Integrated Circuit Enterprises Dataset Using Qichacha and Baidu Map Platforms and Data Mining Technology: Taking the Pearl River Delta Region of China (1980–2020) as an Example

Wang, P. H.[1,2]   Zha, R. B.[2*]   Huang, J. C.[1,3,4*]   Huang, Y.[2]   Du, S. Y.[2]   Xu, J. R[2]

1. Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China;
2. College of Cultural Tourism and Public Administration, Fujian Normal University, Fuzhou 350108, China;
3. Key Laboratory of Regional Sustainable Development Modeling, Chinese Academy of Sciences, Beijing 100101, China;
4. College of Resources and Environmental Sciences, University of Chinese Academy of Sciences, Beijing 100049, China

Abstract: The integrated circuit (IC) industry is the foundation of the electronic information industry, the core of transforming traditional industries, and the key to promoting the continuous development of emerging strategic industries. As the Pearl River Delta (PRD) region hosts one of the agglomerations of China's IC industry and the frontline of China's reforming and opening up, a spatial and temporal study of the distribution of IC enterprises in the PRD region could deepen our understanding of the current state of IC enterprises in China and provide a lesson for the development of IC industries in other regions. This paper collects, processes, and summarizes data based on the enterprise information publishing platform Qichacha and the Baidu map opening platform to obtain data on the geographical locations, establishment dates, and operation statuses of IC enterprises in the PRD region, and further uses the kernel density to explore the spatial-temporal distribution of IC enterprise in the PRD region. The dataset is archived in .shp format and consists of 7 data files with a data size of 45.8 MB (compressed to 1 file, 710 KB).

Data Citation: [1] Wang, P. H., Zha, R. B., Huang, J. C., *et al*. Integrated circuit enterprises dataset using Qichacha and Baidu Map platforms and data mining technology: taking the Pearl River Delta region of China (1980–2020) as an example [J]. *Journal of Global Change Data & Discovery*, 2022, 6(4): 631–637. https:// doi.org/10.3974/geodp.2022.04.14. https://cstr.escience.org.cn/CSTR:20146.14.2022.04.14.
[2] Wang, P. H., Zha, R. B., Huang, Y., *et al*. Integrated circuit enterprises dataset using Qichacha and Baidu Map platforms and data mining technology: taking the Pearl River Delta region of China (1980–2020) as an example [J/DB/OL]. *Digital Journal of Global Change Data Repository*, 2022. https:// doi.org/10.3974/geodb. 2022.10.07.V1. https://cstr.escience.org.cn/CSTR:20146.11.2022.10.07.V1.

1 Introduction

The integrated circuit (IC) industry is a leader in China's national economy. It has produced a new generation of information technologies with epoch-making significance, such as personal computers, the Internet, the Internet of Things, smartphones, digital images, big data, cloud computing, and artificial intelligence. Its applications have penetrated all aspects of China's economy and become an important part of people's production and lives. In the current international environment, the IC industry has become a strategic of China's economy. It is not only an important factor in international political and economic competition but also the industry with the most intense international competition and the most thorough global resource flow and allocation. If a country (or a region) can occupy a commanding position in the IC industry, it can earn large profits on the world market and lead the trend of world economic development. However, too much reliance on overseas IC industry, could lead to constrain in key areas. Based on the dual needs of international competition and national security, the development of China's IC industry is increasingly urgent[1,2].

The development of China's IC industry began in the late 1970s. After nearly 30 years of exploration, it entered a high-speed development stage in 2009. From 2009 to 2019, the average annual sales of China's IC industry increased by 64.53 billion Yuan. At the same time, China has become the world's largest semiconductor market. Despite rapid development and a broad market, the IC industry is generally weak, the industry scale is small, core products are heavily dependent on imports, the scale of trade deficit is expanding year by year, and an independent and controllable industrial chain has not yet been formed. A "lack of core and soul" seriously threatens the industry's safety and restricts its development[5]. Until now, research on the IC industry has mostly focused on industrial chain innovation[6,7], development contexts and patterns[5,8], industrial collaboration[9], and other fields, while there has been less research on the spatial distribution of IC enterprises[10].

