are the number of club members; Obs. represents the total number of observations
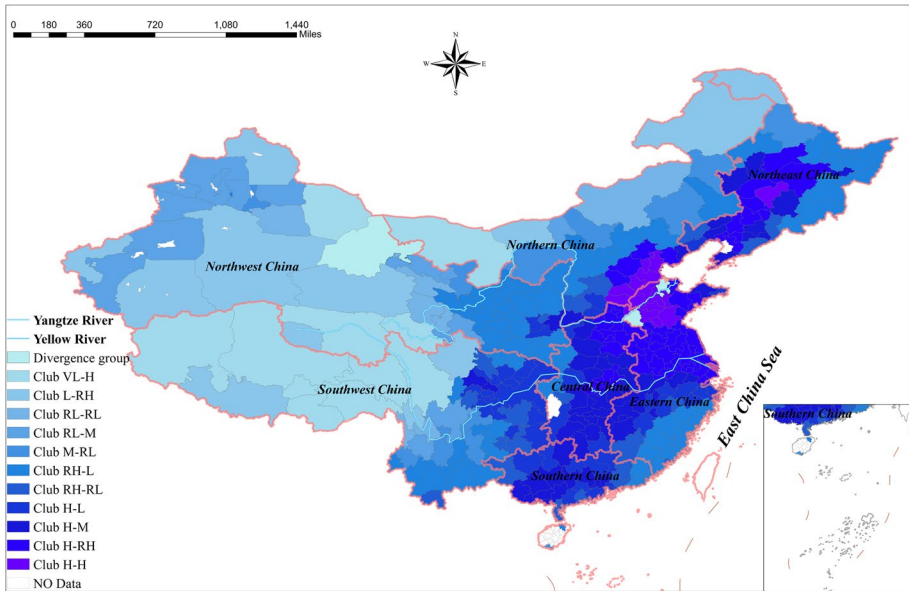
Fig. 1 The distribution of convergence clubs and the divergent group

4 Analysis of the spatial distribution of convergence clubs

The above results confirm that the PM2.5 concentrations in Chinese cities showed club convergence, and the 342 cities were divided into 11 convergence clubs and one divergent group. This means that cities belonging to the same club had become increasingly similar over time in terms of haze pollution, and changes in their PM2.5 concentrations could be explained by common factors. It is important to determine what these common factors are and explore the reasons behind the current club divisions. We first examined the connections among cities from the perspective of geographical distribution.

4.1 Geographic distribution of the convergence clubs

Figure 1 shows the spatial distribution of the 11 convergence clubs and one divergent group; it shows that the cities belonging to the same club tend to be geographically clustered. Specifically, members of the four clubs with high PM2.5 concentrations (Club H-H, Club H-RH, Club H-M, Club H-L) were mainly located in the central and northern parts of East China, the southern part of North China, the central part of Northeast China, and throughout Central China, especially the areas between the middle and lower Yellow River and Yangtze River. Members of the two clubs with low or very low PM2.5 concentrations (Club L-RH and Club VL-H) were mainly in Qinghai–Tibet Region and the central part of Northwest China. Therefore, from the perspective of long-term convergence, the PM2.5 pollution in Chinese cities showed a marked spatial distribution pattern in which the east was more serious than the west and the north was more serious than the south, which is consistent with the conclusions of previous studies (Fan
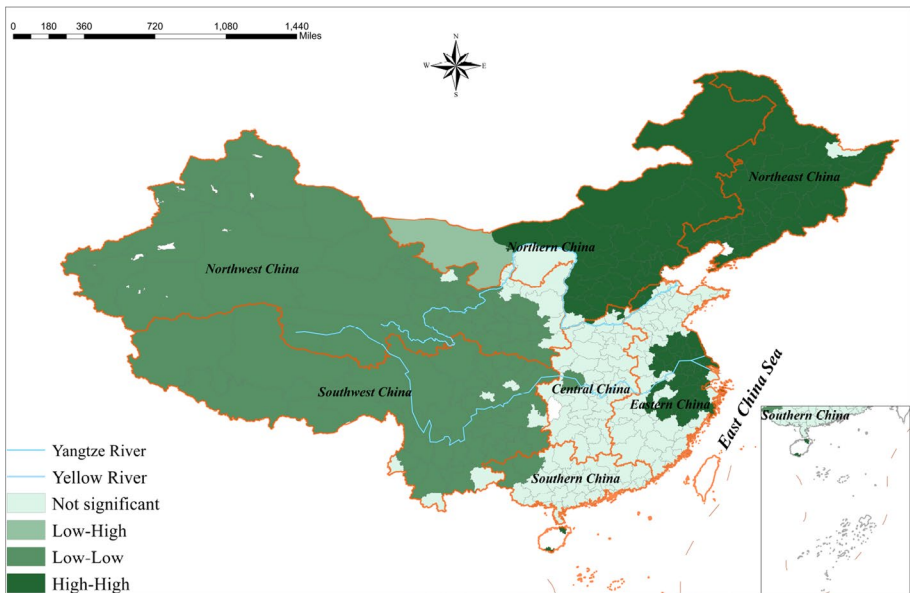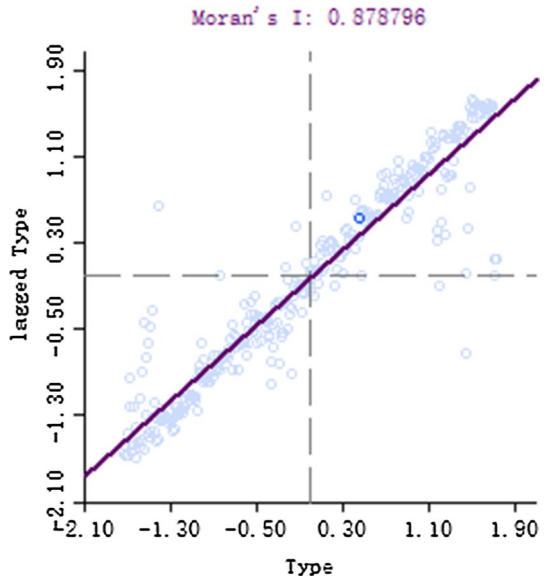
Fig. 2 Moran scatter plot

