
China's Provincial Multi-Regional Input-Output Database for 2018 and 2020

Received: 21 August 2025

Accepted: 24 December 2025

Cite this article as: Li, J., Zhang, Z., Liang, D. *et al.* China's Provincial Multi-Regional Input-Output Database for 2018 and 2020. *Sci Data* (2025). <https://doi.org/10.1038/s41597-025-06543-y>

Jie Li, Zeyi Zhang, Diling Liang, Qianhong Ouyang, Peipei Tian, Dabo Guan & Heran Zheng

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

China's Provincial Multi-Regional Input-Output Database for 2018 and 2020

Jie Li¹, Zeyi Zhang¹, Diling Liang², Qianhong Ouyang³, Peipei Tian¹, Dabo Guan^{2,4}, Heran Zheng^{2,*}

1. Institute of Blue and Green Development, Shandong University, Weihai, 264209, China

2. The Bartlett School of Sustainable Construction, University College London, London, WC1E 7HB, UK

3. School of Statistics, Beijing Normal University, Beijing, 100875, China

4. Department of Earth System Sciences, Tsinghua University, Beijing 100084, China

Correspondent author:

Heran Zheng (Heran.zheng@ucl.ac.uk)

Abstract

The U.S.-China trade friction in 2018 and the COVID-19 pandemic in 2020 have significantly influenced China's domestic supply chains, with their impacts varying considerably across regions and sectors. Multi-regional input-output (MRIO) models are widely used to track supply chains and analyze cross-regional spillover effects, playing a key role in understanding economic linkages and environmental impacts. However, due to data unavailability, existing MRIO tables fail to capture the impact of the U.S.-China trade friction and the COVID-19 pandemic on China's regional supply chains. To address this data gap, we employ hybrid methods to construct Chinese MRIO tables for 2018 and 2020, covering 31 regions and 42 sectors. This dataset is consistent with our previous work on the China provincial MRIO model for the years 2012, 2015, and 2017, offering insights into how regional supply chains and economic structures adapted to the combined impacts of the trade war and the COVID-19 pandemic.

Key words: China, supply chain, multi-regional input-output

Background & Summary

Since China joined the World Trade Organization (WTO) in 2001, its regional supply chains have deeply integrated into the global value chain^{1,2}. However, this integration is challenged by new disruptions such as geopolitical conflicts and natural disasters³⁻⁵. In particular, the technology blockade and tariff barriers imposed by the U.S.-China trade war since 2018 have resulted in an estimated indirect economic burden of USD 6.5 billion on China⁶, raising both imports and exports costs to the United States^{7,8}, and prompting a reconfiguration of trade networks⁹. The lockdown and declining global demand during the COVID-19 pandemic caused China's manufacturing Purchasing Managers' Index (PMI) fell sharply to a historic low of 35.7 in February 2020, exposing structural vulnerabilities in the supply chains^{10,11}. These shocks reshaped China's supply chain by restricting exports, diverting investments, and disturbing supply and demand dynamics^{3,12,13}.

Amid external shocks, the challenges faced by industries vary widely across provinces and cities due to the heterogeneity of their economic characteristics^{14,15}. In the trade war, areas with a high share of exports and exposure to targeted goods (like Guangdong and Shanghai) were impacted first, causing ripple effects across economic system. Industries with more flexible supply chains were more likely to withstand the pressure. In the 2020 pandemic, Hubei Province suffered the most severe blow from the initial outbreak. After that, other provinces experienced disruptions due to production shutdown, logistic failure and global trade halts. Nevertheless, the pandemic opened opportunities for industries like healthcare, research and digital information services. Therefore, strategies for mitigating risks and driving development must be specific to industries and satisfy each province's conditions. We urgently need a dataset at province level to support more comprehensive research and more targeted policies.

The Multi-Regional Input-Output (MRIO) model is widely used to analyze regional economic flows, supply chain dynamics, and associated environmental impacts¹⁶⁻²¹. Numerous MRIO tables for China have been developed by institutions and individual researchers. For example, Zhang developed China's MRIO tables for the years 1997, 2002, and 2007 using a combination of the maximum entropy model and the doubly constrained gravity model²², covering 8 regions with 17 sectors. Shi and Zhang constructed a 2002 MRIO model for China, encompassing 30 provinces and 60 sectors, employing a hybrid approach based on the Chenery-Moses framework²³. Liu and colleagues improved the gravity model by incorporating spatial weights and competition coefficients to construct MRIO tables for 2007, 2010, and 2012²⁴⁻²⁷. Mi proposed an MRIO table for 2012, encompassing 30 provinces and 30 sectors, using a modified gravity model²⁸. Later, Zhao and colleagues constructed a series of MRIO tables covering 1987 to 2017 with carbon emission accounts²⁹, which were later refined to include more detailed regional and sectoral classification. Zheng applied the maximum entropy approach to construct MRIO tables for 2012, 2015, and 2017, incorporating 31 regions and 42 sectors³⁰. Following the approach, Wang developed the city-level MRIO table by firm ownership from 2002 to 2017³¹. However, no existing MRIO datasets beyond 2017 are available, limiting the capacity to accurately capture the changes in China's provincial supply structures amid the U.S.-China trade friction in 2018 and the COVID-19 pandemic in 2020.