The Pearl River Delta (PRD) region is one of the agglomeration areas of China's IC industry[11], and the frontier on which China's mainland is reforming and opening up[5]. A spatial-temporal study of the distribution of IC enterprises in the PRD region could help deepen our understanding of the development of China's IC industry and provide lessons for the development of IC industries in other regions[12]. For this reason, this dataset provides spatial distribution data on IC enterprises in the PRD region. These data include the establishment dates of IC enterprises, their home counties and cities, the industrial division of labor, and other information, that can provide basic data support for related industries and contribute to the development of China's IC industry.

2 Metadata of the Dataset

The metadata of the Integrated circuit enterprises dataset using Qichacha and Baidu Map platforms and data mining technology: taking the Pearl River Delta region of China (1980–2020) as an example[11] is summarized in Table 1.

3 Methods

The basic data for this study came from the Qichacha website[1]. From this website, 14,195 enterprises in nine cities in the PRD region were identified using the keywords "integrated circuit" and "chip", and the name, registration location, and establishment date, of each enterprise were recorded along with other relevant information. As the enterprises' geographic coordinates were not provided by the Qichacha website, they were obtained using the Baidu Map open platform[2] based on the enterprises' geographic locations. The division of labor of

---

[1] https://www.qcc.com/.

[2] https://lbsyun.baidu.com/.

Table 1   Metadata summary of the dataset

Items	Description
Dataset full name	Integrated circuit enterprises dataset using Qichacha and Baidu Map platforms and data mining technology: taking the Pearl River Delta region of China (1980–2020) as an example
Dataset short name	PRD_IC_1980-2020
Authors	Wang, P. H., Institute of Geographic Sciences and Natural Resources Research, CSC, College of Cultural Tourism and Public Administration, Fujian Normal University, wph1996@126.com
	Zha, R. B., College of Cultural Tourism and Public Administration, Fujian Normal University, rbzha@fjnu.edu.cn
	Huang, Y., College of Cultural Tourism and Public Administration, Fujian Normal University, hyueyuey@163.com
	Du, S. Y., College of Cultural Tourism and Public Administration, Fujian Normal University, huoying_ll@126.com
	Xu, J. R., College of Cultural Tourism and Public Administration, Fujian Normal University, xjr9018@163.com
Geographical region	the Pearl River Delta region of China
Year	1980–2020
Data format	.shp
Data size	45.8 MB (710 KB after compressed)
Foundations	National Natural Science Foundation of China (42201267); Ministry of Education of P. R. China (18YJCZH007)
Data publisher	Global Change Research Data Publishing & Repository, http://www.geodoi.ac.cn
Address	No. 11A, Datun Road, Chaoyang District, Beijing 100101, China
Data sharing policy	*Data* from the Global Change Research Data Publishing & Repository includes metadata, datasets (in the *Digital Journal of Global Change Data Repository*), and publications (in the *Journal of Global Change Data & Discovery*). *Data* sharing policy includes: (1) *Data* are openly available and can be free downloaded via the Internet; (2) End users are encouraged to use *Data* subject to citation; (3) Users, who are by definition also value-added service providers, are welcome to redistribute *Data* subject to written permission from the GCdataPR Editorial Office and the issuance of a *Data* redistribution license; and (4) If *Data* are used to compile new datasets, the 'ten percent principal' should be followed such that *Data* records utilized should not surpass 10% of the new dataset contents, while sources should be noted in suitable places in the new dataset[7]
Communication and searchable system	DOI, CSTR, Crossref, DCI, CSCD, CNKI, SciEngine, WDS/ISC, GEOSS

the enterprises in this industrial chain was determined by manually consulting their websites or inquiring over the telephone.

