Supplementary Material

Open-data-based city-scale gridded carbon dioxide emission inventory: supporting urban carbon monitoring in Chengdu, China

Supplementary Material Section 1: Detailed emissions processing and reclassification method

The tabular CO_2 emission inventory was sourced from the "China City Carbon Dioxide Emission Dataset (CCCED) 2020[1]", which was provided by the China City Greenhouse Gas Working Group (CCG). This dataset integrated three sources: 1) CHRED 3.0 database; 2) various official data at the city scale, including statistical yearbooks, government documents, and survey reports; and 3) data obtained from CCG's field research, interviews, telephone inquiries, and official correspondence with relevant departments. The dataset offered comprehensive CO_2 emission data across different sectors for Chinese cities.

This study extracted the emission dataset of Chengdu from this dataset, as presented in Supplementary Table 1, including six major sectors: agriculture, service, industrial energy, industrial processing, residents, and transportation. However, during the compilation of the dataset, we found that the emissions from industrial and industrial processing were only reported at the category level, without accounting for secondary classifications. Significant variations existed in emissions between different sub-sectors, so it was imperative to undertake a more detailed sub-sectoral analysis.

Supplementary Table 1. CO_2 emissions by various sectors in Chengdu from CCCED 2020

CO_2 emission source		**Emissions**
Category	**Type**	**(10 kt)**
Agriculture		26
Service		222
Industrial Energy		1182
Industrial Processing		822
Residents	City Residents	413
	Rural Residents	66
	Total	478
Transportation	Road Transportation	1109
	Railway Transportation	4
	Waterborne Transportation	0
	Air Transportation	435
	Total	1547
Total		4278

Supplementary Table 2. Volume of Main Energy Consumption by Industrial Sector (2020)[2]

	Coal (t)	Gas (10^4 m^3)	Gasoline (t)	Diesel (t)	Fuel Oil (t)	Electricity (10^4 kWh)
Oil and Gas Extraction	-	178359	3562	2211	-	66157
Non-metallic Mineral Extraction	-	-	-	446	-	378
Mining and Dressing	-	3921	1782	204272	-	32256
Agricultural and Sideline Products	22208	6644	223	953	-	42498
Food Manufacturing	-	7752	185	465	-	37901
Beverage Manufacturing	-	5263	144	802	-	30650
Tobacco Products	-	1309	-	2	-	9276
Textile Manufacturing	-	189	54	3	-	7454
Garment, Shoes, and Accessories	-	175	136	36	-	2279
Leather, Fur, and Feather Products	-	339	102	30	-	4632
Wood Processing	-	571	58	604	-	36912
Furniture Products	-	1003	509	878	-	36459
Paper Products	-	7202	96	816	-	41430
Printing and Recording Media	-	3510	512	227	8	33916
Cultural, Educational, and Sports Goods	-	57	38	13	-	942
Petroleum Refining and Coking	-	55125	3	-	100741	190384
Chemical Raw Materials and Products	-	60330	762	1070	-	147929
Pharmaceutical Manufacturing	-	10172	433	262	-	71754
Chemical Fiber Manufacturing	-	3805	2	6	-	13979
Rubber and Plastic Products	15316	1859	305	738	-	102358
Non-metallic Mineral Products	1529881	49059	1938	25784	941	283433
Ferrous Metal Smelting and Rolling	8679	5648	84	279	-	200924
Non-ferrous Metal Smelting and Rolling	-	2614	157	657	-	21817
Metal Products	-	5924	1326	1454	-	76735
General Equipment Manufacturing	30	2632	864	561	-	57025
Special Equipment Manufacturing	-	1101	634	1355	-	27879
Automobile Manufacturing	-	5353	1963	2746	-	117286
Railways, Ships, Aerospace, and Others	-	470	309	237	-	21574
Electrical Machinery and Equipment	-	1190	709	267	-	128291
Computer, Communication, and Electronics	-	5212	246	59	-	480758
Instruments and Meters	-	47	143	20	-	10090
Other Manufacturing	-	33	-	-	-	537
Waste Resources Utilization	-	29	40	196	323	2791
Metal Products, Machinery, and Equipment	-	-	7	-	-	954
Electricity, Heat Production, and Supply	2102815	6190	47	321	-	1735627
Gas Production, and Supply	-	207	322	25	-	3286
Water Production and Supply	-	309	126	230	-	54574

To reclassify industrial energy and industrial processing into secondary types, we utilized “Volume of Main Energy Consumption by Industrial Sector (2020)”[2] (Supplementary Table 2) from the “Chengdu Statistical Yearbook – 2021”. Employing the accounting method specified in “Provincial Greenhouse Gas Inventory Compilation Guidebook”[3], we estimated the CO_2 emissions for various energy consumption and sub-sectors of industrial energy by the following equation:

$$E = \sum\sum\sum\left(EF_{i,j,k} \times AD_{i,j,k}\right)$$

where EF represents emission factor (kg/TJ); AD represents activity data, which refers to fuel consumption (TJ); i represents fuel types; j represents sector activities; k represents technology types.

