

What are the real populations of Chinese cities? On mistakes in city border definition and data interpretation in popular internet sources

Dariusz Sokołowski

Nicolaus Copernicus University in Toruń, Department of Spatial Management and Tourism, Torun, Poland, e-mail: sokol@umk.pl, <https://orcid.org/0000-0003-0361-7017>

How to cite:

Sokołowski, D. (2023). What are the real populations of Chinese cities? On mistakes in city border definition and data interpretation in popular internet sources. *Bulletin of Geography. Socio-economic Series*, 62(62): 7-26. DOI: <http://doi.org/10.12775/bgss-2023-0031>

Abstract. Economic growth and intensive metropolisation processes in the PRC have increased interest in information on the size of the country's cities. Chinese institutions apply population data to administrative units which is considerably larger than its urban area. The publicly available data are often divergent; according to Gibson & Li (2017) hundreds of studies in economics misinterpret China's subnational population and over 80% of articles use these data erroneously. Few specialists are able to use data directly from original Chinese sources (see Chan, 2007; Chan & Wan, 2017), most of them use publicly available sources. Scientists and other users often have at their disposal estimates published by international institutions. A comparative analysis of those data exposes marked dissimilarities. The article is to identify differences between particular sources and to establish their causes. I assume that the reason for the discrepancy in the assessment of the city's population is not a difference in method of defining the city, but rather the fact that the authors of the studies do not apply established criteria and do not provide accurate calculations. Finally, guidelines are created that, if followed, should result in smaller discrepancies between data published by various sources.

Article details:

Received: 16 October 2023
Revised: 09 November 2023
Accepted: 28 December 2023

Key words:

data analysis,
Chinese population,
population data comparison,
China cities

Contents:

1. Introduction	8
2. Cities in the territorial structure of the PRC	9
3. Definition of the city	11
4. Research methods and data sources	13
5. Brief description of sources	14
6. Causes of variations in population of cities, sources of mistakes and differing interpretations ...	18
7. Discussion	23
8. Concluding note	24
Notes	24
References	25

1. Introduction

The economic success of the People's Republic of China, which dates back to the 1980s, led to a dynamic rise in the population of cities and intensification of metropolisation processes. Quantitative data on urban population in the PRC are mentioned in numerous academic papers, popular science literature, expert opinions, reports, media releases, etc. The accuracy of measurement of population in Chinese cities is important for demographic, economic and social reasons: the PRC has the largest urban population of any state in the world; for several decades the country has been witnessing a pace of growth in urban population that is spectacular on a global scale; the importance of the Chinese economy is systematically rising; and the contact between Chinese citizens and the outside world and vice versa (particularly in terms of business and tourism) is reaching an unprecedented level. The above-mentioned factors sustain the constant international interest in the PRC's cities.

The primary sources of population data are Chinese institutions – first and foremost the National Bureau of Statistics of China, which treats cities in terms of areas of municipal jurisdiction, which does not match with definitions accepted by most countries. Need to point out the Chinese “cities” are first and foremost administrative units in the top-down power hierarchy manifested as “cities” controlled by the party-state; hence, all the “abnormalities”, which causes a lot of confusion to unsuspecting observers (Cartier, 2015).

International institutions (e.g. the UN) have been attempting to estimate the size of Chinese cities as urban areas or urban agglomerations, in order to express the numbers in more real terms.

Nevertheless, specific problems are encountered when efforts are made to determine the number of people in Chinese cities. Data published by different institutions, research units and Internet portals often show serious discrepancies: in extreme cases differences are of as much as several hundred percent. What is the reason behind such substantial inaccuracies? What are the methods for calculating the populations of Chinese cities? Are some data sources more accurate than others?

Earlier research (e.g. Chan, 2007; Chang & Wan, 2017) has shown that the overstatement and understatement of Chinese city sizes actually coexist. The problems at the aggregate level (national or provincial) have become a topic discussed by many scientists (e.g. Chan & Hu, 2003; Kirkby 1985; Pannell, 2003; Shen, 2006; Zhou & Ma, 2003, 2005), while the study at the individual-city level remains relatively scant (Chan, 2007). Chan presented an analysis of the problems related to determining the

actual size of Chinese cities, focusing mainly on the largest (provincial- and prefecture-level) cities. Chan (2007) clarifies factors leading to misunderstanding regarding the number of inhabitants of China's major cities, indicating, principally, the multi-layered meanings of the term “city”, the effects of the Chinese *hukou* (household registration) system and the rapid rate of urban growth. This article is to show not only a different approach resulting from the above-mentioned premises, but also the differences between the sources, as well as numerous inconsistencies within individual sources.

There are two main reasons for the provision of incongruous data on urban populations: one is the city delimitation resulting from the method for defining a city; the other is the accuracy of estimates of the population within delineated boundaries. Based on the review of sources, the following hypothesis may be formed: in the case of Chinese cities, there are many more reasons for these discrepancies. They result not only from the problems of the multi-layered meanings of the term “city” and the effects of the *hukou* system discussed by other researchers; in publicly available sources there are also errors and inconsistencies resulting from deviations from the accepted definitions (inattention?), which result in significant (unjustified) annual changes in the data.

Chan (2007), Chan & Wan (2017), Gibson & Li (2017) mainly point to errors due to discrepancies between *hukou* registrations and the actual place of residence (and work), and errors due to a misunderstanding of the complex structures of territorial administration in the PRC (see Fig. 1). Gibson and Li provide numerous examples of data interpretation errors in scientific articles; this study goes in a slightly different direction – it points to errors generated upstream, i.e., in the sources that form the basis for these publications. This is important because most non-Chinese scholars use mainstream, usually English-language sources. The aim of this paper is to make a critical review of online sources publishing data on the population of Chinese cities and to attempt to describe and identify the causes of these errors and differences; only some of the sources of errors are indicated in the above-mentioned works. The study should also make it possible to indicate the most accurate source, i.e. the one publishing data that are closest to reality in the light of established definitions.

Due to limits on article length, the text will refer only to data for recent years (2010–2018), and on account of uniformity of sources the spatial coverage will be restricted to Mainland China, i.e. the area of the PRC excluding the special-status territories (Hong Kong, Macau).

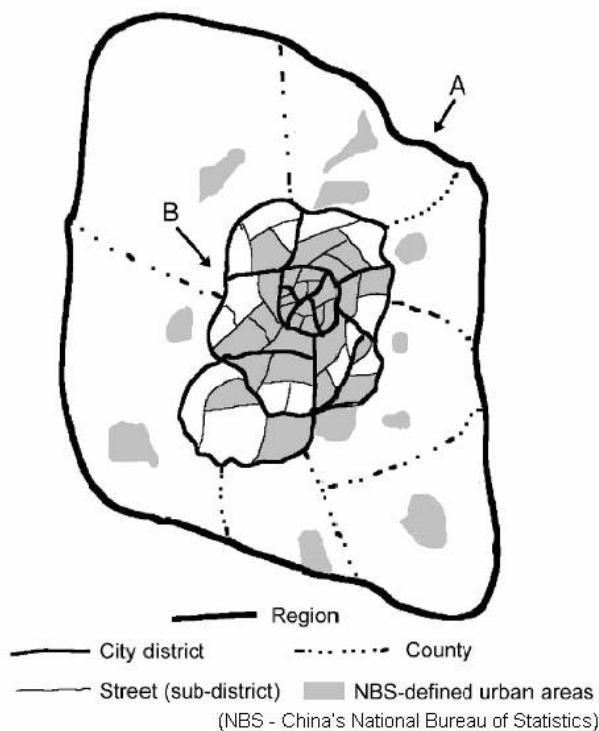


Fig. 1. Conceptual diagram of the spatial/administrative structure of a typical large city in China
Source: Chan, 2007

2. Cities in the territorial structure of the PRC

Problems related to the delimitation of cities are partly consequent upon the political and administrative structure established in the PRC, where cities are not distinguished by borders that would approximately coincide with those of urban areas. The areas of municipal jurisdiction are often – incorrectly – identified with cities (province-level cities, prefecture-level cities, county-level cities), but they cover an area much larger than the urban area. The area typically includes an urbanised core surrounded by extensive rural areas (see, for example: Shen, 2005; Chan, 2007). Quite the contrary situation exists in, for example, the USA, where cities generally have tight boundaries, and the delimitation of urban areas includes adjacent areas belonging to neighbouring administrative units (counties).

The territorial structure of the PRC should be discussed before the description of sources, because it will be referred to through the article.

At the first level of division, alongside provinces and autonomous entities, there are four province-level cities (Fig. 2, Table 1). Chongqing is a good

case in point; its administrative area is larger than Belgium, the Netherlands and Luxembourg combined. Considering this province-level city, which is mostly covered by agricultural lands, as an urban area is inconsistent with the actual state. Although the other three province-level cities are smaller in area, all of them have very extensive borders, and therefore cover substantial non-urbanised areas and their rural inhabitants.

The second level of the division comprises mainly prefecture-level cities, the number of which grew rapidly in the past: 102 in 1982; 185 in 1990; 259 in 2000; 283 in 2007 (Chien, 2010), though nowadays their number is stable and ranges around 294 (Note 1). Fifteen of the largest cities of that level are not administratively controlled by provincial governments and they have the status of deputy-provincial level cities.