4 Data Result

4.1 Data Composition

Our dataset contains a directory of IC enterprises in nine cities in nine cities in the PRD. Information on 14,195 enterprises was collected from the Qichacha website. The dataset was composed of spatial point data stored in a .shp format data file. The attribute fields recorded are shown in Table 2.

Table 2   Attribute fields in the dataset

Number	Field name	Field content example
1	Name	Shenzhen Hangshun Chip Technology Development Co., Ltd., Guangdong Leadyo IC Testing Co., Ltd., Storart Technology (Shenzhen) Co., Ltd.…
2	Longitude	113.278,241,4, 114.277,798,2, 113.925,025,9…
3	Latitude	22.839,827,95、22.722,734,04、22.495,764…
4	Status	Existing, Revocation, Cancelled…
5	Date_of_estibilish	2013/8/15、2017/5/12、2013/10/14…
6	City	Shenzhen, Guangzhou, Huizhou……
7	District	Baoan, Huangpu, Xiangzhou……
8	Enterprise	Company with limited liability, company limited by shares, and others
9	Industrial division	Fabless, Downstream enterprise, Foundry……

4.2 Data Result

4.2.1 Spatial Distribution

At the county and district levels, as shown in Figure 1, the regions with the most IC enterprises are Shenzhen, Dongguan, Huicheng district of Huizhou, Huangpu district of Guangzhou, Shunde district of Foshan, Zhongshan and Xiangzhou district of Zhuhai. The number of enterprises in other regions is small. In general, Shenzhen is the core, and enterprises in the Pearl River estuary are concentrated in an inverted U shape.
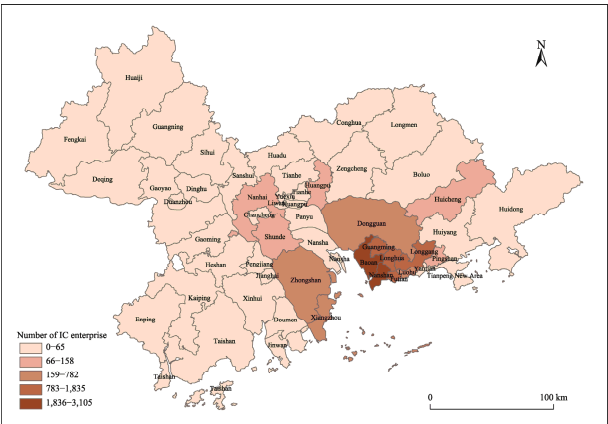

Figure 1   Spatial distribution of IC enterprises in the PRD

ArcGIS software was used to calculate the degree of spatial autocorrelation of IC enterprises in the PRD. Moran's I index was 0.793 and the z value was 10.100, indicating that the spatial distribution of IC enterprises in the PRD is highly agglomerated. Figure 2 shows the kernel density distribution of these enterprises. In Figure 1, we can see that the kernel density in Shenzhen is significantly higher than in other regions. In addition, there is a certain degree of concentration in central Dongguan, Tianhe district, and Huangpu district of Guangzhou, Xiangzhou district of Zhuhai, central Zhongshan, and Huicheng district of Huizhou. The northern industrial park of Baoan district has the highest value of kernel density.
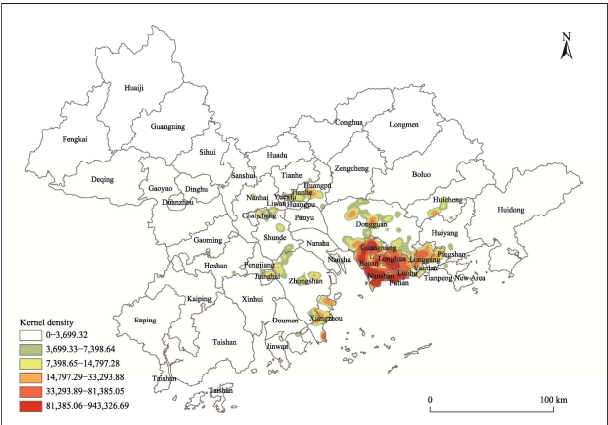