Fig. 3 Local indicators of spatial association map

et al., 2020; Ma et al., 2016; Wang et al., 2017). At a province level, Shandong was the province with the most members of Club H-H. By contrast, all cities in Tibet belonged to Club L-RH or Club VL-H, and their average annual PM2.5 concentrations were all below 4 µg/m³. Qamdo prefecture was the cleanest city in China.

4.2 Spatial correlation analysis

To further investigate the distribution characteristics of the convergence clubs, an autocorrelation test was conducted on their spatial dependence using the GeoDa software (Center for Spatial Data Science, University of Chicago, Chicago). Specifically, the global Moran's I index of the convergence clubs was first calculated. Then, to intuitively show the main types of local spatial agglomeration patterns, the Moran scatter plot for all cities was drawn (Fig. 2). Finally, to clearly display their specific geographic locations, a map of local indicators of spatial association was drawn (Fig. 3).

By constructing a spatial adjacency matrix, the obtained global Moran's I value of the convergence club types in all cities was 0.879, and the P value was 0.001, which indicates a positive correlation at the 1% significance level. In other words, the geographic distribution of the PM2.5 convergence clubs was not random, but showed significant agglomeration and spatial dependence.

As shown in Fig. 2, almost all cities were distributed in the first or third quadrants, indicating that the club convergence mainly presented high-high and low-low clusters with a few high-low and low-high outliers. Figure 3 shows that, except for Central China, Southern China, and the western part of Eastern China, the remaining regions evidenced spatial dependence and spatial spillover effects. Specifically, the low-low cluster types were mainly distributed in Southwest China and Northwest China, while the high-high cluster types were mainly distributed in Northeast China, Northern China, and the Yangtze River Delta. The increasing interconnection of economic activities between cities might be the reason of spatial autocorrelation, especially in developed areas. In fact, due to the increasingly frequent movement of labor and the radiating effects of economic activities among regions, classifying cities by their functional urban areas rather than by their administrative boundaries has become widely used (Ma & Long, 2020).

4.3 Regional distribution difference

Despite the global agglomeration and spatial dependence of the convergence clubs, when comparing the pollution differences among cities in each province, regional heterogeneity emerged. Table 5 displays the clubs of the cities in each province.

Table 5 shows that more than 75% of the cities in Jiangsu, Fujian, Tibet, and Hunan provinces fell into the same club, while cities in Sichuan, Yunnan, Guangdong, Heilongjiang, and Guizhou provinces were almost evenly distributed among different clubs. This indicates that in some provinces, PM2.5 pollutions among cities were relatively similar or gradually converging and thus, can be explained by same factors. However, in other provinces, PM2.5 concentrations varied greatly among cities, and different cities converged to different steady states. The possible reasons for this phenomenon are as follows: First, due to different geomorphic features, the degree of differences of meteorological conditions in constituent cities vary greatly across provinces. Since meteorological conditions are an important factor affecting PM2.5 (Chen et al., 2020), it may have caused the different dispersion of clubs among province. For example, the natural environment of Guizhou Province is complex and diverse, including different landforms such as plateaus, mountain basins, and river terraces, which results in significantly different meteorological conditions and then different PM2.5 among its constituent cities despite a close location. By contrast, the meteorological conditions of 13 cities in Jiangsu Province are relatively similar because its topography is mainly plain. Thus,

Table 5 The clubs of the cities in each province

Province	Club H–H	Club H-RH	Club H-M	Club H–L	Club RH-RL	Club RH-L	Club M-RL	Club RL-M	Club RL-RL	Club L-RH	Club VL-H
Shandong [17]	9	4	1								
Hebei [11]	5	3	1			2					
Henan [18]	1	12	4								
Jilin [9]	1	3	2	1		2					
Tianjin	1										
Anhui [17]		13	3	1							
Jiangsu [13]		13									
Hubei [17]		6	7	1		3					
Liaoning [13]		5	5	1		2					
Beijing		1									
Shanghai		1									
Heilongjiang [13]		2	2			7	1			1	
Zhejiang [11]		1	1	2		3					
Guangxi [14]			10	3	1						
Hunan [13]			10	2	1						
Jiangxi [11]			8	2	1						
Guangdong [21]			3	9	4	5					
Sichuan [21]			2	10	2	3	1	1		1	
Shanxi [11]			1	2	4	4					
Shaanxi [10]			1	2	1	6					
Guizhou [9]			3	3	3	3					
Chongqing			1								
Fujian [9]					1	8					
Yunnan [16]					1	7	4	2		1	
Gansu [14]						8	1	3			
Ningxia [5]						4	1	1			
Inner Mongolia [12]						2	5	1		2	1

Table 5 (continued)

Province	Club H–H	Club H-RH	Club H-M	Club H–L	Club RH-RL	Club RH-L	Club M-RL	Club RL-M	Club RL-RL	Club L-RH	Club VL-H
Qinghai [8]						2		3		1	2
Xinjiang [15]						1	7	1		4	1
Hainan [3]						1	2				
Tibet [7]										1	6

The values in square brackets are the number of cities

pollutions among these cities did not differ due to local climate. Second, since the dispersions of economic development vary among provinces and economic development is an important factor affecting PM2.5 pollution (Xu & Lin, 2016; Xu, Miao, et al., 2019), the different dispersion of club may be explained by the different dispersion of economic development.