To address this gap, we employed a hybrid approach to construct MRIO tables for 31 provinces and 42 sectors in mainland China for the years 2018 and 2020. This approach has been widely applied in the construction of MRIO tables across various countries³²⁻³⁵, and consistent with our previous work of the provincial MRIO tables for 2012, 2015, and 2017. Our dataset reflects changes in China's regional economic structure influenced by U.S.-China trade friction and the COVID-19 pandemic. It is valuable for economic and environmental analysis, especially in assessing regional growth drivers and inter-regional spillover effects amid U.S.-China trade friction in 2018 and the COVID-19 pandemic in 2020.

Methods

Input data

This study constructs province-level MRIO tables for China for the years 2018 and 2020 through regionalizing national IO tables published by the National Bureau of Statistics of China (https://data.stats.gov.cn/ifnormal.htm?u=/files/html/quickSearch/trcc/2020_1.html&h=740). Our hybrid approach combines official survey data with simulation results^{31,36,37}, balancing cost and data quality^{25,38}. The provincial data used in this process include 2017 and 2020 Single-Region Input-Output (SRIO) tables³⁹. Sectoral economic data were obtained from the statistical yearbooks published by the statistical bureaus of all 31 provinces in China. All yearbooks are publicly accessible through the respective provincial statistical bureau websites (search term: “<Province Name> + Statistical Yearbook + <Year>”). In addition, import-export data were sourced from China Customs (<http://stats.customs.gov.cn/>). All data are constrained by the national IO tables to ensure consistency. However, due to the high cost of conducting surveys, China only releases national IO tables every 2-3 years or during the economic censuses, while provincial tables published even less frequently. After 2017, only Beijing, Shanghai, and Tianjin have published their tables in 2020, which largely delay the construction of MRIO table. Meanwhile, sectoral data for interregional trade flows are also unavailable. To address the data limitations, we employed a cross-entropy model^{40,41} to generate the missing SRIO tables for each province. For inter-provincial trade, we applied a gravity model^{42,43} based on 2017 railway transport and electricity transmission data⁴⁴ to estimate the trade matrix.

Form MRIO table

As shown in Figure 1, the construction workflow of the MRIO table involves five stages:(1) Estimating domestic supply and demand;(2) Disaggregating domestic supply and demand;(3) Compiling provincial SRIO accounts;(4) Estimating intra-regional transaction matrices;(5) Estimating inter-regional trade matrices.

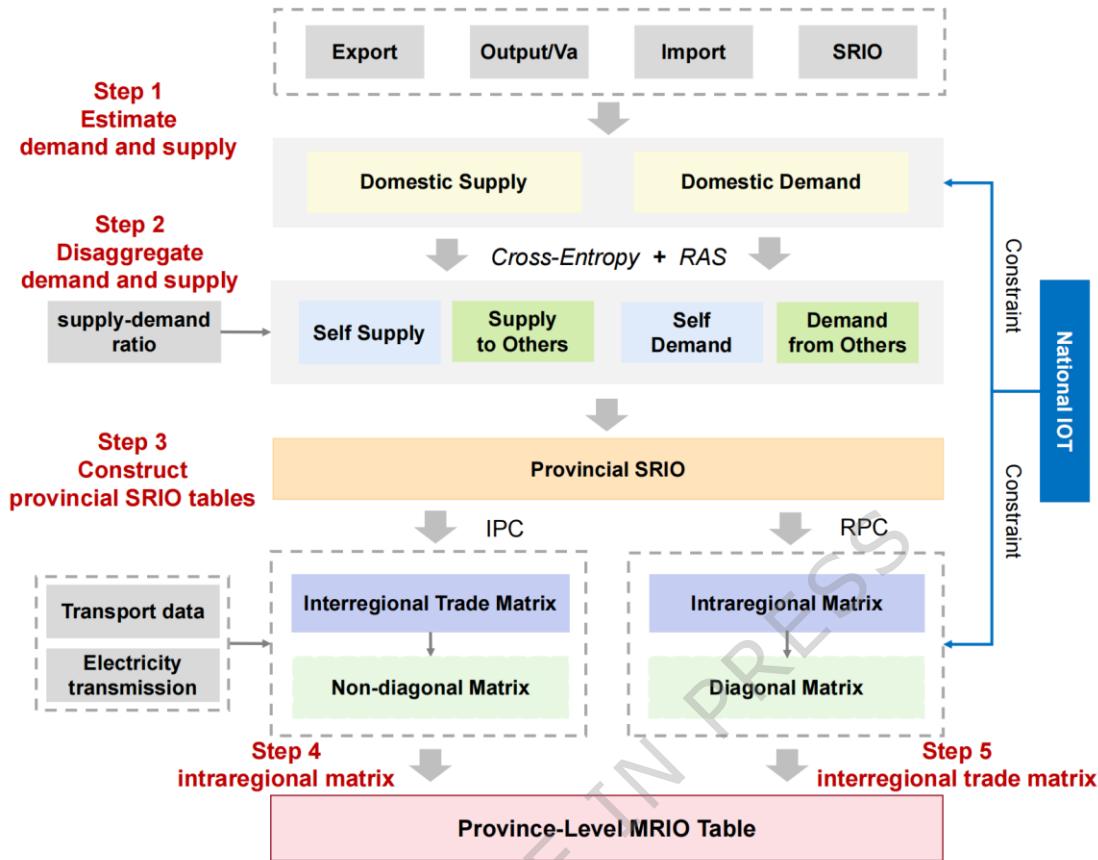


Figure 1. MRIO Table Construction Workflow

The specific construction steps of the MRIO table are as follows:

(1) Estimating domestic supply and demand

First, based on the supply-demand balance principle, we estimate the supply and demand for each region and sector. Domestic supply and domestic demand include both intermediate goods and final products. From the supply perspective, the domestic supply of each region and sector should equal its total output minus foreign exports, as shown in Eq. 1:

$$supply_r^i = output_r^i - export_r^i \quad (1)$$

The output of sector i in province r is represented by $output_r^i$, while $export_r^i$ stands for its exports. $supply_r^i$ reflects the domestic allocation of sector i 's output in province r , indicating the share that remains for domestic supply. Sectoral output data for some provinces in 2018 and 2020 is available from provincial statistical yearbooks or SRIO tables. However, some provincial yearbooks do not provide data on tertiary industry output

or industrial sector value added. In these cases, we estimate the missing data by assuming identical share structures for value added and output. It means that the sectorial distribution of value added is same as the distribution of output. The assumption is used for scenarios either no sectoral value added or no sectoral output.

From the demand perspective, the domestic demand of each region and sector should equal its total demand (intermediate demand plus final demand) minus foreign imports. This data can be obtained from the provincial SRIQ table, as shown in Eq. 2:

$$demand_r^i = z_r^i + f_r^i - import_r^i \quad (2)$$

Here, z_r^i denotes the intermediate demand for sector i in province r , while f_r^i represents its final demand. $import_r^i$ indicates the import volume of sector i in the same province. The domestic demand of sector i in province r is defined as the total demand excluding imports ($demand_r^i$). However, since SRIQ tables are unavailable for certain provinces, the total demand cannot be directly determined. Therefore, we make the following assumptions: (1) The technical coefficients in 2018 are the same as in 2017, and the ratio of intermediate demand to total demand remains unchanged. Thus, we use the 2017 provincial SRIQ table structure to estimate total demand for 2018. (2) For regions where the 2020 provincial SRIQ tables are unavailable, we estimate total demand using the technical coefficients and the ratio of intermediate demand to total demand from the national IO table. The specific calculation formula is as follows:

$$\begin{cases} td_r^{i*} = \frac{\sum_i a_{2017_r^i} \times output_{2018_r^i}}{\sum_i z_{2017_r^i} / td_{2017_r^i}} \\ \hat{td}_r^i = \frac{td_r^{i*}}{\sum_r td_r^{i*}} \times nd_{2018^i} \\ demand_r^i = \hat{td}_r^i - import_r^i \end{cases} \quad (3)$$

The variable td_r^{i*} , marked with a “*”, provides a preliminary estimation of total demand for sector i in province r . $a_{2017_r^i}$ denotes the 2017 technical input coefficient for sector i as required by province r . The total output of sector i in province r is captured by $output_{2018_r^i}$. Variables $z_{2017_r^i}$ and $td_{2017_r^i}$ refer to the intermediate and total demand for sector i in province r in 2017. $demand_r^i$ indicates the required input of sector i in province r , while nd_{2018^i} denotes the national demand for the same sector in 2018. Finally, $import_r^i$ corresponds to the imported quantity of sector i in province r . To ensure data consistency, the total output, total demand, value added, import, and export data for each province and sector are constrained using the national IO table.

(2) Detailed decomposition of domestic supply and demand

From the supply perspective, domestic supply (DS) consists of self-supply (SS) and supply to other provinces (SO). Similarly, domestic demand (DD) can be further divided into self-demand (SD) and demand from other provinces (DO), as shown in the Figure 2.



Figure 2. Provincial flow of supply and demand

The Cross-Entropy (CE) model can effectively adjust unknown distributions to minimize information loss while maintaining consistency with known totals⁴⁵. Therefore, we use the CE model to disaggregate domestic supply and demand based on the following relationships: (1) Self-supply equals self-demand within the same province. (2) The domestic supply within a province can be further decomposed into two components: supply retained for local use (SS) and supply distributed to other provinces (SO). (3) Domestic demand within a province is the aggregate of local demand (SD) and demand from external provinces (DO). (4) The total SO from all provinces must equal the total DO from all provinces. (5) Local supply is prioritized to meet local demand, meaning that each province's self-demand must be greater than 0. Mathematically, this can be described as:

$$\min C(P||Q) = \sum_i \sum_r P_{ir} \cdot \ln \left(\frac{P_{ir}}{q_{ir}} \right) \quad (4)$$

Subject to:

$$\sum_i \sum_j (P_{ir}^{SS} + P_{ir}^{SO}) = 1;$$

$$\sum_i \sum_r (P_{ir}^{SD} + P_{ir}^{DO}) = 1;$$

$$\sum_r p_i^{SO} \times s_i = \sum_r P_i^{DO} \times d_i;$$

$$(P_{ir}^{SS} + P_{ir}^{SO}) \times s_i = S_{ir};$$

$$(P_{ir}^{SD} + P_{ir}^{DO}) \times d_i = d_{ir};$$

$$p_{ir}^{SD} > 0$$

P_{ir} represents the allocation of supply (SS+SO) and demand (SD+DO) for sector i in province r , while q_{ir} denotes the initial distribution for the same sector and province. The variables s_i and d_i correspond to the total domestic supply and demand aggregated across all provinces. Meanwhile, S_{ir} and d_{ir} specify the domestic

supply and domestic demand for sector i in province r .