Figure 2   Kernel density distribution of IC enterprises in the PRD

4.2.2 Temporal Evolution

To explore the establishment and development of IC enterprises in the PRD region, enterprises were divided into eight time periods according to the date of establishment, as shown

in Figure 3: 1980–1984, 1985–1999, 1990–1994, 1995–1999, 2000–2004, 2005–2009, 2010–2014, and 2015–2020.
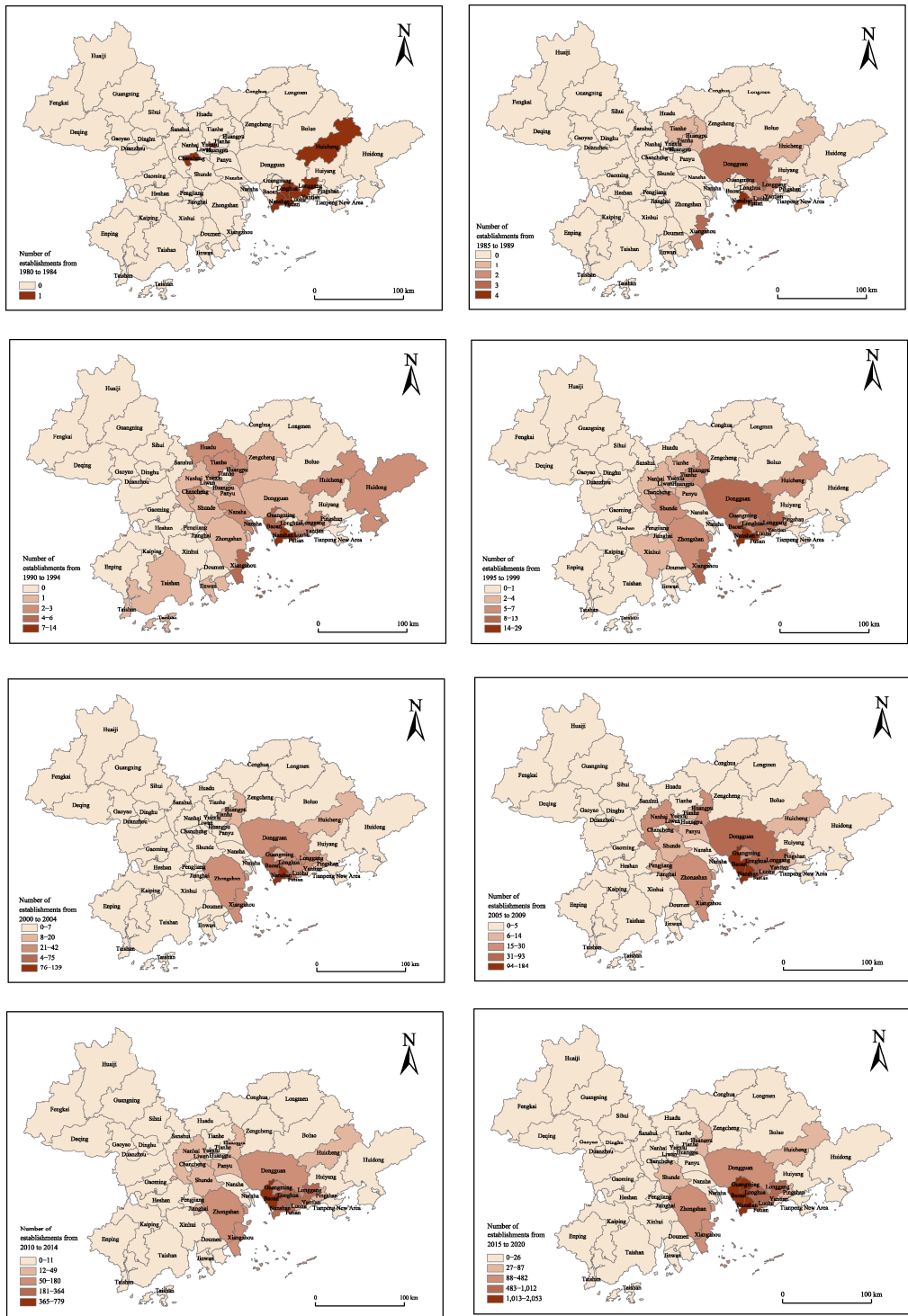
Figure 3 Maps of IC enterprises established in each county and city in the PRD every five (or six) years from 1980 to 2020