They are: Changchun, Chengdu, Dalian, Guangzhou, Hangzhou, Harbin, Jinan, Nanjing, Ningbo, Qingdao, Shenyang, Shenzhen, Wuhan, Xiamen and Xi'an. The use of the term "cities" for units at this level is misleading for the same reasons mentioned in the case of province-level cities (their area usually amounts to several or over a dozen thousand km²) (Note 2).

Third-level units include: districts (*Qu*), county-level cities (*Shi*) and counties (*Xian*). Districts are usually small units whose area measures in the tens or hundreds of km², and which cover – entirely or in large part – urbanised areas connected with a core city (prefecture-level city). The populations of county-level cities sometimes exceed 1,000,000; however, on account of the large area of those units (often ~1,000 km²), these cities also cover non-urbanised areas. Counties are defined as rural areas (Note 3), but their administrative centres are units with an urban character. Third-level units (*Qu*, *Shi* and *Xian*) may be divided into: subdistricts (urban subdistrict – *Jiedao*); towns (town – *Zhen*); townships (rural township – *Xiang*); residential communities; administrative villages. Not all types of low-level units must be represented in a particular unit.

It should be emphasised that none of the types of first-, second-, third- and fourth-level units is by definition exclusively urban: they comprise both urban and rural areas.

By way of example, the actual urbanisation of units that make up the political and administrative structure is demonstrated by the data for the city of Tianjin (Note 4) (Fig. 3, Table 2), where only six central districts are 100% urbanised. The other districts include urban areas, suburbs of diverse levels of urbanisation, as well as totally rural areas. In the case of districts located peripherally,

urbanised areas are not connected spatially with the core city. The administrative centre of the district of Jizhou is 120 km from the centre of Tianjin (a similar distance separates Tianjin and Beijing) and almost 100 km from its suburbs.

The areas subject to administratively designated as a city or urban region (Table 1, Fig. 3) in Chinese, American and European cities is in many cases similar. The Chinese specificity is in the fact that both city centers (core) and peripheral areas are very densely populated (see Fig. 3). For this reason, the delimitation of urban areas in the PRC should be taken higher population density threshold indicator – if it is the main or one of the criteria. It is also advisable to take into account other delimitation criteria, such as the character of buildings, functional connections or sources of livelihood (in PRC some studies use data on non-agricultural populations).

The majority of sources sum up the urban population of all districts subordinate to their core

city administration (in this case – Tianjin), which results in the overestimation of its population. Also, it is worth noting that even if the same basis for calculation is used (2010 Census), individual sources are characterised by remarkable discrepancies: 9.583 million (Table 1) and 10.278 million (Table 2). Based on lower-level cities (prefecture-level cities), it can also be demonstrated that not only the population of the whole administrative unit but also the population of districts (*Qu*) is far larger than that of the core city's urban area (Table 3).

One example of the administrative structure typical of eastern provinces of the PRC is Suzhou (prefecture-level city in Jiangsu), whose land area (Note 5) is 6,094 km² (Table 3). It includes nine third-level units: five districts (*Qu*) and four county-level cities (*Shi*). The core city comprises (exclusively or in its largest part) four districts and part of a fifth district (*Wujiang*). The other units include urban and



Fig. 2. Provincial-level administrative divisions in PRC (Mainland)

Source: adaptation of <http://www.ibiblio.org/chinesehistory/imagemap.html>

Table 1. Provincial-level cities of the People's Republic of China (compared to other selected metropolises)

Name	Year	Population [thou.] [4]	Land area [km ²] [4]	Density [/km ²]	Urban area					
					Population [4]		Population [3]		Land area [3]	Density [/km ²]
					[thou.]	%	[thou.]	%		
Beijing	2020	21,893	16,411	1,334	18,961	86.6	19,433	88.8	4,172	4,658
Chongqing	2020	32,054	82,403	389	9,581	29.9	7,739	24.1	1,537	5,035
Shanghai	2020	24,871	6,341	3,923	21,910	88.1	22,120	88.9	4,068	5,438
Tianjin	2020	13,866	11,610	1,194	10,900	78.6	10,800	77.9	2,813	3,839
Chicago	2020	9,619 ¹	18,634 ¹	516 ¹	8,608 ²	91.0 ²	9,014	97.4	7,006	1,287
Paris	2020	12,263 ³	12,012 ³	1,032 ³	10,859 ⁴	88.6 ⁴	11,020	89.9	2,509	4,392

¹Chicago–Naperville–Elgin Metropolitan Statistical Area, ²Chicago Urban Area [4] in 2010, 6,327 km², MSA population: 9,461 thou.; ³Île-de-France Region in 2019; ⁴Paris Urban Agglomeration [4] in 2019, 2,853 km².

Source: [4], [3], U.S. Census Bureau (web), Institut National de la Statistique et des Études Économiques, France (web) and own calculations

rural areas and in their core there are large cities, such as Kunshan and Changshu.

The above examples show that administrative units do not match urbanised areas; therefore, they cannot be identified as cities, even though they are referred to as a “city”. Neither are the smallest administrative units (townships and villages) appropriate for the delimitation of urban areas in the PRC.

The complexity of the administrative structure and its inadequacy for the delimitation of urban areas is the main cause of ambiguity in the estimates of population of the PRC's cities.

3. Definition of the city

To determine city size it is necessary to consider how a city is defined. The city can be considered *de jure* or *de facto*. The concept of the city *de jure* is connected with the above-discussed political and administrative approach. Therefore, data should be refer here to the concept of the city *de facto*, which – in practice – entails approaching the city irrespective of its administrative borders.

For centuries people have been trying to define “the city”. Classical definitions postulated by geographers, urban planners or sociologists refer to different characteristic features of the city and vary according to time and place. However, it is possible to distinguish several characteristics of the city that define it in a universal manner (Sokołowski, 1998, cf Maik, 1997; Szymańska, 2007): size of unit; population, urban and infrastructure density, and other density indexes; type of infrastructure; heterogeneity (social, functional, urban, etc.); non-

agricultural character; landscape highly transformed by human activity; well-developed infrastructure; etc. The figures that make those criteria more precise vary depending on time and space (for example, in some Asian countries, including around large PRC cities, rural population densities are much higher than in the suburbs of U.S. and European cities), but their significance remains unchanged. Assuming that those features are the ones that define the city, this term can be apply not only to the core area of the city, but to adjacent urban areas as well. Numerous studies ascribe some meaning to functional and economic connections, too; however, it should be remembered that connections often refer to areas of greater vastness and, in principle, they are associated with the concept of metropolitan areas. It is not so often that functional/economic connections occur within an area smaller than an urban area: however, they can occur in the case of two or more cities that lie close to one another and are joined to one another by continuous urban development, in which the strength of connections will determine to which city a given urban area belongs.

In international sources, terms related to the concept of ‘city’ are understood differently from country to country. For example, U.S. Metropolitan Areas are not defined in morphological, but primarily in functional terms. MAs in the USA include (except for New England) entire counties demonstrating adequate strength of connections to the central city/cities. Their areas can be huge (even tens of thousands of square kilometres, such as in the case of Los Angeles), their population densities low, and most of the county land included in an MA may be