For the 42 sectors, the initial shares of each province's SS, SO, SD, and DO are derived from the provincial SRIO tables. The detailed calculation steps are as follows:

$$q_{ir}^{SS} = SS_{ir}/(SS_{ir} + SO_{ir}); \quad (5)$$

$$q_{ir}^{SO} = SO_{ir}/(SS_{ir} + SO_{ir}); \quad (6)$$

$$q_{ir}^{SD} = SD_{ir}/(SD_{ir} + DO_{ir}); \quad (7)$$

$$q_{ir}^{DO} = DO_{ir}/(SD_{ir} + DO_{ir}); \quad (8)$$

q_{ir}^{SS} , q_{ir}^{SO} , q_{ir}^{SD} and q_{ir}^{DO} represent the proportions of SS, SO, SD, and DO for sector i in province r , respectively. Specifically, SS_{ir} denotes the self-supply of sector i in region r , SO_{ir} represents the outflow, SD_{ir} refers to the self-demand, and DO_{ir} indicates the inflow. After disaggregating the domestic supply and domestic demand by province, the sum of domestic exports from all provinces matches the total domestic imports for each sector.

(3) Compilation of provincial SRIO accounts

Firstly, we estimate the final demand for each sector in each region using regional GDP, export and import data, external supply (supply to others), and external demand (demand from others). Secondly, the intermediate demand is obtained by multiplying the existing technical coefficient matrix with the output vector. It is important to note that although the China-U.S. trade war in 2018 brought about significant shifts in the global trade landscape, the production technologies and input-output structures inside China are unlikely to change rapidly over a short period. Accordingly, we adopt the 2017 input-output coefficients derived from SRIO tables to approximate the technological structure of each province in 2018. This is a common approach for short-term MRIO construction in data-scarce contexts to provide a reasonably accurate estimation of each province's economic characteristics. For 2020, the technical coefficient matrices of Beijing, Tianjin, and Shanghai are sourced from the SRIO tables published by these provinces, while for other provinces without SRIO tables in 2020, the average technical coefficient matrix from the national IO table is used as a substitute. Then, the optimized sectoral outflows, inflows, exports, and imports for each province from the previous step are integrated with the provincial intermediate use and final demand matrices. Finally, the SRIO tables for each province are developed using a cross-entropy approach, ensuring compliance with two equilibrium conditions: (1) For each sector, the sum of intermediate demand and final demand should equal the total output minus net exports; and (2) intermediate demand is also derived by subtracting value added from total output. These relationships can be formally expressed as:

$$\min C(P \parallel Q) = \sum_i \sum_r |q_r| \cdot p_r \cdot \ln\left(\frac{p_r}{e}\right) \quad (9)$$

Subject to:

$$\sum_j q_r^{ij} \cdot p_r^{ij} = Output_r^j - Valve\ added_r^j;$$

$$\sum_i q_r^{ij} \cdot p_r^{ij} = Output_r^i - Net\ export_r^i;$$

p_r^{ij} denotes the adjusted distribution obtained by updating the known prior through the cross-entropy procedure. e represents the natural logarithm. q_r^{ij} , as the prior distribution, includes the intermediate demand and final demand. $Output_r^j$ refers to the input of sector j in province r , while $Output_r^i$ indicates the output of sector i in province r . $Net\ export_r^i$ captures net exports of product i , calculated as total (foreign + domestic) exports minus total (foreign + domestic) imports. $Valve\ added_r^j$ represents the value added of each sector in each region. Foreign trade data are sourced from provincial SRIO tables or China Customs.

It is important to note that when constructing the 2020 provincial SRIO tables (except for Beijing, Tianjin, and Shanghai), we used national technical coefficients as proxies for the technical coefficients of each province. These coefficients were then multiplied by the actual 2020 output of each region to estimate preliminary intermediate demand, which was subsequently used to calculate total demand ($Z + F$). Technical coefficients reflect each region's production structure and the proportions of intermediate inputs. Provinces display pronounced heterogeneity in industrial structure and supply chain integration. Applying national technical coefficients uniformly across provinces reduces the actual and often significant technological differences among them. However, this influence is mainly confined to the technological dimension and does not eliminate regional heterogeneity. The subsequent RAS adjustment is performed to satisfy the constraints, which incorporate actual provincial data on output, demand, and trade. This helps preserve differences in economic scale and industrial structure across provinces. As a result, the final matrix retain distinct provincial characteristics.

(4) Estimation of intraregional matrix

To develop a competitive MRIO table, it is essential to transform the provincial competitive SRIO tables from Step (3) into their non-competitive counterparts. This transformation requires decomposing each province's intermediate demand and final demand into three components: (1) demand met by local production, (2) demand fulfilled by other provinces, and (3) imports. To facilitate this disaggregation, we employ the Regional Purchase Coefficient (RPC), which estimates the proportion of total demand that is sourced locally. It is assumed that the local share of supply within both intermediate demand and final demand remains constant across each SRIO table. Therefore, local demand components in the MRIO framework are obtained by applying the RPC to the respective intermediate demand and final demand figures from the SRIO tables. The relationship is expressed mathematically as follows:

$$rpc_r^i = \frac{(Output_r^i - export_r^i - so_r^i)}{(Output_r^i - export_r^i - so_r^i + import_r^i + do_r^i)} \quad (10)$$

$$z_{rr}^{ij} = rpc_r^i \times z_r^{ij} \quad (11)$$

$$f = rpc_r^i \times f_r^i \quad (12)$$

rpc_r^i denotes the share of intermediate demand and final demand for sector i in region r fulfilled by local production. $Output_r^i$, $export_r^i$ and $import_r^i$ represent total output, foreign exports, and foreign imports, respectively. so_r^i denotes the outflow of sector i in province r , whereas do_r^i denotes the inflow of sector i in the same province. z_{rr}^{ij} and f represent the amounts of intermediate demand and final demand in province r that are sourced domestically within the same province.