From 1980 to 1984, IC enterprises appeared sporadically. In Shenzhen, Guangzhou, Foshan, and Huizhou, these enterprises began to appear, but the overall number was small. Their quantity did not vary significantly across counties and districts.

From 1985 to 1989, newly established enterprises were mainly concentrated in the east of the Pearl River Estuary. Xiangzhou district of Zhuhai, Huangpu district of Guangzhou, and Conghua also established IC enterprises, and the overall number increased. Nanshan district of Shenzhen had the largest increase.

From 1990 to 1994, the number and scope of enterprises increased. Jiangmen and Zhongshan also established them, and Shenzhen gradually opened a quantitative gap with other cities.

From 1995 to 1999, enterprises were gradually concentrated along the Pearl River Estuary. Shenzhen and Dongguan gradually opened a gap with other cities in terms of the number of establishments.

From 2000 to 2004, Shenzhen's growth rate continued to accelerate. The number of enterprises in Nanshan district increased by more than 100, followed by Baoan district, Futian district, and Longgang district.

From 2005 to 2020, the number of IC enterprises increased rapidly, and an inverted U-shaped distribution pattern gradually formed along the Pearl River estuary, with the largest increase in Shenzhen.

In general, since the 1980s, selected cities in the PRD have taken the lead in establishing IC enterprises, which were initially concentrated on the east bank of the Pearl River estuary. In the 1990s, the number of enterprises and cities covered gradually increased. Finally, in the 21st century, their number along the Pearl River Estuary increased rapidly and gradually formed an inverted U-shaped distribution pattern along the Pearl River Estuary. The number of companies established in Shenzhen is far greater than that in other cities.

4.2.3 Industrial Division of Enterprises

With the development of the IC industry, a mature division of subindustries has emerged, including upstream support, core design, manufacturing, seal testing, and downstream demand enterprises. At the same time, some enterprises participate in multiple divisions. To explore the division of labor among IC enterprises in the PRD region, their positions in the industrial chain were divided into eight categories based on telephone inquiries and the self-descriptions on their websites: supporting, fabless, foundry, sealing testing, downstream; fabless and foundry; foundry and sealing testing; and fabless, foundry, and sealing testing enterprises (Table 3.)

Table 3   Industrial division of IC enterprises

Industrial division	Supporting enterprises	Core						Downstream enterprises
		Fabless	Foundry	Sealing testing	Fabless and foundry	Foundry and sealing testing	Fabless, foundry, and sealing testing	
Number of enterprises	23	257	41	68	22	3	12	13766
Percentage/%	0.16%	1.81%	0.29%	0.48%	0.15%	0.02%	0.08%	96.98%

More than 97% of the IC enterprises in the PRD are downstream demand enterprise, totaling 13,766. Most uses finished ICs to manufacture products, mainly in home appliance manufacturing, LED lighting, smart cards, etc. Upstream support enterprises are very rare, accounting for only 0.16%. There are 403 core IC enterprises in total, accounting for 2.84%, of which design enterprises account for 1.81% at most, seal testing enterprises account for 0.48%, and manufacturing enterprises account for 0.29%.