In addition, there were some singularities where the PM2.5 concentrations or growth rates were significantly different from the surrounding cities. As shown in Fig. 1, this included cities with significantly more severe pollution than the surrounding areas, such as Xiamen, Shihezi, Changchun, Guiyang, Weinan, Yuncheng, Chengdu, and Ziyang. It also included cities with better air quality than the adjoining areas, such as Enshi Tujia and Miao autonomous prefecture, Yangjiang and Zhanjiang. Similar to the above analysis, the possible reason for this may also be due to the differences in meteorological and economic factors.

4.4 Further analysis

Many studies have similarly examined the geographical distribution characteristics of PM2.5 pollution, including by spatial autocorrelation test (Zhang et al., 2016; Ye et al., 2018; Xu, Chen, et al., 2019). However, although the research methods were similar, the subjects and perspectives were completely different. The subject of previous studies was regional PM2.5 concentrations, and thus, the spatial distribution pattern exhibited the regions with similar PM2.5 from a static perspective. In contrast, the subject of this section was the convergence clubs to which the cities converged, which depicted the geographical distribution characteristics of cities with similar PM2.5 change trends from a dynamic perspective.

Due to different research subjects, there are some new findings regarding to the geographical distribution characteristics of PM2.5 pollution. First, although cities in the same urban agglomeration are located adjacently and have close economic ties, they might still be divided into different clubs. For example, as shown in Fig. 1, not all cities in the Beijing-Tianjin-Hebei urban agglomeration converged to the same club, but were scattered over different clubs with differences in PM2.5 concentrations or growth rates. The same was also true for the Pearl River Delta and Yangtze River Delta urban agglomerations. This indicates that the PM2.5 differences among these closely connected cities were increasing, and they showed different patterns of long-term variation. Therefore, in addition to the studies on the differences among typical urban agglomerations (Lin et al., 2018; Shen et al., 2019), the exploration of differences in PM2.5 pollution within each agglomeration should not be ignored.

Second, cross-regional convergence was a common occurrence. For example, Changchun and Jinan, which are more than 1,000 km apart, both belonged to the Club H-H. This not only indicates that there are other factors affecting the convergence of PM2.5 besides geographical location, but also that there will indeed be a bias if samples are divided in advance according to geographical location when identifying convergence clubs.

5 Analysis of the influencing factors of club convergence

The spatial distribution of convergence clubs showed a significant clustering characteristic of same type, indicating the role of geographical location of a city in determining its convergence club. However, geographical location reflects the cause, but it is not the direct determining factor. Further, to reveal the reasons and propose targeted solutions for pollution, we explored its diverse determinants.

Table 6 The mean values of each club for the factors affecting the club convergence

Club	tempe	humid	industry	popula	conratio	pubtrans	regula	elec	rd	estate
Club H-H	12.84	65.07	53.21	675.13	13.28	97.23	5.46	1.20	1.34	3.18
Club H-RH	13.75	71.43	50.68	676.94	13.48	113.78	4.90	1.29	1.64	4.86
Club H-M	16.19	74.92	48.98	386.15	8.71	112.30	4.87	0.68	1.77	3.05
Club H-L	17.46	77.34	50.01	451.97	8.49	133.35	5.13	1.02	1.35	5.20
Club RH-RL	15.64	73.04	48.07	324.18	6.10	119.34	5.47	0.52	1.69	3.76
Club RH-L	12.62	69.51	46.18	276.46	5.14	104.62	5.16	0.39	1.51	2.67
Club M-RL	10.46	61.39	46.95	143.44	4.18	102.60	5.25	0.50	1.44	5.30
Club RL-M	8.46	57.26	60.86	51.67	2.23	86.71	5.46	0.81	1.75	3.52
Club RL-RL	8.55	57.79	56.05	63.04	1.71	168.10	6.33	0.50	1.17	4.48
Club L-RH	6.08	54.11	47.37	14.23	2.45	41.72	4.62	0.12	1.41	2.12
Club VL-H	4.41	52.38	33.41	18.71	9.39	414.39	5.65	/	3.18	3.13
F-stat	17.25***	28.12***	1.46	80.03***	17.23***	1.98*	1.20	11.00***	0.45	3.61***

***, **, and * represent significance at the 1%, 5%, and 10% levels, respectively

To obtain an intuitive understanding before the formal regression analysis, we first carried out a comparative analysis of the characteristic variables of 11 clubs. Table 6 shows the mean values of each club for the 10 variables listed in Table 1. Due to variance heterogeneity, the Welch test is applied for one-way Analysis of Variance and the corresponding F statistic is showed in the last row.

Table 6 shows that the F statistics of more than half variables are significant at the 1% level, such as *tempe, humid,* and *popula*, which indicates that there are indeed significant differences in these variables among 11 clubs. Thus, these may be the possible factors characterizing the observed clustering patterns. Additionally, Table 6 also reveals some trends. For example, the population density (*popula*) of clubs with high PM2.5 concentrations is significantly higher than that of clubs with low concentrations, and there is an apparent trend that the higher the PM2.5 concentration of the club, that is from Club VL-H to Club H-H, the higher the population density.

Although the comparative analysis provides the intuitions, the correlation among different variables may bias the result. In addition, the question of interest is not only the difference among the clubs in these factors, but also their specific relations with the club type, that is, how changes in these variables affect the club to which the city converges. Therefore, we further examined the driving factors that characterize the clubs based on normative regression analysis. As mentioned above, the explained variable was set as the convergence club type. Specifically, Club H-H to Club VL-H in Table 4 were successively assigned as 11 to 1, that is, the higher the PM2.5 concentration of the club, the higher the assigned value. Based on the Ordered Probit model with a robust standard error, Table 7 lists the estimated results, in which column (2) is based on 276 prefecture-level cities.[7] As a robustness test, column (3) is based on the samples remaining after removing the municipalities and provincial capitals, and columns (4)–(7) show the average marginal effect of the four high PM2.5 concentrations club type.