(5) Estimation of interregional trade matrix

We need to estimate the share of total demand met by external regions for each province to derive the off-diagonal entries in the MRIO table for both intermediate demand and final demand.

Firstly, similar to the approach in Step (4), we introduce the Import Purchase Coefficient (IPC) to quantify the portion of each province's intermediate and final demand supplied by other provinces. It is assumed that this interregional import ratio remains consistent across intermediate demand and final demand within each SRIO framework. Formally:

$$ipc_r^i = \frac{do_r^i}{(Output_r^i - export_r^i - so_r^i + import_r^i + do_r^i)} \quad (13)$$

$$zm_r^{ij} = ipc_r^i \times z_r^{ij} \quad (14)$$

$$fm_r^i = ipc_r^i \times f_r^i \quad (15)$$

ipc_r^i denotes the share of intermediate demand and final demand for sector i in region r fulfilled by imported products. $Output_r^i$, $export_r^i$ and $import_r^i$ represent the total output, foreign exports, and foreign imports, respectively. so_r^i denotes the outflow of sector i in province r , whereas do_r^i denotes the inflow of sector i in the same province. zm_r^{ij} and fm_r^i represent the amounts of intermediate demand and final demand in province r that are met by imported products, respectively.

Secondly, due to the lack of detailed interregional trade matrices for all 42 sectors, we employ a gravity model using railway transportation data for 11 categories of goods to estimate interregional trade flows across 24 sectors within the input-output framework. For non-transportable sectors (e.g., services and construction), transport costs are not considered; rather, we assume allocation is proportional to regional supply and demand. For electricity, interregional electricity transmission data from the China Electricity Power Yearbook is used to estimate related trade flows. This method and data are consistent with our previous study, thereby ensuring consistency over the time series. The general structure of the gravity model is defined as:

$$trade_{rs}^i = G^\alpha \frac{(outflow_{ro}^i)^{\beta_1} \times (inflow_{os}^i)^{\beta_2}}{(distance_{rs})^\gamma} \quad (16)$$

$trade_{rs}^i$ represents the volume of interregional trade for product i from province r to s , while $distance_{rs}$ denotes the geographic distance between the two regions, serving as a proxy for transport-related costs. $outflow_{ro}^i$

indicates the amount of product i that province r supplies to other provinces, whereas inflow_{os}^i denotes the volume of product i that province s imports from other provinces. However, the initial trade matrix obtained from the above steps does not satisfy the row and column constraints for domestic exports and imports in the provincial SRIO tables. To reconcile these discrepancies, we utilize the RAS balancing method to adjust the trade matrix, thereby ensuring coherence with the supply and demand constraints of the SRIO tables.

Thirdly, based on the balanced trade matrix, we compute the share of total domestic imports supplied by each province-referred to as the regional purchase proportion (RP). By applying this proportion to each region's total domestic demand, we derive the volume of demand met by other provinces. Mathematically:

$$rp_{rs}^i = \frac{\text{trade}_{rs}^i}{\sum_s \text{trade}_{rs}^i} \quad (17)$$

$$z_{rs}^{ij} = rp_{rs}^i \times zm_s^{ij} \quad r \neq s \quad (18)$$

$$f_{rs}^i = rp_{rs}^i \times fm_s^i \quad r \neq s \quad (19)$$

Here, rp_{rs}^i represents the share of domestic imports of product i flowing from province r to province s , while trade_{rs}^i reflects the volume of sector i 's trade from r to s . zm_s^{ij} and fm_s^i denote the intermediate and final demand for the imported products i to province s . z_{rs}^{ij} and f_{rs}^i represent the intermediate and final demand fulfilled by domestic imports flowing from province r to province s .

Data Records

Multi-regional IO table for 31 provinces											Other components							
Multi-regional Input-Output table			Intermediate Use						Final Demand				Export	Total Output				
			Province 1		...	Province 31		Province 1		...	Province 31							
Sector 1	...	Sector 42	Sector 1	...	Sector 42	F1	...	F5	...	F1	...	F5				
Intermediate input	Province 1	Sector 1	$Z_{1,1}$...	$Z_{1,31}$		$F_{1,1}$...	$F_{1,31}$		EX TO					
									
		Sector 42										
	Province 31								
		Sector 1	$Z_{31,1}$...	$Z_{31,31}$		$F_{31,1}$...	$F_{31,31}$							
									
Import			IM _{inter use}						IM _{final demand}									
Value added	Compensation of employees	V _{1,1}	...	V _{1,32}														
	Net taxes on production	V _{2,1}	...	V _{2,32}														
	Depreciation on the fixed capital	V _{3,1}	...	V _{3,32}														
	Operating surplus	V _{4,1}	...	V _{4,32}														
	Total input	TI																

Figure 3. The layout of China's province-level MRIO table

The dataset is available at the China Emission Accounts and Datasets (CEADs, www.ceads.net) and Figshare⁴⁶ (<https://doi.org/10.6084/m9.figshare.29927291>). The 2018 and 2020 China MRIO tables reflect the economic linkages and industrial structure evolution among regions during the U.S.-China trade war and the COVID-19 pandemic. The structure of our China's MRIO tables for 2018 and 2020 is shown in Figure 3. It consists of the intermediate demand matrix (1302×1302), final demand matrix (1302×155), value-added matrix (5×1302), import vector (1×1457), export vector (1302×1), total output vector (1×1302), and total output vector (1302×1). Economic relationships among 31 provinces and 42 sectors in China are captured by the intermediate demand matrix. The final demand matrix comprises five categories: urban household consumption, rural household consumption, government consumption, fixed capital formation, and changes in inventories. The value-added matrix consists of four components: compensation of employees, net taxes on production, depreciation of fixed capital, and operating surplus. Imports and exports, categorized by province and sector, are captured in separate vectors. The import vector distinguishes between intermediate demand (1×1302) and final demand (1×155).