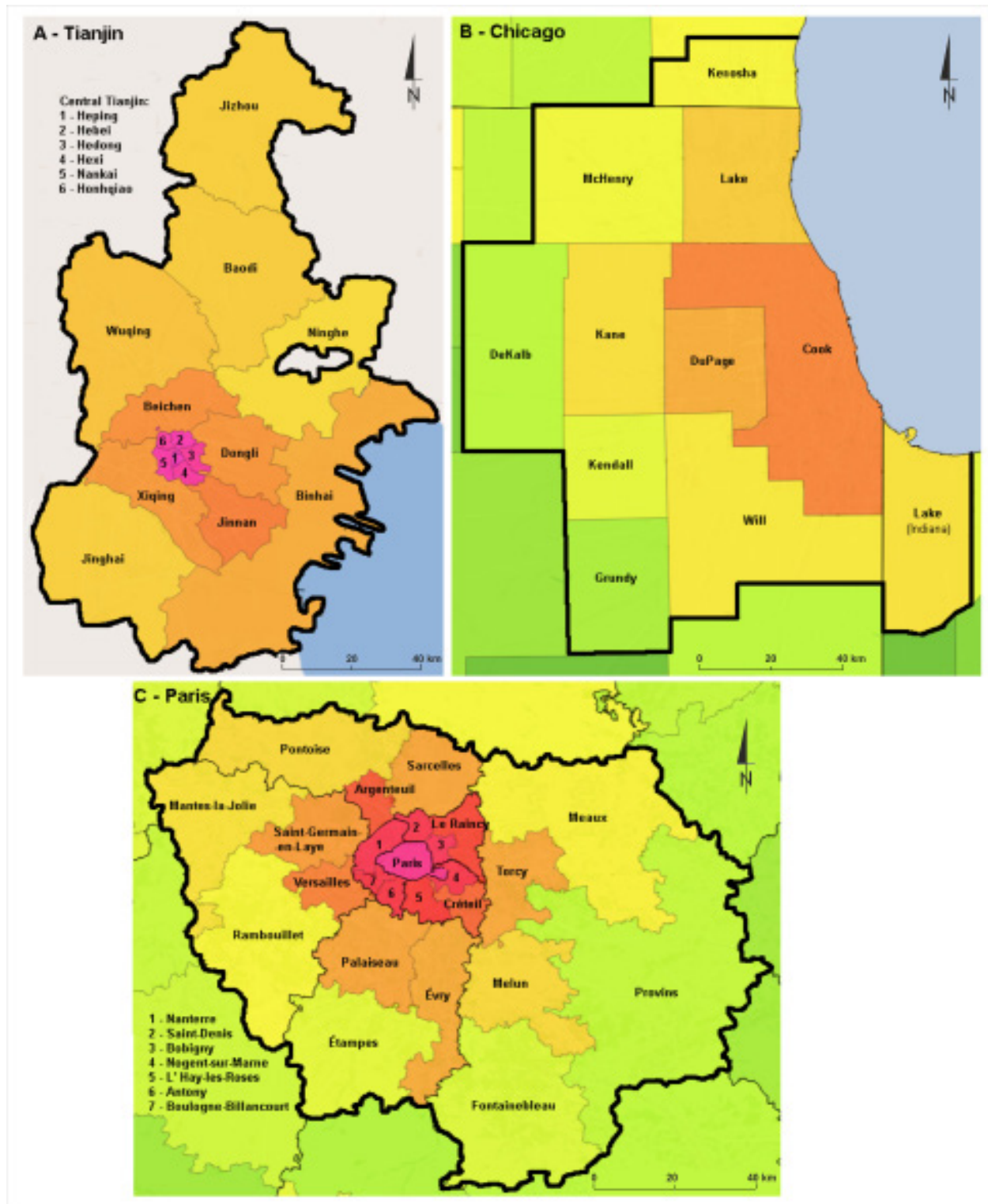


Fig. 3. Comparison of the area and structure of administrative units in selected metropolises: A – Tianjin (districts), B – Chicago (counties), C – Paris (arrondissements). Borders are distinguished: A – Municipal province (Province-level city), B – Chicago–Naperville–Elgin Metropolitan Statistical Area, C – Île-de-France Region. Urbanized areas cover only a part of the distinguished units. The colors indicate the population density (see Tables: 1, 2); extreme sizes: Tianjin (2020) – Ninghe (343/km²), Heping (35,500/km²), Chicago (2020) – Grundy (49/km²), Cook (2,156/km²), Paris (2019) – Provins (79/km²), Paris (20,545/km²).

Source: [4], U.S. Census Bureau (web), Institut National de la Statistique et des Études Économiques, France (web) and own calculations

Table 2 Population in districts of Tianjin

Division	Population 2010			Population 2020
	Total [thou.][a]	Urban area thou.[b]	%	Total [thou.][c]
Tianjin	12,939	10,278	79.4	13,866
Heping	273	273	100.0	355
Hedong	861	861	100.0	859
Hexi	871	871	100.0	822
Nankai	1,018	1,018	100.0	917
Hebei	788	788	100.0	648
Hongqiao	532	532	100.0	483
Dongli	599	591	98.7	936
Xiqing	713	525	73.6	1,181
Jinnan	593	590	99.5	928
Beichen	669	575	85.9	910
Wuqing	951	353	37.1	1,153
Baodi	799	272	34.0	722
Binhai	2,423	2,313	95.5	1,974
Ninghe	416	152	36.6	395
Jinghai	647	293	45.3	787
Jizhou	785	270	34.4	796

Compiled by [a] the Census Office of the State Council, [b] the Department of Population and Social Science and Technology Statistics of the National Bureau of Statistics (2012) and [c] – source [4].

Data for the 2010 and 2020 – China Census by County

unurbanized. The term ‘urban agglomeration’ is not precisely defined and its understanding varies not only from country to country, but even from study to study. International sources publishing lists of cities for all countries of the world most often use the concepts of ‘urban area’ and ‘urban agglomeration’ in the morphological sense (continuous, compact buildings, relatively high population density for a given country, the absence or a small share of agricultural land and employment in agriculture, etc.) and they can be identified with the concept of ‘city’ understood in the non-administrative sense. A similar understanding of the terms in question is evidenced by the similar population sizes of some cities reported in various sources, despite the use of different names for urban units (Note 6).

Individual sources use complex criteria and point to the political and administrative status, population density and spatial continuity (Note 7). Some sources refer to the concepts of urban area or urban agglomeration without any precise definition; in those cases it is reasonable to presume that the assumptions on which their delimitations are

made are (theoretically) similar. One of the most important causes of variations in city sizes (Note 8) based on different estimates, is a dissimilar approach to the criterion of spatial continuity.

4. Research methods and data sources

The hypothesis regarding the inconsistency of sources and inconsistency in the interpretation of data cannot be verified on the basis of statistical inference, only the comparative method, whereby the analysis involves data on population and area of units acquired from different sources. The spatial range of units was mostly based on Google Maps – a tool that is commonly available and provides enough details to suit the research needs. The estimation of the population of urban areas based on ranges of urban development and population within the administrative borders of units at various levels enabled the identification of fundamental mistakes in sources.

For such comparative studies, several public online sources were used. The numbers in square brackets should be treated as reference numbers for those used later in the text:

- [1] United Nations; data from different years, contains data (archival, current and forecasts for the period 1950-2035 in one table) in 5-year cuts. Last Revision: 2018;
- [2] World Population Review; all of the data used are for 2020, annually changes the date of 'Population of All Cities in China'; but the data does not change from 2018 to 2023;
- [3] Wendell Cox, Demographia World Urban Areas (releases 2014-2022), updates data annually;
- [4] Thomas Brinkhoff: China; this source uses information from 2000, 2010 and 2020 Censuses; this source also provides data on the unit area, provides city population data in PRC based on National Censuses (every 10 years); for some cities, estimates for other selected years are published. Last Revision: 2022;
- [5] PopulationData.net, does not provide the date of the data; data changes irregularly, some remaining identical since at least 2018;
- [6] PopulationStat. World statistical data (data for 2020-2022), the data are updated data irregularly, much of it has not been updated for several years, like the source [2]; archive data were sourced from:
- [7] World Gazetteer. The data is not updated, only archival source.

Table 3. The share of urban areas in the population of selected prefecture-level cities in the PRC (2010 and 2020 Census)

Name	Unit	Area in km ² (2020)	Population in thou.		Population density /km ²		Population of main city			
			2010	2020	2010	2020	thou.		% of unit	
							2010	2020	2010	2020
Suzhou (Jiangsu)	A	6,094 ¹	10,460	12,748	1,716	2,092				
	B	2,945 ¹	5,346	6,716	1,815	2,280	4,084	5,893	39.0	46.2
Hefei (Anhui)	A	11,473	7,457	9,370	650	817				
	B	1,348	3,466	5,118	2,571	3,797	3,099	5,056	76.4	87.7
Zhengzhou (Henan)	A	7,565	8,627	12,601	1,140	1,666				
	B	1,033	4,122	6,453	3,990	6,427	3,560	6,461	41.6	54.0
Changsha (Hunan)	A	11,819	7,041	10,048	596	850				
	B	2,154	3,744	5,981	1,738	2,777	3,193	5,630	89.4	98.8
Nanchang (Jiangxi)	A	7,194	5,043	6,255	701	869				
	B	2,778	3,153	3,930	1,135	1,415	2,614	3,519	41.3	51.3
Nanning (Guangxi)	A	22,108	6,659	8,742	301	395				
	B	9,836	3,979	5,977	405	608	2,661	4,583	86.4	100.1 ²
Guiyang (Guizhou)	A	8,051	4,323	5,987	537	744				
	B	2,519	3,051	4,506	1,211	1,789	2,520	2,661	45.3	45.3
Taiyuan (Shanxi)	A	6,905	4,202	5,304	609	768				
	B	1,416	3,427	4,529	2,420	3,198	3,154	4,304	85.3	88.3

A – all sub-provincial city, B – sub-provincial city without counties and county-level cities (only districts); 1 land area only, the whole area is 9,102 km² and 5,091 km² respectively (with Lake Taihu); 2 the population of the main city is not the sum of the population of individual districts, so values exceeding 100% are possible

Source: [4] and own calculations

5. Brief description of sources

Differences between individual compilations of city data mostly lie in:

1. different numbers of cities being taken into account,
2. different populations being provided for the same city,
3. different populations being provided for the same city in the same source for different years, which is not reflected in the actual dynamics of growth; as well as differences in data for the same year in various editions of the compilation.