Column (2) of Table 7 shows that the coefficient of the relative humidity is significantly positive at the 10% level. This indicates that the higher the humidity is in a city, the more likely the city is to converge to a club with a higher PM2.5 concentration, which is consistent with the findings of Zhang et al. (2020). The response of PM2.5 to humidity is mainly due to the complex characteristics of aerosol components under different meteorological conditions. One possible reason is that higher humidity makes the soluble particles in PM2.5 absorb water and expand more easily, thus increasing the size and mass concentration of PM2.5 (Chen et al., 2017).

The coefficient of the industrial structure is significantly positive at the 5% level, indicating that cities with more developed industries are more likely to converge to a club with a higher PM2.5 concentration. If the development and application of clean technology are limited, extensive industrial development with a high energy input and high pollution emissions usually causes a conflict between the environment and development (Zhang et al., 2013).

The coefficient of population density is significantly positive at the 5% level, demonstrating that cities with a denser population also tend to converge to a club with a higher PM2.5 concentration. This is because continuous population growth greatly raises energy consumption. Without significantly improved consumption structures, fossil fuels are still the main energy source in some areas. Their burning, especially inferior coal, produces

---

[7] Due to missing values in the explanatory variable data, 66 cities out of 342 cities were excluded.

Table 7 Regression results for factors affecting the club convergence

Variables	All samples	No municipality samples	Marginal effect			
			Club H-H	Club H-RH	Club H-M	Club H-L
humid	0.0259*	0.0329**	0.0021*	0.0047*	0.0018*	− 0.0003
	(1.8151)	(2.2023)	(1.7534)	(1.7237)	(1.7165)	(− 1.1436)
industry	0.0154**	0.0150**	0.0013**	0.0028**	0.0010**	− 0.0002
	(2.4682)	(2.1310)	(2.2802)	(2.3938)	(2.2763)	(− 1.3506)
popula	0.0016**	0.0016*	0.0001**	0.0003**	0.0001**	− 0.0000
	(2.0729)	(1.8433)	(2.0550)	(2.2518)	(2.2287)	(− 1.2615)
conratio	0.0322**	0.0389**	0.0026**	0.0058**	0.0022**	− 0.0004
	(2.4307)	(2.4774)	(2.0976)	(2.3138)	(2.1431)	(− 1.2726)
pubtrans	− 0.0019*	− 0.0020*	− 0.0002*	− 0.0003*	− 0.0001*	0.0000
	(− 1.8414)	(− 1.6568)	(− 1.7379)	(− 1.7948)	(− 1.7902)	(1.1024)
regula	− 0.1492***	− 0.1437**	− 0.0122**	− 0.0269**	− 0.0101**	0.0017
	(− 2.5829)	(− 2.3814)	(− 2.4081)	(− 2.5324)	(− 2.4297)	(1.2643)
tempe	− 0.0301	− 0.0331	− 0.0025	− 0.0054	− 0.0020	0.0003
	(− 1.2948)	(− 1.3777)	(− 1.2804)	(− 1.2622)	(− 1.3183)	(0.9103)
elec	0.0290	0.1686	0.0024	0.0052	0.0020	− 0.0003
	(0.2300)	(1.4143)	(0.2280)	(0.2302)	(0.2301)	(− 0.2265)
rd	0.1027	0.0657	0.0084	0.0185	0.0069	− 0.0012
	(1.5761)	(0.8771)	(1.5744)	(1.6029)	(1.5653)	(− 1.0963)
estate	− 0.0168	− 0.0455	− 0.0014	− 0.0030	− 0.0011	0.0002
	(− 0.6701)	(− 1.5838)	(− 0.6577)	(− 0.6928)	(− 0.6811)	(0.6306)
N	276	247	276	276	276	276
Pseudo R^2	0.1086	0.1117				
Log likelihood	− 465.82229	− 417.73078				

Due to the limited layout, we only list the marginal effect of the four club with the high PM2.5 concentration

***, **, and * represent significance at the 1%, 5%, and 10% levels, respectively; the values in parentheses are robust *z*-statistics

many gaseous pollutants that easily convert to secondary aerosol species, which are main components of PM2.5 (Wang et al., 2014).

The coefficient of the proportion of construction land is significantly positive at the 5% level, indicating that the more rapidly construction expands, the more likely the city is to converge to a club with a higher PM2.5 concentration. This relation can be attributed to the impact of urban construction design and its thermodynamic properties. Studies have found that inappropriate building layouts may block airflow and result in urban heat islands, making it difficult for ground-level aerosols to diffuse outward, thus exacerbating PM2.5 pollution (Dobrovolny & Krahula, 2015; Hang et al., 2012).

However, the coefficient of the public transportation utilization rate is significantly negative at the 10% level, indicating that cities with more frequent public transportation use tend to converge to a club with a lower PM2.5 concentration. Emissions released by vehicles are an important source of PM2.5 pollution in urban areas (Yang et al., 2011); thus, cities with a high per capita car ownership usually have more severe PM2.5 pollution (Xu & Lin, 2016; Zhang

et al., 2018). In contrast, public transportation, as an alternative to private transportation, can significantly reduce PM2.5 pollution by decreasing vehicle exhaust.