These emissions were statistically analyzed and reclassified according to the activity data categories (Supplementary Table 3) outlined in "Provincial Greenhouse Gas Inventory Compilation Guidebook". Following this, the emissions for sub-sectors were obtained through calculation and reclassification. The emissions were proportionally reassigned according to the calculated emission of sub-sectors, resulting in the emissions for sub-sectors in industrial energy and industrial energy.

Supplementary Table 3. Correspondence between activity data category and national industrial classification

Activity Data Category	National Industrial Classification
Power, Heat Production	Electricity, Heat Production, and Supply
Oil, Gas Extraction and Processing	Oil and Gas Extraction
Solid Fuels, Other Energy	Petroleum Refining and Coking
	Gas Production, and Supply
Steel Industry	Ferrous Metal Smelting and Rolling
Non-ferrous Metals Industry	Non-ferrous Metal Smelting and Rolling
Chemical Industry	Chemical Raw Materials and Products
	Pharmaceutical Manufacturing
	Chemical Fiber Manufacturing
	Rubber and Plastic Products
Building Materials Industry	Non-metallic Mineral Products
Other Industries	Sub-sectors outside the national industrial classification

Using the aforementioned method, the emissions for each sub-sector of industrial energy were estimated. Subsequently, emissions were proportionally assigned from total emissions reported in CCCED dataset. This approach enabled the reclassification of sub-sectors, as shown in Supplementary Table 4. The total industrial energy emissions estimated from the statistical yearbook data closely approximated the emissions provided by CCCED, indicating the feasibility of this reclassification approach.

Supplementary Table 4. Secondary classification results for industrial energy and industrial processing

CO_2 emission source		Estimated emissions	CCCED dataset
Category	Type	(10 kt)	(10 kt)
Industrial Energy	Power, Heat Production	413	392
	Oil, Gas Extraction and Processing	34	32
	Solid Fuels, Other Energy	1	1
	Steel Industry	14	13
	Non-ferrous Metals Industry	6	6
	Chemical Industry	169	160
	Building Materials Industry	406	385
	Other Industries	205	194
	Total	1247	1182
Industrial Processing	Cement, Limestone Production Processing	819	-
	Steel Production Processing	3	-

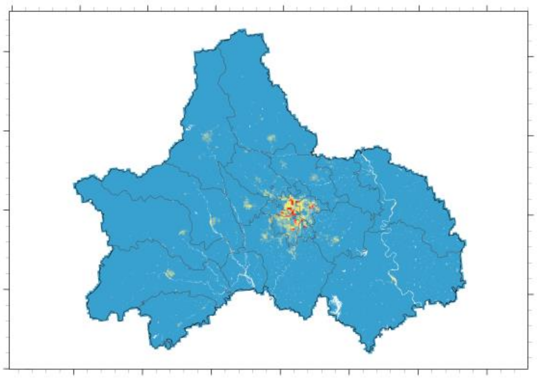
Supplementary Material Section 2: Detailed retrieval process for point sources proxy data

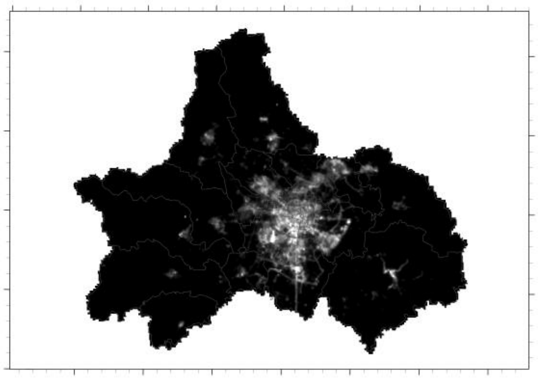
Industrial point source emissions represented a significant sector of overall emissions, yet obtaining accurate location and scale information for these sources remained a challenging task. This study adopted a two-step approach to address this issue: first, this study utilized an online map API to batch-collect location data for factories and enterprises; second, this study retrieved corresponding enterprise-scale information based on the compiled list of names, thereby enabling the precise identification and quantification of emission sources. The retrieval process was as follows:

TianYanCha[4], an enterprise information query platform, was selected as the platform for obtaining registration information of POIs using the batch query function. The query results provided four required attributes: name, registered capital, industry, and business scope. The registered capital was used as an indicator of enterprise scale, serving as proxy data for the spatial allocation of emissions. The industry and business scope helped identify the emissions type. However, the batch query function was only suitable for precise queries, meaning the input enterprise name had to exactly match the name or former name in TianYanCha's database to output results. Since the POIs data contained many abbreviated or approximate names, the batch query results were not perfect, necessitating a secondary screening of the remaining POIs data. To address this issue, a network data acquisition was employed for secondary screening. The acquisition was an automated script for obtaining webpage content, acquiring the required data through repeated browsing. Given the relatively small data demand and to avoid interception by anti-scraping technologies, the Selenium module was applied as the acquisition. Selenium is a web-based technology that directly interacts with browsers, simulating user operations such as entering URLs, scrolling, and clicking, which makes it highly capable of bypassing website anti-scraping mechanisms[5]. In addition to the automated script, several auxiliary modules were required to complete the acquisition, including Xpath library for webpage navigation and positioning, the Beautiful Soup (BS4) library for information collection, and JSON library for data object encoding and decoding, thus achieving the full workflow of data acquisition. Finally, the POIs that could not be identified were removed, and the database was formed by combining the batch query POIs. Through manual screening of industry and business scopes, enterprises and factories related to CO_2 emissions were selected, resulting in a total of 5,335 enterprise information. This database included relevant location and scale information on the emitting enterprises as proxy data for industrial point source emissions.