The study by the United Nations [1] is very extensive; it includes tabular data on population

in cities for 1950–2035, based on population censuses, estimates and forecasts. Data in this compilation are usually not remarkably different from other compilations, in particular [3], [4], [6], but there are also some exceptions, e.g. Chongqing, which according to [2], [3], [4] has 7–8 million people, while [1] informs about 13.372 million in 2015 and (forecast) 15.872 million in 2020 (cf Table 4). The analysis of data by districts proved that data from [1] are here significantly inflated; the whole population (urban and non-urban) of 11 core and suburban districts (9,824 km²) equals 10.512 million (2017 [4]). The number of almost 15 million applies to the urban population of the entire district administered by Chongqing (82,403 km²) and includes such remote cities as Wanzhou and Kaizhou, which lie 270 km and 340 km, respectively, from the core city. The significant data

revision (Chongqing and other cities) during the 2020 census is noteworthy.

Another example of artificially high values is the case of Handan (Hebei province), with a population of 2.248 million (2015) and 2.727 million (2020); these figures are about twice as high as those provided by other sources. No case was found of any drastic underestimation of population in [1]. The source provides data on 424 cities >300,000 in the PRC. The number of cities included is lower than the actual number, as county administration centres (*Xian*) are generally ignored. An example: the urban network in the north-west province of Jiangsu is relatively regular; larger cities are reasonably regularly distanced from each other (40–50 km). In this area, the cities of Xuzhou, Suqian and Huai'an have the status of prefecture-level city and in all comparable sources they are considered as cities. Other cities administering lower-level areas: Pizhou and Xinyi belong to county-level cities (*Shi*), while Suining, Shuyang, Siyang and Sihong are categorised under counties (*Xian*) (Note 9). According to [1], the last four are not in the category of cities, even though their size approximates that of Pizhou and Xinyi.

As suggested by [1], it is the administrative status that decides whether a unit is categorised as a city. Those cities that are administrators for counties are usually overlooked. For the sake of comparison, [4] lists all of the above-mentioned cities.

The fact that [1] incorrectly identifies administrative units is confirmed by data for other countries: for instance, data on cities within their administrative borders are taken as the size of urban agglomeration in Germany, Indonesia, Poland, Republic of Korea and Russian Federation.

World Population Review [2] compiles a list of 400 Chinese cities of over 100,000 inhabitants (Note 10). It includes a lot of inexplicable mistakes of either overestimation or underestimation of population in cities. For some cities the entire population of the prefecture-level city is given, including the population of remote cities and villages, which results in unjustified inflation of data by as much as several hundred percent. For example: Tai'an (Shandong) 5.499 million (7,762 km²), Tianshui (Gansu) 3.500 million (14,359 km²), Shiyan (Hubei) 3.460 million (23,680 km²), Yunfu (Guangdong) 2.613 million (7,779 km²), Ordos (Inner Mongolia) 1.941 million (86,882 km²). It was calculated that the city of Nanchong (Sichuan) has 7.150 million inhabitants, which is more than the whole of Nanchong Shi (6.418 million in 2017 [4], 12,480 km²). One interesting

example is Dadonghai in Hainan Province (2.000 million). No city of such a name exists; Dadonghai is part of the city of Sanya. The population of all Sanya prefecture-level city (1,905 km²) in 2020 equalled 1,031 thousand [4].

Data on some cities are significantly underestimated in that source. It suffices to mention Beijing 11.717 million, the city which in other sources is said to have circa 20 million. Another preposterously low data point on population of cities provided by [2] is, for example (for the sake of comparison the figures in the brackets show the population of cities according to [1] and the estimate for 2020; all data are in millions): Foshan (Guangdong) – 3.600 (7.327); Fuzhou (Fujian) – 1.180 (3.686); Changzhou (Jiangsu) – 0.949 (3.625); Wenzhou (Zhejiang) – 0.866 (3.624); Nanning (Guangxi) 0.804 (3.860); Yantai (Shandong) – 0.719 (2.527); Huai'an (Jiangsu) – 0.555 (2.655); Zhuhai (Guangdong) – 0.501 (1.759); Putian (Fujian) – 0.377 (1.907) (Note 11); Linyi (Shandong) – 0.271 (1.937); Weihai (Shandong) – 0.153 (1.304); Puning (Guangdong) – 0.118 (1.160). Xuzhou (Jiangsu) was not even mentioned among cities of >100,000.

Source [2] does not give exact information on the method for identification of cities. As in [1], most cities that are administration centres for counties are ignored.

Demographia World Urban Areas [3] annually presents data on population of cities in all countries of the world. The PRC's cities are defined as urban areas, which is the same as in the method for defining the city *de facto* – the method favoured by the author of this paper. The 2019 publication includes data on 317 cities in the PRC, revision 2020 shows 248 cities, and rev. 2021 – 214 cities.

The inaccuracy of that source has at least two causes. Firstly, it does not include the population of most cities that administer counties (*Xian*), as is the case in [1] and [2]. Secondly, the inaccuracies in the data in [3] originate from frequent changes in the spatial range of cities (delimited for the purposes of the study) and inaccurate calculations. As a result, in subsequent publications cities not only differ in terms of the total population growth, which is consequent upon the birth rate and net migration rate, but they also sometimes present significant changes due to the delineation of new borders of urban areas. Erroneous completion of these procedures is observed in numerous situations when together with a considerable increase in area (or if the area is constant), population decreases incrementally. The above problem is illustrated by a few examples in Table 5 and Figure 4.

Table 4. Population of largest cities/urban areas [in thousands] in the period 2010–2020(21) based on different sources; order as in 2020 [4] (bold)

Location	2010 [1]	2015 [1]	2020 ¹ [1]	~2020 [2]	2010 [3]	2015 [3]	2020 [3]	2010 [4]	2020 [4]	2019 [5]	2021 [5]	2020 [6]
Guangzhou-Foshan	16,931	18,722	20,629	14,671	16,400	20,597	20,902	17,413	25,139	23,426	23,426	18,372
Shanghai	20,314	23,482	27,058	22,315	21,550 ²	23,416	22,120	20,217	21,910	26,317	24,871	26,778
Beijing	16,441	18,421	20,463	11,717	19,000 ²	21,009	19,443	16,704	18,961	21,542	21,893	20,301
Shenzhen	10,223	11,275	12,357	10,358	11,500	12,084	15,929	10,358	17,445	12,905	12,357	12,271
Chengdu	7,573	8,360	9,136	7,416	8,750	10,376	11,309	7,824	13,568	14,677	14,677	9,074
Tianjin	10,150	12,516	13,589	11,090	8,775	10,920	10,800	9,736	11,052	15,621	13,866	13,516
Wuhan	7,515	7,921	8,365	9,785	6,900	7,509	8,962	7,542	10,495	10,893	10,893	8,328
Dongguan	7,118	7,325	7,408	8,000	8,100	8,442	7,981	7,271	9,645	8,254	8,397	7,397
Chongqing	11,244	13,372	15,872	7,458	6,200	7,217	7,739	6,264	9,581	15,384	22,252	15,676
Xi'an	5,526	6,657	8,001	6,501	5,175 ³	5,977 ³	6,680 ³	5,596	9,393	12,906	12,906	7,895
Hangzhou	5,758	6,658	7,642	6,242	5,950	7,275	6,446	5,850	9,236	10,018	10,018	7,565
Nanjing	6,162	7,392	8,847	7,165	5,525	6,155	7,496	5,828	7,520	8,436	8,436	8,733
Shenyang	5,849	6,498	7,220	6,256	5,500	6,078	7,105 ⁴	5,718	7,026	8,294	8,294	7,163
Zhengzhou	3,630	4,401	5,323	4,254	4,175	4,942	6,765	3,560	6,461	5,810	5,810	5,250
Qingdao	4,512	5,041	5,620	3,719	5,100	5,816	5,911	4,556	6,165	6,353	6,353	5,574
Suzhou	3,997	5,336	7,070	5,346	4,000	5,246	4,979	4,084	5,893	10,722	10,722	NA
Jinan	3,965	4,613	5,360	4,336	3,325	3,789	4,026	4,120	5,648	4,694	4,694	5,301
Changsha	3,630	4,401	4,578	3,094	3,125	3,657	4,242	3,193	5,630	5,289	5,289	4,534
Kunming	3,470	3,927	4,443	3,855	3,275	3,649	3,755	3,385	5,273	4,677	4,677	NA
Harbin	5,140	5,730	6,387	5,879	4,350	4,815	4,458	4,596	5,243	5,282	5,282	6,335
Shijiazhuang	3,353	3,717	4,114	2,835	2,875	3,367	3,706	3,317	5,090	4,304	4,304	NA
Hefei	3,027	3,608	4,242	3,310	3,350	3,665	4,628	3,099	5,056	3,865	3,865	NA

¹projection; ²2012, ³incl. Xianyang, ⁴incl. Fushun

Sources: [1], [2], [3], [4], [5], [6]

Based on Shenyang, it can be seen that the city grew steadily in the period of 2015–2016 and then again in 2017–2018. The decrease in its population by ca. 1 million in 2019 (with the area unchanged) is unjustified. Just as big a decrease (in absolute amounts), and even bigger with regard to the number of inhabitants, is the decrease in the population of the urban complex of Zhangjiaggang-Jiangyin in the period of 2018–