The coefficient of environmental regulation is also significantly negative at the 1% level, which indicates that the greater the intensity of environmental regulation, the more likely a city is to converge to a club with a lower PM2.5 concentration. Thus, from the perspective of long-term convergence, government actions can effectively alleviate urban PM2.5 pollution, which is consistent with the findings of Zhou et al. (2019). In fact, environmental regulation not only improves regional air quality directly through environmental governance, but also affecting micro-enterprise decisions. It can influence the emission and pollution control decisions of enterprises through incentives such as taxes and subsidies or direct central supervision (Laplante & Rilstone, 1996; Liu et al., 2021; Shapiro & Walker, 2018), forcing polluting enterprises to relocate or close (Zeng & Zhao, 2009). Moreover, environmental regulation can also encourage enterprises to innovate energy-efficient and environmentally friendly technology, encourage the use of clean products, and improve environmental quality; this is known as the Porter hypothesis (Porter, 1991).

In addition, the coefficients of temperature, electricity consumption, R&D intensity, real estate investment and construction intensity are not significant, indicating that these variables have no significant impact on the club convergence.

Furthermore, the coefficients showed in column (3) are similar to those in column (2), which indicates that the above conclusions are robust and not affected by the economic and policy conditions of municipalities or provincial capitals. In addition, columns (4)–(7) shows that the average marginal effects of 10 variables on the probability of converging to the Club H-H, Club H-RH or Club H-M are consistent with the above discussion. By contrast, there is no variable that has a significant marginal effect on the probability of converging to the club H-L. This indicates that higher humidity, higher secondary industry proportions, higher population density, and more rapid construction land expansion not only increase the likelihood that a city would converge to a club with a higher PM2.5 concentration, which is the result from columns (2) and (3), but more precisely, increase the likelihood that a city would converge to the Club H-H, Club H-RH or Club H-M, the top three clubs with the highest PM2.5 concentrations.

When considering the correlation among the explanatory variables studied, the continuous advancement of urbanization has made the increasing trends in the urban population density and construction land ratio inevitable. The significant positive coefficient of the industrial structure also indicates that, for most cities in China, industrial production still extensively consumes resources. This indicates that there is a conflict between environmental protection and economic development. The key to resolving this issue and realizing coordinated improvement of the two should rely on upgrading industry through knowledgeable input (Zhao et al., 2021). In addition, reasonable urban planning, combined with effective transportation facilities, can also serve as important supplementary means to alleviate this issue. However, the key to realizing these changes is the implementation and enforcement of governmental environmental regulations.

6 Conclusions and policy implications

By performing the log t regression test, club clustering algorithm, spatial autocorrelation test, and Ordered Probit regression, this study performed an empirical convergence analysis of PM2.5 concentrations for 342 Chinese cities, which not only identified its overall

divergence and club convergence but also explored the spatial distribution pattern as well as its influencing factors.

The main conclusions are as follows: (1) There was no evidence of full panel convergence, but the results showed club convergence. Eleven convergence clubs and one divergent group were identified. (2) The geographical distribution of the convergence clubs showed significant spatial interdependencies. (3) Higher humidity, higher secondary industry proportions, higher population density, and more rapid construction land expansion increased the likelihood that a city would converge to a club with a higher PM2.5 concentration. However, a city with more frequent use of public transportation and a greater intensity of environmental regulation tended to converge to a club with a lower concentration.

Based on these conclusions, we propose following two policy implications. First, "one-size-fits-all" environmental policies should be avoided and replaced by a set of city-specific plans. This is because the PM2.5 pollution in Chinese cities converged to different clubs over time. This was true even among geographically adjacent cities in the same urban agglomeration with a close economic linkage. Second, because higher utilization of public transport increased the likelihood that a city would converge to a lower pollution club, efforts should be made to build a convenient and effective urban transit systems and encourage residents to travel by it. Conversely, because excessive urban expansion influenced cities to converge to a club with a higher pollution, it is critical to plan urban layouts reasonably. The distribution, height, and density of buildings should be designed in consideration of the local meteorological conditions, especially the wind speed and direction, to improve the city's capacity of dispersing ground-level pollutants. Active research and development of new technologies, such as the Internet of Things, may provide a way (Bai et al., 2020; Gonzalez-Zamar et al., 2020). As the conclusion shows, since government actions can effectively alleviate urban PM2.5 pollution, it is necessary to give full play to the positive role of the regulation in environmental protection. To achieve energy efficiency and emissions reduction, regulations should force industrial upgrading, change the energy consumption structure, and commit to building environmentally friendly cities.

Although this study reveals the convergence patterns of PM2.5 in Chinese cities and its main influencing factors, which provides recommendations for the haze control, there are two deficiencies. First, because this study mainly focuses on the convergence characteristics of PM2.5 and is limited by space, we do not further investigate the specific mechanism of each factor affecting the current convergence patterns. In the future research, we will conduct more detailed analyses in this regard. Second, this study mainly focuses on the convergence characteristics of PM2.5 in Chinese cities, and thus, the political and institutional influencing factors cannot be examined in depth. Our future research will extend the dataset to include other countries and cities to further explore more potential influencing factors.

Availability of data and materials The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of interest The authors declare that they have no competing interests.

Consent for publication Not applicable.

Ethics approval and consent to participate Not applicable.