5 Discussion and Conclusion

In the face of an increasingly complex international situation, the spatial-temporal evolution of the PRD region holds great significance as a key cluster of China's IC industry. The dataset created for this study provides a directory of 14,195 IC enterprises established in the PRD from 1980 to 2020 and includes detailed information on their attributes, such as the name, business status, location, industrial division, establishment date, and registered capital of each enterprise. In terms of spatial distribution, it was found that the IC enterprises are concentrated in an inverted U shape along the Pearl River estuary, with Shenzhen as the core. Since the 1980s, IC enterprises first appeared in selected cities and then gradually spread to all cities. In the 21st century, especially the 2010s, the number of IC enterprises increased rapidly, and the number of enterprises established in Shenzhen was far higher than that in other cities. In terms of industrial divisions, the number of downstream demand enterprises was highest, while that of upstream support enterprises was lowest. Among the core IC enterprises, the number of fable enterprises was highest, and that of foundry enterprises was lower. The enterprise information provided by this dataset has great potential for data information mining and is expected to provide a reference and data repository for research on the IC industry in the PRD region.

Author Contributions
Wang, P. H. designed the algorithms of dataset and wrote the data paper. Huang, Y., Du, S. Y., and Xu, J. R. contributed to the data processing and analysis. Zha, R. B. and Huang, J. C. modified datasets and papers.

Conflicts of Interest
The authors declare no conflicts of interest.

References

[1]   Zhu, Y. W. Review of IC Industry in the Past 50 Years [M]. Beijing: Publishing house of electronics industry, 2016.
[2]   Liu, W., Ma, X. H., Liu, W. Development trend of IC industry in the mainland China and some recommendations [J]. *China Soft Science*, 2015(11): 186–192.
[3]   Wang, Y. Y. Solidly promoting the integrated circuit (IC) industry of China: mastering law and adhering to innovation [J]. *Science ＆ Technology Review*, 2021, 39(3): 31–51.
[4]   Qu, Y. Y., Li, X. J. Innovation chain catches up: innovation and development of China's IC industry [J]. *Business and Management Journal*, 2022, 44(9): 5–26.
[5]   Wu, F. F., Han, C. X., Huang, L. C. Analysis of the characteristics of integrated circuit industry R&D cooperation network—based on the perspective of industry Chain [J]. *Science & Technology Progress and Policy*, 2020, 37(8): 77–85.
[6]   Wang, Y. N., Liao, L. How to take the road of breaking the situation of "China Core" [J]. *People's Tribune*, 2022(17): 60–63.
[7]   Sun, Q., Liu, J. J., Hu, B. B. Research on integrated circuit Industry chain and Innovation chain integration development in China [J/OL]. *Studies in Science of Science*, 2022: 1–2 [2022-11-06]. DOI:10.16192/j.cnki. 1003-2053.20221011.002.
[8]   Ren, Y. W., Yang, Y. Spatial distribution and location correlation of the semiconductor industry in the Pearl River Delta region [J]. *Progress in Geography*, 2022, 41(9): 1622–1634.
[9]   Ye, Y. Y., Wu, K. M., Zhang, H. O., *et al.* Geographical agglomeration and location factors of the new-born cross-border manufactual firms in the Pearl River Delta [J]. *Progress in Geography*, 2019, 38(10): 1583–1595.
[10]   He, C. F., Wang, W. Y., Guo, Q. The inheritance, development and transformation of industrial geography in China: Based on analysis of articles published in Acta Geographica Sinica during 1934–2019 [J]. *Acta Geographica Sinica*, 2021, 76(8): 1815–1834.
[11]   Wang, P. H., Zha, R. B., Huang, Y., *et al*. Integrated circuit enterprises dataset using Qichacha and Baidu Map platforms and data mining technology: taking the Pearl River Delta region of China (1980–2020) as an example [J/DB/OL]. *Digital Journal of Global Change Data Repository*, 2022. https:// doi.org/10.3974/geodb.2022.10.07.V1. https://cstr.escience.org.cn/CSTR:20146.11.2022.10.07.V1.
[12]   GCdataPR Editorial Office. GCdataPR data sharing policy [OL]. https://10.3974/dp.policy.2014.05 (Updated 2017).