2019, when the area grew, or in Xiamen (2016–2017). Also inconsistent with the actual situation were remarkable decreases in the population of other large cities of Jiangsu (Suzhou, Wuxi, Changzhou) in the period of 2018–2019. In Fuzhou, there was a three-fold growth in the area in the period of 2016–2017, while in 2018–2019 it shrank again; even though the area in 2019 is 64% bigger than in 2016, the population is markedly

Table 5. Changes in population and area of selected cities/urban areas in source [3]

City/Urban Areas	Province	2015		2016		2017		2018		2019		2020		2021		2022	
		P	A	P	A	P	A	P	A	P	A	P	A	P	A	P	A
Chongqing	CQ	7,217	932	7,440	971	7,990	1,010	8,875	1,489	8,300	1,489	7,739	1,537	7,261	1,536	12,135	1,580
Shenyang-Fushun ¹	LN	6,078	1,010	6,200	1,010	7,935	1,502	8,095	1,502	7,055	1,502	7,105	1,515	7,208	1,515	7,964	1,551
Suzhou	JS	5,246	1,127	5,380	1,269	5,845	1,269	6,175	1,360	5,250	1,373	4,979	1,387	5,103	1,386	6,031	1,386
Wuxi	JS	3,597	738	3,670	738	3,745	842	3,825	842	3,480	842	3,642	852	3,681	852	4,479	982
Changzhou	JS	3,425	829	3,515	829	3,670	958	3,770	958	3,210	958	3,647	1,172	3,711	1,171	4,164	1,251
Xiamen	FJ	4,420	583	4,715	583	4,000	751	4,265	751	3,975	699	4,773	715	4,964	715	5,201	715
Fuzhou	FJ	3,962	440	4,080	440	5,245	1,243	5,400	1,243	3,940	725	4,039	746	4,160	746	4,443	746
Ningbo	ZJ	3,753	738	3,895	738	3,600	932	3,735	932	3,525	932	3,649	944	3,823	943	4,339	943
Zhangjiagang-Jiangyin	JS	3,056	673	3,225	673	3,405	712	3,595	712	2,570	816	2,420	824	2,546	824	2,742	824

¹2015-2016 – only ShenyangExplanations: P – Population [thou.], A – Area [km²]

smaller. In Ningbo, a decrease in the population (2016–2017) was recorded despite its area having been extended; and in the period of 2018–2019 the population declined again although the area remained constant. Source [3] abounds in similar cases, which makes the data presented there very unreliable.

Source [4] includes detailed data on populations of cities in census years: 1990, 2000, 2010, 2020 and populations of different-level administrative units (province-level units, prefecture- and county-level units), among others: 2000, 2010, 2020 censuses and towns: 2010. The publication is enriched with interactive maps.

Generally, the data apply to the whole population of urban areas – within the borders of districts connected to a particular city. In cases where the prefecture-level city includes (besides the core centre) county-level cities, the urban population of the latter is treated as the population of these cities. For example: the prefecture-level city of Suzhou Shi (Jiangsu) comprises five districts and four county-level cities (Kunshan, Changshu, Zhangjiagang and Taicang). The urban population of the four mentioned cities is counted separately; it is not included in the calculations for Suzhou (which is correct under the accepted assumptions). Such an approach inflates the population of the city much less than the method suggested in source [5], which will be presented later in the text. This does not mean a complete lack of interpretative problems; while still on the case of Suzhou, it suffices to mention its south district – Wuijiang – which covers, for instance, several towns separated from Suzhou by rural areas. Located in the south of the district, Shengze Zhen is 50 km from the centre of Suzhou and 25 km from the border of the urban area. The centre of Jiaxing in the province of Zhejiang is much closer (17 km). Such a unit as Shengze Zhen could be treated as a separate city (244 thousand in 2010).

One huge merit of source [4] is the method by which the category of cities includes not only the urban population of centres that administer prefecture-level cities and county-level cities, but also centres administering counties, which are equal with the latter in the administrative hierarchy. As a result, the number of cities listed in the discussed source is the biggest (over 1,000) and the closest to reality.

Source [5] presents population data for 66 cities of over 1 million people in the PRC with no information on the year, nor any precise definition; yet, they are called “urban areas” (Fr. *aire urbaine*). The figures suggest that the entire population

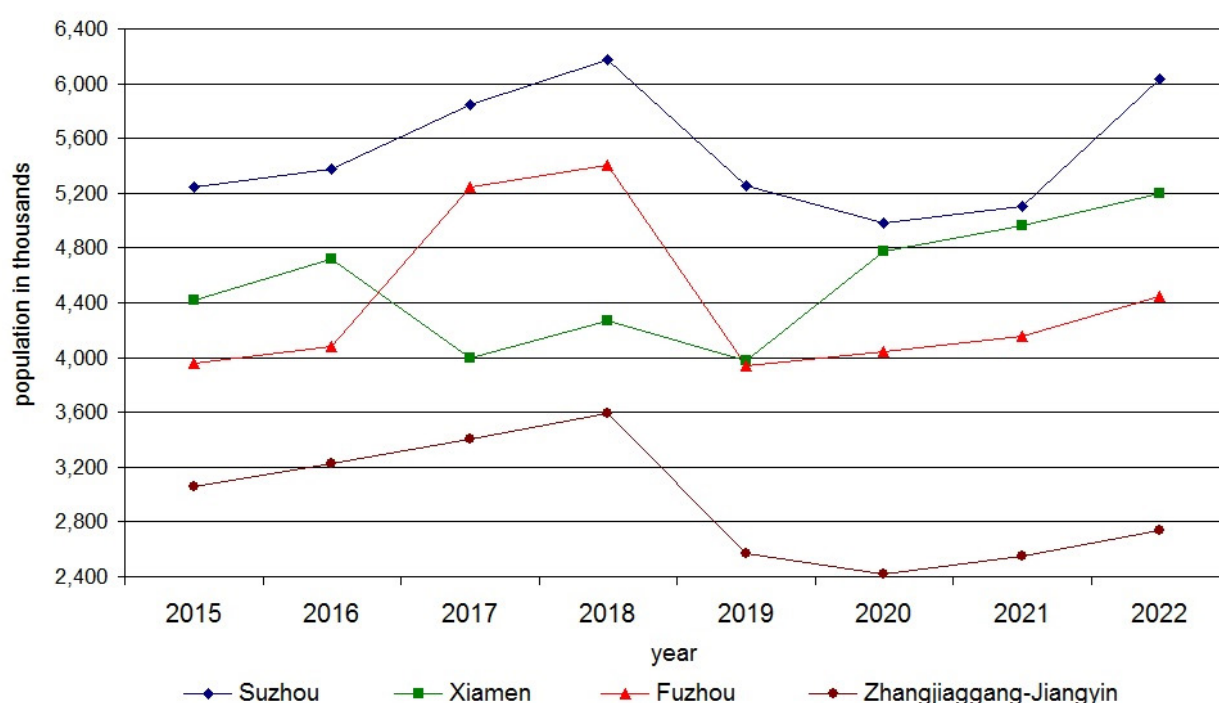


Fig. 4. Population discrepancies in selected cities/urban areas by source [3]

Source: data from Table 5

(urban and non-urban) of administrative units related to a particular city (prefecture-level city) is taken into account, and in many cases even the total population of adjoining units that do not necessarily share administrative or economic ties with the same city. For example, the population of Shantou (14.252 million) includes three prefecture-level cities (Shantou, Chaozhou and Jieyang) of a total area of 10,660 km²; the data for Tianjin (15.621 million) include the population of the whole province-level city (11,610 km²); and the data for Suzhou (10.722 million) includes the population of the prefecture-level city (6,094 km²), which includes Kunshan, Changshu, Zhangjiaggang and Taicang. The analysed source, in principle, highly overestimates the actual population of cities.

Source [6] provides data on cities and urban areas listed as numbering 418. The information on cities is full of mistakes that cannot be reasonably explained, such as Xi'an (360 thousand in 2020), Suzhou in Jiangsu (205 thousand in 2020), and the population of urban areas is, in many cases, smaller than that of cities (Note 12). Juxtaposing it with other sources, and based on the present author's own estimates, it can be concluded that in some cases the data from the "cities" column are more correct and in other cases it is the data from the "urban areas" column that can be more relied on.

6. Causes of variations in population of cities, sources of mistakes and differing interpretations

The differences in city population data are clearly associated with the delimitation of borders of a spatial unit (urban area) and with estimates of changes occurring in a given period. More precisely, the greatest impact on the variations in population of cities is ascribed to the following factors: 1) birth rate, 2) migrations, and 3) spatial changes involving transformation of areas in the direct neighbourhood of the city into urban areas (or the merging of separate urban areas as the areas between them become urbanised). In reality, the list of factors affecting the disparities in population analyses and estimates is more extensive.

• The Hukou system and temporary population

Hukou (*hujī*) is the system of household registration in China (mainland), and has its origins in ancient China. A household registration record identifies a person as a resident of an area (Miller, 2012; Kroeber, 2016; Gibson & Li, 2017). In its current form, the *hukou* system came into being with the

1958 People's Republic of China Hukou Registration Regulation (Note 13). Until very recently, each citizen was classified into an agricultural (rural) or non-agricultural (urban) *hukou* (Young, 2013). This organisational structure was linked to the social policy. Inflows of migrant workers without local *hukou* to major cities are an important part of the China's economic growth (Miller, 2012; Gibson & Li, 2017) and growth of urban population (Chan & Wan, 2017).