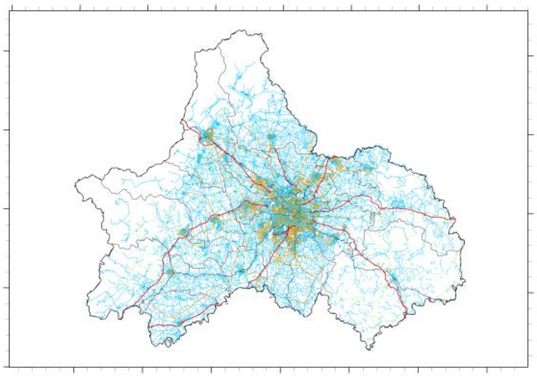
Supplementary Material Section 3: Detailed proxy data for non-point source

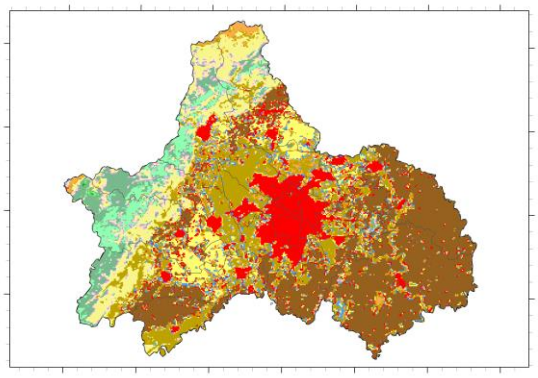
This section presented the detailed proxy data, as shown in Supplementary Figure 1.





A B





C D

Supplementary Figure 1. Detailed proxy data for non-point source: A. WorldPop population data; B. Nighttime light data; C. Road network; D. Land cover.

Supplementary Material Section 4: Preprocessing methods for non-point source emission proxy data beyond population distribution data for residents' emissions

This section described the preprocessing methods for proxy data, excluding population distribution data. The focus was on traffic volume distribution data for transportation emissions and cropland distribution data for agriculture emissions.

Transportation emissions were spatial allocated using traffic volume distribution data. Prior to spatial allocation, road network density for each grid was computed based on road levels. The calculation equation was as follows:

$$\rho_{ij} = \frac{\sum L_{ij}}{S_{ij}}$$

where ρ_{ij} represents the road network density of each grid (m/m^2); L_{ij} represents the road length within each grid; S_{ij} represents the area of each grid.

In this study, roads were categorized into four levels: expressways, arterial roads, collector roads, and branch roads. The road network density for each level was computed. Given the substantial differences in traffic volumes among all road levels, the spatial allocation must consider the influence of traffic volumes on the road network density of each level. According to "Code for Design of Urban Road Engineering"[6] and "Technical Standard of Highway Engineering"[7], traffic volumes were converted to average daily traffic volumes: 55,000 vehicles for expressways, 25,000 vehicles for arterial roads, 10,000 vehicles for collector roads, and 2,000 vehicles for branch roads. Traffic volumes were overlaid on road network density, and the traffic volume distribution for each road level was aggregated to obtain the traffic volume distribution in Chengdu. Finally, spatial allocation of transportation emissions was conducted based on the traffic volume distribution.

Agriculture emissions mainly include CO_2 emissions generated from agricultural machines. In this study, croplands and cropland/natural vegetation mosaics were extracted as proxy data based on the land cover data in Supplementary Figure 1D. However, due to the limited availability of detailed agricultural machine usage data, agricultural emissions were uniformly spatially allocated to target grids to develop the gridded agriculture emission inventory.

REFERENCES

1. Group CCGGW. China City Carbon Dioxide Emission Dataset (2020). Available from: http://www.cityghg.com/toCauses?id=4. [Last accessed on 28 August 2024]
2. Bureau CMS. Chengdu Statistical Yearbook - 2021. Beijing:China Statistics Press;2021.
3. Commission NDaR. Provincial Greenhouse Gas Inventory Compilation Guidebook. In; 2011.
4. Tianyancha. Tianyancha. Available from: https://www.tianyancha.com/. [Last accessed on 21 June 2024]
5. García B, Munoz-Organero M, Alario-Hoyos C, Kloos CD. Automated driver management for Selenium WebDriver. *Empirical Software Engineering* 2021;26:107. (https://doi.org/10.1007/s10664-021-09975-3)
6. China MoHaU-RDotPsRo. Code for Design of Urban Road Engineering.Beijing:China Architecture & Building Press;2012.p.8-9. (in Chinese)
7. China MoTotPsRo. Technical Standard of Highway Engineering.Beijing:China Communication Press;2003.p.1-9. (in Chinese)