Technical Validation

MRIO table balancing

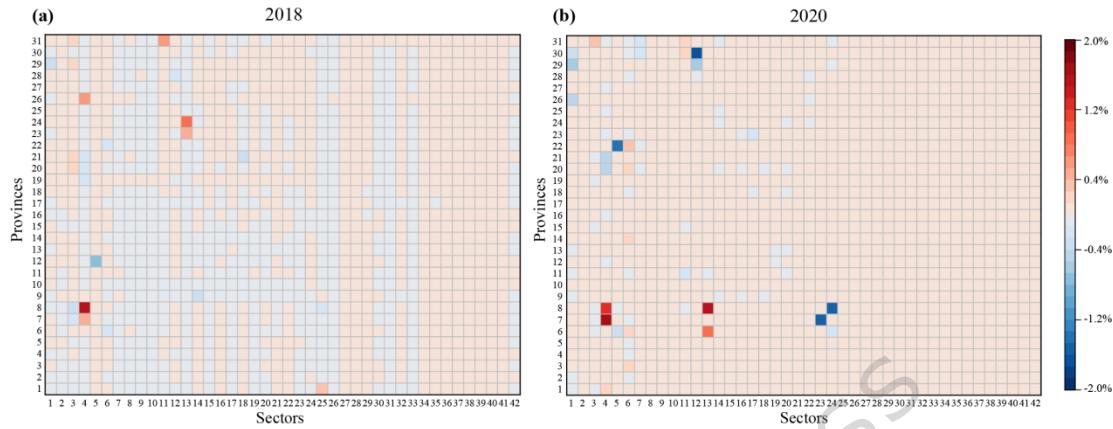


Figure 4. Heatmap of Error Terms Across 31 Provinces and 42 Sectors in 2018 and 2020

This study constructs the "Errors" column vector (1302×1) in the China MRIO table to reflect total output balancing errors. In all provinces, the discrepancy in total output across the 42 sectors stays within a $\pm 5\%$ range. The "Errors" column vector represents the difference between the total output of each sector in each province and the sum of its intermediate demand, final demand, and net exports. Figure 4 presents heatmaps of error terms across 31 provinces and 42 sectors for 2018 and 2020. Deeper red indicates a greater positive deviation of sectoral total output from the actual value, while deeper blue indicates a greater negative deviation. For example, in 2018, the largest output discrepancy was observed in Sector 4 (Mining and Processing of Metal Ores) in Province 8 (Heilongjiang), with a deviation of 1.5%. In 2020, the most significant discrepancies occurred in Sector 4 of Province 7 (Jilin) and Sector 12 (Manufacture of Chemical Products) in Province 30 (Ningxia), with deviations of 1.7% and -1.7%, respectively. These deviations primarily stem from balancing inconsistencies in the provincial input-output tables and mismatches between total inflows and outflows across provinces.

Discrepancies between MRIO and national IO table

Currently, no other institutions have published MRIO tables for China for the years 2018 and 2020. Therefore, to validate the reliability of our MRIO tables for these years, we compare them with the official national SRIIO tables. The results show that, in both 2018 and 2020, the differences in domestic intermediate inputs across most sectors remain within $\pm 10\%$, indicating a high degree of overall consistency between the two datasets. Only a few sectors exhibit relatively large discrepancies (see Figure 5). Specifically, in 2018, the sectors "Extraction of petroleum and natural gas," "Processing of petroleum, coking, and processing of nuclear fuel," and "Production and distribution of gas" show deviations of -24%, +21%, and +25%, respectively. By contrast, in 2020, the largest discrepancy is concentrated in "Mining and processing of metal ores" (11%). These discrepancies mainly

stem from the assumption that intermediate and final demand share the same import ratio when converting the competitive SRIO table into a non-competitive format. For example, in 2018, the discrepancy for the "Extraction of Petroleum and Natural Gas" sector reached -24%, primarily due to an overestimation of intermediate imports in the MRIO table—88% higher than the actual value. Since this sector is highly import-dependent, with approximately 54% of its supply sourced from abroad, the discrepancy is further amplified under the import estimation assumptions described above. Therefore, it is reasonable that sectors with high import reliance are more prone to larger discrepancies during the MRIO construction process. Access to more detailed import data—especially distinguishing between intermediate and final demand—would enhance the consistency and accuracy of the MRIO tables relative to national SRIO benchmarks.