From the perspective of calculating the urban population, the problem was that, before 2014, some sources (e.g. [7]) only included the population formally registered as urban (in-migrants were considered as temporary population). In the case of some cities, particularly those dynamically growing and economically attractive to rural population, the differences were very large. For example, the population of Shenzhen presented by [7] was 719 thousand as of 2000, i.e. 11.1% of the actual population at that time according to [4], and 11.0% following [1].

In 2014 the state published and partially implemented the "National New-type Urbanisation Plan (2014–2020)" to tackle various problems derived from China's fast urbanisation process; the plan aimed to narrow the inequalities between urban residents who do not hold urban *hukou* and urban residents who do hold urban *hukou* (Chan, 2015). The plan also aims to eliminate the differences between agricultural and non-agricultural *hukou* status (Wang et al., 2015). As a result, sources are now more likely to ignore the issue of *hukou* and only count real population.

Potentially a fairly accurate way to identify the number of temporary population seems to be the use of the mobile-phone signaling data. However, it should be borne in mind that the use of this method in PRC may be associated with at least two problems. First: data on owners of mobile phones may be confidential to international organizations or research units. Second: the data on the number of mobile phones does not cover all people, for example some children. When researching on appropriate statistical samples, the latter problem can be minimized with a satisfactory accuracy.

• Cities left out of compilations

Some sources ([1], [2], [3]) take into account only cities with a particular administrative status: individual cities at the province level, cities that administer prefecture-level cities, and county-level cities. All of the mentioned units cover both stricte

urban areas and suburbia, as well as rural areas of agricultural nature. At the third administrative level, besides county-level cities, there are two other types of spatial units: district (*Qu*) and county (*Xian*), which may also cover different kinds of areas (urban development, suburbia, rural development). The population of urban areas within district borders is, in principle, categorised as urban population in statistical sources. The same is not true of counties; it is usually the city that is the administrative centre of this unit (officially known as *Zhen* – a town). It usually has several hundred thousand inhabitants, who are omitted in some compilations of cities on account of the status of the unit, whereas the city-defining features should be taken into consideration instead, e.g. size, density, heterogeneity, functions, character of the economy, building development, infrastructure, etc. Two examples were given above: Jiangsu and Zhejiang; similar cases are found all across the country.

The omission of some cities in compilations sometimes has one more cause. Some cities belong to third-level administrative units (usually county-level cities or counties), where they are not core centres. Some of them are topographically coupled with neighbouring county-level cities. Statistics also leave out urban areas, by including the populations of individual *Zhens* into other more remote units. An example: in the south part of Wenzhou Shi (prefecture-level city), on opposing banks of the Aojiang River, there are: Longgang Zhen, which is administratively subordinate to Cangnan Xian (Note 14) (county); and Aojiang Zhen, which is administratively subordinate to Pingyang Xian (Fig. 5). In compilations they are listed neither as one urban area nor separately. Longgang Zhen (Note 15) (337 thousand in 2010) together with Lingxi Zhen (260 thousand) and others are treated as parts of Cangnan City (648 thousand of urban population), while Aojiang Zhen (169 thousand), together with Kunyang Zhen (121 thousand), Xiaojiang Zhen (65 thousand) and others belong to Pingyang City (375 thousand of urban population) – all of the figures are taken from [4]. Neither do some sources (e.g. [1], [2], [3]) include Cangnan City and Pingyang City on account of their status (Xian).

In 2019, approved by the Ministry of Civil Affairs, Longgang was proclaimed to be a county-level city by the provincial government of Zhejiang. In 2020, only source [4] distinguishes Longgang as a separate city (450 thousand). Situated across the river, Aojiang Zhen continues to be part of Pingyang City.



Fig. 5. Unusual spatial connections of cities determined administratively: Cangnan City and Pingyang City in Zhejiang Province (discussion in the text)

Explanations: line – border between counties, yellow letters – towns.

Source: adaptation of Google Maps

• Units with two-way connections

Another issue is related to units whose spatial and functional connections with larger cities differ from their administrative subordination. An example: Chao'an is a district (*Qu*) located partly between the cities of Chaozhou and Shantou (Guangdong). Some towns in that district, primarily Anbu Zhen (155 thousand in 2010) and Caitang Zhen (137 thousand) are situated closer to and have stronger topographical connections with Shantou rather than Chaozhou; nevertheless, due to their administrative subordination (Chao'an *Qu* belongs to Chaozhou Shi) they are treated as parts of Chaozhou City (see Fig. 6). Another case in point is Huiyang *Qu* (Guangdong). Here, because of their administrative subordination, the majority of urban subdistricts

(particularly Danshui Jiedao and Qiuchang Jiedao) are considered in most sources as belonging to Huizhou, which is separated from them by a small mountain range, rather than being considered part of nearby Shenzhen, which they are connected to by continuous urban development.

• Delimitation of cities in zones with several neighbouring urban areas

Generally, vast urbanised zones are separated by small rural areas, and are usually found in coastal regions, where intensive industrialisation and urbanisation processes are conspicuous in areas close to harbours. For example, Shantou is surrounded by numerous large towns (Fig. 6) that can hardly

Table 6. Examples of discrepancies in population (thousands) between sources for selected cities/urban areas

Location	2010 [1]	2015 [1]	2019 [1]	2019 [2]	2010 [3]	2015 [3]	2020 [3]	2010 [4]	2020 [4]	2020 [5]	2020 [6]
Chongqing	11,244	13,372	15,354	7,458	6,200	7,217	7,739	6,264	9,581	15,384	15,679
Zhengzhou	3,630	4,401	5,131	4,254	4,175	4,942	6,765	3,677	4,459 ⁵	5,810	5,251
Quanzhou	1,139	1,391	1,631	184	5,600 ¹	6,710 ¹	6,345 ¹	1,155	1,469	-	1,671
Xiamen	3,040	3,395	3,652	3,531	3,200	4,420	4,773	3,119	4,617	13,765 ³	3,695
Shantou	3,623	3,958	4,249	5,329	2,225	2,419	2,408	3,644	3,839	14,252 ¹	4,298

¹incl. Jinjiang & Shishi, 2 incl. Jinjiang, Shishi & Quanzhou, 3 incl. Quanzhou, 4 incl., Chaozhou, Jieyang, Puning, 5 in 2018.

Source: [1], [2], [3], [4], [5], [6]



Fig. 6. Mosaic-like, urbanised zone between Shantou, Chaozhou, Jieyang and Puning – the separation of the urban areas of individual cities is blurred (especially between Shantou and Puning, Shantou and Chaozhou); the non-arbitrary separation of individual cities is practically impossible and leads to different delimitations.

Explanations: line – borders of prefecture-level cities, capital yellow letters (eg Jieyang) – prefecture-level cities, lower case yellow letters (eg Jiedong) – districts or county-level cities.

Explanations: line – border between counties, yellow letters – towns.

Source: adaptation of Google Maps

be unambiguously linked with neighbouring cities. Between Shantou in the east, and Puning and Jieyang in the west, there are Chaoyang Qu and Chaonan Qu, which do not have a clearly developed central part with big-city characteristics. For that reason, the towns and urban districts that belong to

Chaoyang Qu and Chaonan Qu (Lugang Zhen, Heng Zhen, Tongyu Zhen, Liangyin Zhen, and others) are considered by some sources [4] as being part of Shantou, while [3] distinguishes a unit comprising Chaoyang and Chaonan joined. Elsewhere, i.e. in

[1] and [2], Chaoyang is distinguished with no mention of Chaonan.

As a result of the problems arising from the ambiguous delimitation of units in vast, partly urbanised zones which resemble a mosaic, there are different numbers of cities and substantial discrepancies in their population (cf Table 6).

The situation is similar for Quanzhou (Fujian): when considered in its narrow borders ([1], [4]) it has approximately 1.5 million inhabitants; however, when its vast, mosaic-like, urbanised area is taken into consideration (similar to Fig. 6), it has circa 6.5 million inhabitants [3]. Source [2] provides preposterously low figures for this city (184 thousand), while [5] hugely enlarges the range of the urban area, from Quanzhou up to Xiamen (Table 6), in spite of the considerable distance between those cities (about 65 km as the crow flies, about 80 km by main roads). In the middle variant, the urban area of Quanzhou also covers the nearby cities of Jinjiang and Shishi, which – according to [1] – are populated by only 394 thousand and 466 thousand, respectively, thus – all in all – the total population cannot amount to ~6.5 million. Interestingly, the total population of the four units which comprise the centre of Jinjiang: Chendai Zhen (370 thousand), Chidian Zhen (154 thousand), Qingyang Jiedao (104 thousand) and Meiling Jiedao (67 thousand) greatly exceeds the figure provided by [1]. These discrepancies are due to different categorisations of units with other administrative statuses: only six Street Resident Committees/urban subdistricts (*Jiedao*): Qingyang, Meiling, Xiyuan, Luoshan, Xintang, Lingyuan, are included, while all towns (*Zhen*) with urban characteristics and coupled with them morphologically are left out.