References

Apergis, N., & Payne, J. E. (2017). Per capita carbon dioxide emissions across U.S. states by sector and fossil fuel source: Evidence from club convergence tests. *Energy Economics, 63*, 365–372. https://doi.org/10.1016/j.eneco.2016.11.027

Bai, C. Q., Yan, H., Yin, S. G., Feng, C., & Wei, Q. (2021). Exploring the development trend of internet finance in China: Perspective from club convergence. *North American Journal of Economics and Finance*. https://doi.org/10.1016/j.najef.2021.101505

Bai, C. Q., Zhou, L., Xia, M. L., & Feng, C. (2020). Analysis of the spatial association network structure of China's transportation carbon emissions and its driving factors. *Journal of Environmental Management*. https://doi.org/10.1016/j.jenvman.2019.109765

Becken, S., Jin, X., Zhang, C., & Gao, J. (2017). Urban air pollution in China: Destination image and risk perceptions. *Journal of Sustainable Tourism, 25*(1), 130–147. https://doi.org/10.1080/09669582.2016.1177067

Chen, Z., Chen, D., Zhao, C., Kwan, M., Cai, J., Zhuang, Y., Zhao, B., Wang, X., Chen, B., Yang, J., Li, R., He, B., Gao, B., Wang, K., & Xu, B. (2020). Influence of meteorological conditions on PM2.5 concentrations across China: A review of methodology and mechanism. *Environment International*. https://doi.org/10.1016/j.envint.2020.105558

Chen, Z. Y., Cai, J., Gao, B. B., Xu, B., Dai, S., He, B., & Xie, X. M. (2017). Detecting the causality influence of individual meteorological factors on local PM2.5 concentration in the Jing-Jin-Ji region. *Scientific Reports*. https://doi.org/10.1038/srep40735

Chen, C. H., Liu, W. L., & Chen, C. H. (2006). Development of a multiple objective planning theory and system for sustainable air quality monitoring networks. *Science of the Total Environment, 354*(1), 1–19. https://doi.org/10.1016/j.scitotenv.2005.08.018

Dobrovolny, P., & Krahula, L. (2015). The spatial variability of air temperature and nocturnal urban heat island intensity in the city of Brno, Czech Republic. *Moravian Geographical Reports, 23*(3), 8–16. https://doi.org/10.1515/mgr-2015-0013

Du, K. (2017). Econometric convergence test and club clustering using Stata. *Stata Journal, 17*(4), 882–900. https://doi.org/10.1177/1536867X1801700407

Fan, H., Zhao, C. F., & Yang, Y. K. (2020). A comprehensive analysis of the spatio-temporal variation of urban air pollution in China during 2014–2018. *Atmospheric Environment, 220*(1), 117066. https://doi.org/10.1016/j.atmosenv.2019.117066

Gonzalez-Zamar, M. D., Abad-Segura, E., Vazquez-Cano, E., & Lopez-Meneses, E. (2020). IoT technology applications-based smart cities: Research analysis. *Electroniccs*. https://doi.org/10.3390/electronics9081246

Guan, W. J., Zheng, X. Y., Chung, K. F., & Zhong, N. S. (2016). Impact of air pollution on the burden of chronic respiratory diseases in China: Time for urgent action. *Lancet, 388*(10054), 1939–1951. https://doi.org/10.1016/S0140-6736(16)31597-5

Hang, J., Li, Y. G., Sandberg, M., Buccolieri, R., & Di Sabatino, S. (2012). The influence of building height variability on pollutant dispersion and pedestrian ventilation in idealized high-rise urban areas. *Building and Environment, 56*, 346–360. https://doi.org/10.1016/j.buildenv.2012.03.023

Han, L. J., Zhou, W. Q., Li, W. F., & Qian, Y. G. (2018). Urbanization strategy and environmental changes: An insight with relationship between population change and fine particulate pollution. *Science of the Total Environment, 642*, 789–799. https://doi.org/10.1016/j.scitotenv.2018.06.094

Hao, Y., Peng, H., Temulun, T., Liu, L. Q., Mao, J., Lu, Z. N., & Chen, H. (2018). How harmful is air pollution to economic development? New evidence from PM25 concentrations of Chinese cities. *Journal of Cleaner Production, 173*(20), 743–757. https://doi.org/10.1016/j.jclepro.2017.10.195

He, J. H., Ding, S., & Liu, D. (2019). Exploring the spatiotemporal pattern of PM2.5 distribution and its determinants in Chinese cities based on a multilevel analysis approach. *Science of the Total Environment, 659*, 1513–1525. https://doi.org/10.1016/j.scitotenv.2018.12.402

Heck, T., & Hirschberg, S. (2011). China: Economic impacts of air pollution in the country. *Encyclopedia of Environmental Health.* https://doi.org/10.1016/B978-0-444-52272-6.00415-3

Huang, W., Long, E. S., Wang, J., Huang, R. Y., & Ma, L. (2015). Characterizing spatial distribution and temporal variation of PM10 and PM2.5 mass concentrations in an urban area of Southwest China. *Atmospheric Pollution Research, 6*(5), 842–848. https://doi.org/10.5094/APR.2015.093

Jiang, Z. J., Jolleys, M. D., Palmer, P., Ma, Y. P., Tian, H., Li, J., & Yang, X. (2020). Spatiotemporal and probability variations of surface PM2.5 over China between 2013 and 2019 and the associated changes in health risks: An integrative observation and model analysis. *Science of the Total Environment.* https://doi.org/10.1016/j.scitotenv.2020.137896

Liu, G. X., Dong, X. C., Kong, Z. Y., & Dong, K. Y. (2021). Does national air quality monitoring reduce local air pollution? The case of PM2.5 for China. *Journal of Environmental Management.* https://doi.org/10.1016/j.jenvman.2021.113232

Lin, C. Q., Liu, G., Lau, A. K. H., Li, Y., Li, C. C., Fung, J. C. H., & Lao, X. Q. (2018). High resolution satellite remote sensing of provincial PM2.5 trends in China from 2001 to 2015. *Atmospheric Environment, 180*, 110–116. https://doi.org/10.1016/j.atmosenv.2018.02.045