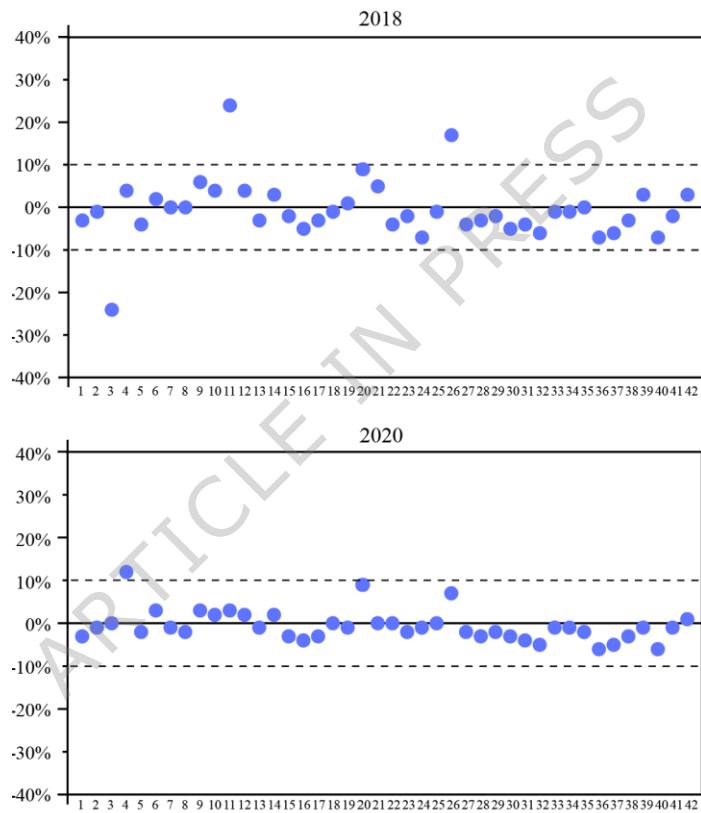


Figure 5. Discrepancies in Domestic Intermediate Inputs Between National IO and MRIO Tables in 2018 and 2020

Regional Value Chain Accounting for 2018 and 2020

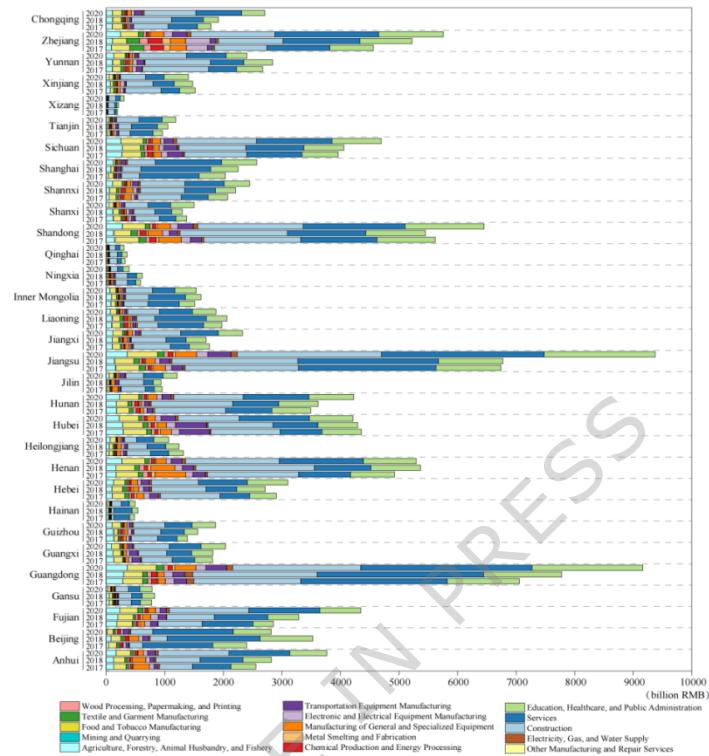


Figure 6. Consumption-Based Value Added across Regions and Sectors in China for 2017, 2018, and 2020

To verify the validity and dynamic stability of the 2018 and 2020 China multi-regional input-output (MRIO) tables, this study compared consumption-based value added (CBVA) across Chinese provinces in 2017, 2018, and 2020 using data adjusted to 2020 prices. Figure 6 presents the CBVA across 15 sectors in 31 provinces for the years 2017, 2018, and 2020. The results reveal four distinct provincial patterns: (1) Continuously Rising Type: Mainly in eastern coastal and emerging central and western regions such as Guangdong, Jiangsu, Zhejiang, Hunan, and Sichuan. Jiangsu (+246.8 billion yuan) and Guangdong (+131.7 billion yuan) showed the largest CBVA growth, driven by services, construction, and high-end manufacturing. For example, Jiangsu's CBVA in transportation equipment manufacturing rose from 176.8 billion to 376.4 billion yuan. Education, healthcare, public administration (+745.5 billion yuan), services (+501.8 billion yuan), and construction (+405.4 billion yuan) also steadily grew, serving as key drivers of CBVA expansion. (2) Increase-Then-Decrease Type: Includes traditional industrial and service-oriented provinces such as Beijing, Liaoning, and Henan, as well as provinces like Qinghai, Ningxia, Inner Mongolia, Gansu, and Hainan. Beijing's service CBVA rose 32% from 114.1 billion to 150.7 billion yuan in 2018 but dropped to 130.9 billion yuan in 2020. Hainan's service CBVA increased slightly in 2018 but fell sharply by 56% in 2020. This decline was mainly due to the impact of the COVID-19 pandemic, which severely affected high-contact sectors such as tourism, catering, and offline services, leading to a significant reduction in service consumption and thereby suppressing CBVA growth in the service sector. (3)

Continuous Decline Type: Provinces such as Heilongjiang, Xinjiang, and Hubei showed steady CBVA decreases over the period. For instance, Xinjiang's service CBVA fell from 34.5 billion yuan in 2018 to 29.83 billion yuan in 2020, mainly due to declines in transportation (-80.2 billion yuan). However, education, healthcare, and public administration sectors continued growing, with CBVA rising from 300.8 billion to 396.9 billion yuan. This indicates that spending on public services and healthcare increased significantly during the pandemic, becoming a key pillar supporting regional CBVA. (4) Decrease-Then-Increase Type: Includes Shanxi, Shandong, Jiangxi, Jilin, and Hebei. For example, Shandong's service sector grew from 127.1 billion to 165.1 billion yuan, while public service spending rose from 96.4 billion to 128.3 billion yuan. Overall, provincial changes in consumption-based value-added highlight regional disparities in industrial structure and resilience, supporting the validity and stability of the MRIO model for regional economic analysis.