The third example of problems with the mosaic urbanisation is encountered with a complex of cities in Jiangsu province: Zhangjiagang and Jiangyin are treated individually as cities ([1], [4], [6]) or jointly [3] (Table 5), or are altogether omitted ([2], [5]).

- **Inclusion of exceedingly vast areas**

This usually results in inflated populations. Chongqing is a good case in point: ([1], [5], [6]) establishes its population at about 15 million, i.e. twice the figure found in the other sources (cf Table 4, 6). Such cases are common on a smaller scale, which is due to the inclusion of the whole urban population of the area administered by a particular

city into the population of that city, including the population of towns (*Zhen*) that are a considerable distance from the city and separated from it by agricultural lands.

- **Different qualification of complexes of integrating cities**

This problem has objective causes; the interpretation of particular cases differs between individual sources on the one hand, and, on the other, it may also change over time as spatial development and functional connections between areas change. It is possible to indicate at least a few pairs of neighbouring cities that constitute separate administrative units at the prefecture level and that – due to their spatial closeness and plethora of functional connections – are integrating with each other. That integration is reflected, among other things, in merging building development, and sometimes in the construction of joint systems of public transportation. The above can be exemplified for example by: Xi'an-Xianyang (Shaanxi), Taiyuan-Jinzhong (Shanxi), Guangzhou-Foshan (Guangdong), and Shenyang-Fushun (Liaoning). Similar pairings are also created between cities with different administrative statuses (e.g. a prefecture-level city with a county-level city), such as: Wenzhou-Rui'an (Zhejiang).

Individual sources list such units as separate cities or integrated urban areas. As cities develop to gain a big-city character, the connections between cities will progressively strengthen; also, new pairs and complexes of cities with similar characteristics will be increasingly encountered (e.g. Hangzhou-Shaoxing in Zhejiang, Changsha-Xiangtan-Zhuzhou in Hubei).

- **Changes in qualification of urbanised areas in connection with changes in administrative status**

This situation is observed, for instance, in the suburbs of Chengdu (Sichuan). Central districts of Chengdu are surrounded by a ring of five external neighbourhoods that also have the status of district (*Qu*): Longquanyi, Shuangliu, Wenjiang, Xindu and Pidu. Before November 2016, the latter had the status of county (it was then known as Pixian Xian). Until 2016 the population of Pixian was left out from calculations of the population of Chengdu urban area; however, upon the status change, the people of Pidu Qu were included into the population of Chengdu [4].

Similar cases of misclassification of units were recorded after the last national population census (2020). Some cities appear twice in the source [4], which is due to the separate counting of urban population within district boundaries (Qu) and separately, under the same name, within county boundaries (Xian). Thus, Changsha in Hunan has a population of 5,630 thousand, and the urban population of Changsha County (section of the city adjacent to the central districts from the east) – 1,024 thousand. Correctly calculated city population should be the sum of these two values. Analogous situations occur, for example, in the case of Xiangtan in Hunan (978 thousand and 357 thousand), or Nanchang in Jiangxi (3,519 thousand and 916 thousand).

• **Inaccuracy of calculations within a single source**

This problem was discussed above in the case analysis – source [3] (Table 5).

7. Discussion

Mistakes arise for diverse reasons, despite the fact that units are defined similarly as urban areas and that the differences arising from the estimation of growth dynamics are generally small. Such quantitative discrepancies are unavoidable and may present particular cities in different lights in different compilations. However, more serious mistakes were found to abound; they are most often due to:

- differences between the actual and registered (hukou system) population – that problem is becoming less important, but has not been eliminated;
- erroneously identifying administrative units of various levels as being cities, when they may be called “cities” but in fact include vast rural areas;
- mistakenly including into urban areas only those units of a particular administrative status irrespective of their actual level of urbanisation, e.g. omitting units that are county seats (Xian) in compilations of cities, omitting numerous towns (Zhen) in the total population of cities;
- inexplicable fluctuations in data within one source over subsequent years that cannot be justified by area changes (best exemplified in populations that decreased as their area increased, or vice versa);

- mistakes in calculating the population of delineated spatial units (observed as considerable overestimation or underestimation of figures, even by several hundred percent in extreme cases).

The differences related to the delimitation of urban areas in zones characterised by mosaic urbanisation are more objective.

The comparative analysis proved that the main cause of discrepancies in data on the populations of Chinese cities should not be seen in different methods of defining “a city”, because all of the cited sources declare that population was calculated within the borders of urban areas (or urban agglomerations). Where the problems originate is in the diversity of approaches to the qualification of spatial units with various administrative statuses, and in mistakes and inconsistencies caused by departing from the established definitions. Thus, the hypothesis formed at the beginning was verified positively.

To conclude, it is possible to provide a general guideline on a calculation method applicable to the population of urban areas in China that would reduce discrepancies between data sources. Three basic criteria should be consistently met:

1. Populations of urban areas should include urban populations of all political and administrative units, regardless of their status (e.g. not only county-level cities, but counties as well); errors of this type were shown, inter alia on the examples of cities of north Jiangsu (Suining, Shuyang, Siyang, Sihong), the cities around Jiaying in north Zhejiang, and cities in the country's southern provinces: Changsha, Xiangtan and Nanchang.
2. The total population of a particular city should include the urban populations of all spatial units connected with that city, regardless of their status (not only urban districts, but towns as well, bearing in mind that towns in the PRC may count as many as several hundred thousand inhabitants and may have big-city characteristics); the problem is illustrated by the example of Quanzhou in Fujian province, where large population differences result from the omission of large towns.
3. Proper distinction should be made between urban areas (and urban population) within the borders of political and administrative units of the first, second or third level so that they are better identified with various cities belonging to a particular unit (if there is more than

one city), instead of pooling the populations of smaller towns with that of the core city; this problem and its consequences have been demonstrated by the examples of towns situated between Shantou and Chaozhou in Guangdong Province, Cangnan City and Pingyang City in Zhejiang Province, and others.

8. Concluding note

Population lists of Chinese cities based on different registration or estimation systems, as well as the naming of cities in administrative units of different levels must result in numerous misunderstandings and errors. In addition, the matter is complicated by wide boundaries of these units, usually covering vast areas inhabited by agricultural population. Other authors' publications (e.g. Chan, 2007) discuss various systems of population statistics in China for province- and prefecture-level cities. Still in the 21st century the administrative system, especially the *hukou*, remains the mainstay of statistical methods for calculating population. Already at the end of the last century, the Chinese scholars (Zhou & Shi, 1995) postulated the need to simplify China's systems of population statistics, revise the terminology and setup.

Due to the careful study by other authors of the basic source problems related to the application of various systems of population statistics, the main aim of the article was to focus on the consequences of various approaches. It has been shown that the data disseminated by individual international institutions differ not only from one another (differences between sources), but often there is no consistency within one source: only some units of a certain level are included in cities (why not all?), and the data for some cities are characterized by significant population fluctuations in the following years (the population cannot decrease by a few million from year to year and then increase again, or *vice versa*).

Prior research focuses primarily on problems arising from the poor relation of *hukou* registration with *de facto* population and incorrect delimitation of city/urban area boundaries. This study identifies diverse source and interpretive problems and demonstrates them through numerous examples of cities. These are problems of a different nature from those identified by the authors cited above, and they are sources of error significant enough that they cannot be ignored. A major issue is the incorrect classification of small administrative units (towns,

urban subdistricts, and others) as urban areas and the variability in methods for calculating population due to inconsistent treatment of definitions.

The comparison of the data on the populations of cities in the PRC leads to the conclusion that no source is free of mistakes. Irregularities are fewest in [1] and [4], and most numerous in [2] and [5], where a lot of information might be deemed not so much inaccurate as fictitious. As far as [3] is concerned, its huge merit is in its systematic data updates and, if it were not for its ungrounded "fluctuations" in data, that source would be in line with [1] and [4] and could be considered alongside them as being relatively reliable.

Results of the seventh national population census in the PRC (1 November 2020) will allow for more precise estimates of the size of cities, such as will narrow the gap between the calculations and the actual situation in the forthcoming years. However, it is still a priority that precise criteria be set for delimiting urban areas and that they be followed consistently, because the census will only establish exact population numbers for administrative units. It is worth noting here that the population of cities can also be reasonably estimated based on alternative data (if such data is available in the PRC), such as data on night lighting, telephone signalling, remote sensing, etc. Such information sets could complement and verify each other. However, the authors of databases most often rely on official statistics as source data (usually these are census data collected every 10 years), which are updated in subsequent years based on estimates and projections.

The projection method is generally not characterized in detail; the recipient receives the finished 'product' in the form of numbers. Some sources, e.g. [3] also provide information on urban areas and monitoring of changes in the terrain. Year-to-year fluctuations in these figures (rather than a fairly steady increase) indicate that this methodology is not being applied consistently.