Lin, Y. F., Yuan, X. Y., Zhai, T. L., & Wang, J. (2020). Effects of land-use patterns on PM2.5 in China's developed coastal region: Exploration and solutions. *Science of the Total Environment.* https://doi.org/10.1016/j.scitotenv.2019.135602

Laplante, B., & Rilstone, P. (1996). Environmental inspections and emissions of the pulp and paper industry in Quebec. *Journal of Economics and Management, 31*(1), 19–36. https://doi.org/10.1006/jeem.1996.0029

Luo, Y., Teng, M. F., Yang, K., Zhu, Y., Zhou, X. L., Zhang, M., & Shi, Y. (2019). Research on PM2.5 estimation and prediction method and changing characteristics analysis under long temporal and large spatial scale -A case study in China typical regions. *Science of the Total Environment.* https://doi.org/10.1016/j.scitotenv.2019.133983

Ma, S., & Long, Y. (2020). Functional urban area delineations of cities on the Chinese mainland using massive Didi ride-hailing records. *Cities.* https://doi.org/10.1016/j.cities.2019.102532

Ma, Z. W., Hu, X. F., Sayer, A. M., Levy, R., Zhang, Q., Xue, Y. G., Tong, S. L., Bi, J., Huang, L., & Liu, Y. (2016). Satellite-Based spatiotemporal trends in PM2.5 concentrations: China, 2004–2013. *Environmental Health Perspectives, 124*(2), 184–192. https://doi.org/10.1289/ehp.1409481

Mi, K., Zhuang, R. L., Zhang, Z. H., Gao, J., & Pei, Q. (2019). Spatiotemporal characteristics of PM2.5 and its associated gas pollutants, a case in China. *Sustainable Cities and Society, 45*, 287–295. https://doi.org/10.1016/j.scs.2018.11.004

Ministry of Ecology and Environment of the People's Republic of China. Ecological Environment Status Bulletin in China in 2018; Ministry of Ecological Environment: Beijing, China, 2019.

Parker, S., & Liddle, B. (2017). Economy-wide and manufacturing energy productivity transition paths and club convergence for OECD and non-OECD countries. *Energy Economics, 62*, 338–346. https://doi.org/10.1016/j.eneco.2016.07.018

Panopoulou, E., & Pantelidis, T. (2009). Club convergence in carbon dioxide emissions. *Environmental and Resource Economics, 44*(1), 47–70. https://doi.org/10.1007/s10640-008-9260-6

Porter, G. (1991). *Global environment politics[M].* Westview Press.

Phillips, P. C. B., & Sul, D. (2007). Transition modeling and econometric convergence tests. *Econometrica, 75*(6), 1771–1855. https://doi.org/10.1111/j.1468-0262.2007.00811.x

Schnurbus, J., Haupt, H., & Meier, V. (2017). Economic transition and growth: A replication. *Journal of Applied Econometrics, 32*(5), 1039–1042. https://doi.org/10.1002/jae.2544

Shapiro, J. S., & Walker, R. (2018). Why is pollution from US manufacturing declining? The roles of environment regulation, productivity and trade. *American Economic Review, 108*(12), 3814–3854. https://doi.org/10.1257/aer.20151272

Shen, Y., Zhang, L. P., Fang, X., Ji, H. Y., Li, X., & Zhao, Z. W. (2019). Spatiotemporal patterns of recent PM2.5 concentrations over typical urban agglomerations in China. *Science of the Total Environment, 655*, 13–26. https://doi.org/10.1016/j.scitotenv.2018.11.105

Song, Y., Wang, X., Maher, B. A., Li, F., Xu, C., Liu, X., Sun, X., & Zhang, Z. (2017). The spatial temporal characteristics and health impacts of ambient fine particulate matter in China. *Journal of Cleaner Production, 163*, S352–S358. https://doi.org/10.1016/j.jclepro.2015.05.006

Tai, A. P. K., Mickley, L. J., & Jacob, D. J. (2010). Correlations between fine particulate matter (PM2.5) and meteorological variables in the United States: Implications for the sensitivity of PM2.5 to

climate change. *Atmospheric Environment, 44*(32), 3976–3984. https://doi.org/10.1016/j.atmosenv. 2010.06.060

Van Donkelaar, A., Martin, R. V., Brauer, M., & Boys, B. L. (2015). Use of satellite observations for long-term exposure assessment of global concentrations of fine particulate matter. *Environmental Health Perspectives, 123*(2), 135–143. https://doi.org/10.1289/ehp.1408646

World Bank, China SEPA, 2007. Cost of pollution in China: Economic estimates of physical damages. http://siteresources.worldbank.org/INTEAPREGTOPENVIRONMENT/Resources/China_ Cost_of_Pollution.pdf (Accessed 07 August 2017).

Wang, Z. B., & Fang, C. L. (2016). Spatial-temporal characteristics and determinants of PM2.5 in the Bohai Rim Urban Agglomeration. *Chemosphere, 148*, 148–162. https://doi.org/10.1016/j.chemo sphere.2015.12.118

Wang, Y. S., Yao, L., Wang, L. L., Liu, Z. R., Ji, D. S., Tang, G. Q., Zhang, J. K., Sun, Y., Hu, B., & Xin, J. Y. (2014). Mechanism for the formation of the January 2013 heavy haze pollution episode over central and eastern China. *Science China Earth Sciences, 57*(1), 14–25. https://doi.org/10. 1007/s11430-013-4773-4

Wang, S. J., Zhou, C. S., Wang, Z. B., Feng, K. S., & Hubacek, K. (2017). The characteristics and drivers of fine particulate matter (PM2.5) distribution in China. *Journal of Cleaner Production, 142*, 1800–1809. https://doi.org/10.1016/j.jclepro.2016.11.104

Xu, G. Y., Ren, X. D., Xiong, K. N., Li, L. Q., Bi, X. C., & Wu, Q. L. (2020). Analysis of the driving factors of PM2.5 concentration in the air: A case study of the Yangtze River Delta, China. *Ecological Indicators*. https://doi.org/10.1016/j.ecolind.2019.105889

Xu, S. C., Miao, Y. M., Gao, C., Long, R. Y., Chen, H., Zhao, B., & Wang, S. X. (2019). Regional differences in impacts of economic growth and urbanization on air pollutants in China based on provincial panel estimation. *Journal of Cleaner Production, 208*, 340–352. https://doi.org/10.1016/j.jclep ro.2018.10.114

Xu, L. Z., Chen, F., Zhong, X. F., Zhang, L. E., Ye, R. P., Cai, W. Y., Rao, Q. H., & Li, J. B. (2019). Spatial disequilibrium of fine particulate matter and corresponding health burden in China. *Journal of Cleaner Production, 238*, 117840. https://doi.org/10.1016/j.jclepro.2019.117840

Xu, B., & Lin, B. Q. (2016). Regional differences of pollution emissions in China: Contributing factors and mitigation strategies. *Journal of Cleaner Production, 112*, 1454–1463. https://doi.org/10. 1016/j.jclepro.2015.03.067

Yang, Y., Christakos, G., Yang, X., & He, J. Y. (2018). Spatiotemporal characterization and mapping of PM 2.5, concentrations in southern Jiangsu Province, China. *Environmental Pollution, 234*, 794–803. https://doi.org/10.1016/j.envpol.2017.11.077

Yang, S. Y., Fang, D. L., & Chen, B. (2019). Human health impact and economic effect for PM2.5 exposure in typical cities. *Applied Energy, 249*, 316–325. https://doi.org/10.1016/j.apenergy.2019.04. 173

Yang, F., Tan, J., Zhao, Q., Du, Z., He, K., Ma, Y., Duan, F., Chen, G., & Zhao, Q. (2011). Characteristics of PM2.5 speciation in representative megacities and across China. *Atmospheric Chemistry and Physics, 11*(11), 5207–5219. https://doi.org/10.5194/acp-11-5207-2011

Yan, D., Lei, Y. L., Shi, Y. K., Zhu, Q., Li, L., & Zhang, Z. (2018). Evolution of the spatiotemporal pattern of PM2.5 concentrations in China-A case study from the Beijing-Tianjin-Hebei region. *Atmospheric Environment, 183*, 225–233. https://doi.org/10.1016/j.atmosenv.2018.03.041

Ye, W. F., Ma, Z. Y., & Ha, X. Z. (2018). Spatial-temporal patterns of PM2.5 concentrations for 338 Chinese cities. *Science of the Total Environment, 631–632*, 524–533. https://doi.org/10.1016/j.scito tenv.2018.03.057

Yun, G. L., He, Y. R., Jiang, Y. T., Dou, P. F., & Dai, S. Q. (2019). PM2.5 Spatiotemporal evolution and drivers in the Yangtze River Delta between 2005 and 2015. *Atmosphere, 10*, 2. https://doi.org/10. 3390/atmos10020055

Zhao, J., Dong, X. C., & Dong, K. Y. (2021). How does producer services' agglomeration promote carbon reduction? The case of China. *Economic Modelling*. https://doi.org/10.1016/j.econmod.2021.105624

Zhang, L. C., An, J., Liu, M. Y., Li, Z. W., Liu, Y., Tao, L. X., Liu, X. T., Zhang, F., Zheng, D. Q., Gao, Q., Guo, X. H., & Luo, Y. X. (2020). Spatiotemporal variations and influencing factors of PM2.5 concentrations in Beijing, China. *Environmental Pollution, 262*, 2. https://doi.org/10.1016/j.envpol. 2020.114276

Zhang, N., Huang, H., Duan, X. L., Zhao, J. L., & Su, B. N. (2018). Quantitative association analysis between PM2.5 concentration and factors on industry, energy, agriculture, and transportation. *Scientific Reports*. https://doi.org/10.1038/s41598-018-27771-w

Zhang, H. F., Wang, Z. H., & Zhang, W. Z. (2016). Exploring spatiotemporal patterns of PM2.5 in China based on ground-level observations for 190 cities. *Environmental Pollution, 216*, 559–567. https://doi.org/10.1016/j.envpol.2016.06.009

Zhang, R., Jing, J., Tao, J., Hsu, S. C., Wang, G., Cao, J., Lee, C. S. L., Zhu, L., Chen, Z., Zhao, Y., & Shen, Z. (2013). Chemical characterization and source apportionment of PM2.5 in Beijing: seasonal perspective. *Atmospheric Chemistry and Physics, 13*(14), 7053–7074. https://doi.org/10.5194/acp-13-7053-2013

Zhou, Q., Zhang, X. L., Shao, Q. L., & Wang, X. L. (2019). The non-linear effect of environmental regulation on haze pollution: Empirical evidence for 277 Chinese cities during 2002–2010. *Journal of Environmental Management*. https://doi.org/10.1016/j.jenvman.2019.109274

Zeng, D. Z., & Zhao, L. X. (2009). Pollution havens and industrial agglomeration. *Journal of Environmental Economics Management, 58*(2), 141–153. https://doi.org/10.1016/j.jeem.2008.09.003

Zou, B., Pu, Q., Bilal, M., Weng, Q. H., Zhai, L., & Nichol, J. E. (2016). High-Resolution satellite mapping of fine particulates based on geographically weighted regression. *IEEE Geoscience and Remote Sensing Letters, 13*(4), 495–499. https://doi.org/10.1109/LGRS.2016.2520480