Notes

1. T. Brinkhoff: China, <https://www.citypopulation.de/en/china/> (accessed 2021.12.20).
2. Except peripheral administrative units of the first level (Gansu, Heilongjiang, Inner Mongolia, Jilin, Ningxia, Qinghai, Xinjiang, Yunnan), whose areas are often very large due to a small population density (mountainous areas, deserts, etc.) and two county-level cities in the mountainous areas of Sichuan (Kangding and

- Barkam), the average area of the county-level city (altogether 271 units) equals 1,783 km².
3. China's Political System, <http://www.china.org.cn/english/Political/28842.htm> (accessed 2020.01.30).
 4. In the structure of province-level cities there is no distinction into the second hierarchical level: prefecture-level cities.
 5. T. Brinkhoff: China, <https://www.citypopulation.de/en/china/> (accessed 2020.01.30).
 6. The similarity of meaning between the terms, excluding MAs, is explicitly mentioned by some sources: "An urban area (built-up urban area or urban agglomeration) is fundamentally different from a metropolitan area. A metropolitan area is a labor market (and a housing market). It includes a principal built-up urban area (the largest built-up urban area in the metropolitan area) as well as economically connected rural areas (and smaller urban areas) to the outside." Wendell Cox, 'Demographia World Urban Areas', 15th Annual Editions (PDF), (St. Louis: Demographia, 2019), p.6.
 7. "For 2000, population of city districts with average population density of at least 1,500 persons per square kilometre, population of suburban-district units and township-level units meeting certain criteria, such as having contiguous built-up area, being the location of the local government, or being a street (jiedao) or having a resident committee. For 2010, urban residents meeting the criterion defined by the National Bureau of Statistics of China in 2008, i.e., the criteria used in the 2000 census plus residents living in villages or towns in outer urban and suburban areas that are directly connected to municipal infrastructure and that receive public services from urban municipalities." United Nations, Department of Economic and Social Affairs, Population Division 'World Urbanization Prospects: The 2018 Revision, Online Edition', <https://population.un.org/wup/Download> [File 12: Population of Urban Agglomerations with 300,000 Inhabitants or More in 2018, by country, 1950-2035 (thousands)], p.84.
 8. Further in the paper, in the section discussing the city (excluding the word 'city' used to mean a unit of the political and administrative division of the PRC), the author always refers to the urban area or urban agglomeration – depending on the cited source.
 9. The largest unit of the urban district type in that group of cities is Shucheng Jiedao (Shuyang), with 410 thou inhabitants (2010).
 10. In the light of the estimates made by [1], which inform that there are 424 cities with population >300 thou and which leave out a lot of units, the discussed compilation is even more incomplete.
 11. In Putian only the core urban subdistricts (Jiedao): Chiwei, Liushaxi, Liushabei, Liushanan i Liushadong, altogether had 501 thou in 2010.
 12. The term 'cities' does not refer here to political and administrative units.
 13. More information about the hukou population can be found, for example, in Kamal-Chaoui, L., Leeman, E. & Rufeï, Z. (2009), Lu (2012).
 14. In August 2019 Longgang was proclaimed to be a county-level city.
 15. County-level city from August 27, 2019.

References

- Cartier, C.** (2015). Territorial Urbanization and the Party-State in China. *Territory, Politics and Governance*, 3(3): 294-320. DOI: <https://doi.org/10.1080/21622671.2015.1005125>.
- Chan, K.W.** (2007). Misconceptions and complexities in the study of China's cities: Definitions, statistics, and implications. *Eurasian Geography and Economics*, 48(4): 383-412. DOI: <https://doi.org/10.2747/1538-7216.48.4.383>.
- Chan, K.W.** (2015). Five Decades of the Chinese Hukou System. In: *Handbook of Chinese Migration: Identity and Wellbeing*, 23-47, Edward Elgar Publishing, Inc.: Northampton, MA.
- Chan, K.W. & Hu, Y.** (2003). Urbanization in China in the 1990s: New Definition, Different Series, and Revised Trends. *The China Review*, 3(2): 49-71.
- Chan, K.W. & Wan, G.** (2017). The size distribution and growth pattern of cities in China, 1982–2010: analysis and policy implications. *Journal of the Asia Pacific Economy*, 22(1): 136-155. DOI: <https://doi.org/10.1080/13547860.2016.1266829>.
- Chien, S.-S.** (2010). Prefectures and prefecture-level cities: the political economy of administrative restructuring. In: Chung J. H., Lam, T.-C. (Eds.). *China's Local Administration. Traditions and changes in the sub-national hierarchy*, 127-148). Routledge: New York, NY.
- Gibson, J. & Li, C.** (2017). The Erroneous Use of China's Population and Per Capita Data: a Structured Review and Critical Test. *Journal of Economic Surveys*, 31(4): 905-922. DOI: 10.1111/joes.12178.
- Kamal-Chaoui, L., Leeman, E. & Rufeï, Z.** (2009). *Urban Trends and Policy in China*, OECD Regional Development Working Papers 2009/1, OECD Publishing. DOI: [10.1787/225205036417](https://doi.org/10.1787/225205036417).
- Kirkby, R.J.** (1985). *Urbanization in China: Town and Country in a Developing Economy 1949–2000 A.D.*, Columbia University Press: New York, NY.
- Kroeber, A.R.** (2016). *China's Economy: What Everyone Needs to Know?*, Oxford University Press.

- Lu, D.** (2012). *The Great Urbanization of China*, Series on contemporary China, 30, World Scientific Publishing: Singapore.
- Maik, W.** (1997). *Podstawy geografii miast* (Basics of city geography – in Polish). UMK: Toruń.
- Miller, T.** (2012). *China's Urban Billion: The Story behind the Biggest Migration in Human History*, Zed Books: London-New York. DOI: [10.1007/s11366-015-9379-6](https://doi.org/10.1007/s11366-015-9379-6).
- Pannell, C.W.** (2003). China's Demographic and Urban Trends for the 21st Century. *Eurasian Geography and Economics*, 44(7): 479–496. DOI: <https://doi.org/10.2747/1538-7216.44.7.479>.
- Shen, J.** (2005). Counting urban population in Chinese censuses 1953–2000: changing definitions, problems and solutions. *Population, Space and Place*, 11(5): 381–400. DOI: <https://doi.org/10.1002/psp.382>.
- Shen, J.** (2006). Estimating Urbanization Levels in Chinese Provinces in 1982–2000. *International Statistical Review*, 74(1): 89–107. DOI: <https://doi.org/10.1111/j.1751-5823.2006.tb00163.x>.
- Sokołowski, D.** (1998). Niektóre problemy definiowania pojęć geograficzno-osadniczych związanych z urbanizacją (Some problems of defining geographical notions related to urbanisation processes – in Polish). *Czasopismo Geograficzne*, 69(2): 169–192. DOI: <https://doi.org/10.5281/zenodo.1146842>.
- Szymańska, D.** (2007). *Urbanizacja na świecie* (Urbanization in the world – in Polish). PWN: Warszawa.
- Wang, X.-R., Hui, E.C., Choguill, C. & Jia, S.-H.** (2015). The New Urbanization Policy in China: Which Way Forward. *Habitat International*, 47: 279–284. DOI: <http://hdl.handle.net/10397/27398>.
- Young, J.** (2013). *China's Hukou System: Markets, Migrants and Institutional Change*, Palgrave Macmillan: Basingstoke. DOI: [10.1057/9781137277312](https://doi.org/10.1057/9781137277312).
- Zhou, Y. & Ma, L.C.J.** (2003). China's Urbanization Levels: Reconstructing a Baseline from the Fifth Population Census. *The China Quarterly*, 173: 176–196.
- Zhou, Y. & Ma, L.C.J.** (2005). China's Urban Population Statistics: A Critical Evaluation. *Eurasian Geography and Economics*, 46(4): 272–289. DOI: <https://doi.org/10.2747/1538-7216.46.4.272>.
- Zhou, Y. & Shi, Y.S.** (1995). Jianli zhongguo chengshi de shiti diyu gainian (Toward Establishing the Concept of Physical Urban Area in China - in Chinese). *Dili xuebao (Acta Geographica Sinica)*, 50(4): 289–301.
- <https://www.worldometers.info/demographics/china-demographics/#urb>.
- [3] Wendell Cox, Demographia World Urban Areas. 10-18th Annual Editions (PDF). (St. Louis: Demographia, 2014-2022). Available at: <http://www.demographia.com/>.
 - [4] Thomas Brinkhoff: China, Available at: <https://www.citypopulation.de/en/china/>.
 - [5] PopulationData.net, <https://www.populationdata.net/pays/chine/aires-urbaines>.
 - [6] PopulationStat. World statistical data, <https://populationstat.com/china/>.
 - [7] World Gazetteer, <http://world-gazetteer.com> (Retrieved 2004-10-10, 2010-12-18).

Main sources

- [1] United Nations, Department of Economic and Social Affairs, Population Division (2018). *World Urbanization Prospects: The 2018 Revision*, Online Edition. Available at: <https://population.un.org/wup/Download> [File 12: Population of Urban Agglomerations with 300,000 Inhabitants or More in 2018, by country, 1950-2035 (thousands)].
- [2] World Population Review. Available at: <http://worldpopulationreview.com/countries/china-population/>