(4) Decrease-Then-Increase Type: This group includes regions such as Shanxi, Shandong, Jiangxi, Jilin, and Hebei. For example, Shandong and Hebei experienced declines in construction and manufacturing in 2018, but saw a recovery in 2020, driven by growth in the service sector and public sectors such as education and healthcare. In Shandong, the service sector and public service spending were the top two contributors to CBVA growth, rising from 127.08 billion yuan to 165.04 billion yuan and from 96.42 billion yuan to 128.28 billion yuan, respectively. Similarly, in Hebei, these sectors led the CBVA rise, with the service sector growing from 49.56 billion to 79.12 billion yuan and public service spending from 46.21 billion to 65.96 billion yuan.

Data availability

These MRIO tables are publicly available via the China Emission Accounts and Datasets (CEADs, www.ceads.net) and Figshare⁴⁶ (<https://doi.org/10.6084/m9.figshare.29927291>).

Code availability

The programs used in the data generation is based on MATLAB and GAMS. The code can be found in https://github.com/LiJie20230/China_MRIO.

References

- 1 Bednarski, L., Roscoe, S., Blome, C. & Schleper, M. C. Geopolitical disruptions in global supply chains: a state-of-the-art literature review. *Production Planning & Control* **36**, 536-562 (2025).
- 2 Xing, Y. China and global value chain restructuring. *China Economic Journal* **15**, 310-329 (2022).
- 3 Guan, D. *et al.* Global supply-chain effects of COVID-19 control measures. *Nature human behaviour* **4**, 577-587 (2020).
- 4 Xing, Y. *Decoding China's export miracle: A global value chain analysis.* (World Scientific, 2021).
- 5 Handfield, R. B., Graham, G. & Burns, L. Corona virus, tariffs, trade wars and supply chain evolutionary design. *International Journal of Operations Production Management* **40**, 1649-1660 (2020).
- 6 Wu, J., Wood, J., Oh, K. & Jang, H. Evaluating the cumulative impact of the US–China trade war along global value chains. *The World Economy* **44**, 3516-3533 (2021).
- 7 Bo, M. & Ankai, X. Rising risks to global value chains. (2021).
- 8 Itakura, K. Evaluating the impact of the US–China trade war. *Asian Economic Policy Review* **15**, 77-93 (2020).
- 9 Huang, Y., Lin, C., Liu, S. & Tang, H. Trade networks and firm value: Evidence from the US–China trade war. *Journal of International Economics* **145**, 103811 (2023).
- 10 Pu, M. & Zhong, Y. Rising concerns over agricultural production as COVID-19 spreads: Lessons from China. *Global food security* **26**, 100409 (2020).
- 11 NBSC. *Monthly Report on Manufacturing Purchasing Managers Index*, <https://www.stats.gov.cn/english/PressRelease/202007/t20200701_1771424.html> (2020).
- 12 Lin, J. *et al.* China's international trade and air pollution in the United States. *Proceedings of the National Academy of Sciences* **111**, 1736-1741 (2014).
- 13 Panwar, R., Pinkse, J. & De Marchi, V. The future of global supply chains in a post-COVID-19 world. *California Management Review* **64**, 5-23 (2022).
- 14 Xia, C. *et al.* Outsourced carbon mitigation efforts of Chinese cities from 2012 to 2017. *Nature Cities* **1**, 480-488 (2024).
- 15 Xia, C. *et al.* Heterogeneity in carbon footprint trends and trade-induced emissions in China's urban agglomerations. *Communications Earth & Environment* **6**, 723 (2025).
- 16 Zheng, H. *et al.* Rising carbon inequality and its driving factors from 2005 to 2015. *Global Environmental Change* **82**, 102704 (2023).
- 17 Zheng, H. *et al.* Regional determinants of China's consumption-based emissions in the economic transition. *Environmental Research Letters* **15**, 074001 (2020).
- 18 Li, S. *et al.* Revisiting Copenhagen climate mitigation targets. *Nature Climate Change* **14**, 468-475 (2024).
- 19 Huo, J., Meng, J., Zheng, H., Parikh, P. & Guan, D. Achieving decent living standards in emerging economies challenges national mitigation goals for CO₂ emissions. *Nature Communications* **14**, 6342 (2023).
- 20 Wang, Q. *et al.* Implications of demographic policies on China's food-related environmental footprints amid population ageing. *Global Environmental Change* **95**, 103082 (2025).
- 21 Zhang, Z., Yu, Y., Zheng, H. & Tian, P. The intersecting impact of aging and affluence on China's household

carbon emissions. *Environmental Impact Assessment Review* **115**, 107997 (2025).

22 Zhang, Y., Liu, Y. & Li, J. The methodology and compilation of China multi-regional input-output model. *Stat. Res* **29**, 3-9 (2012).

23 Zhang, Z., Shi, M. & Zhao, Z. The compilation of China's interregional input–output model 2002. *Economic Systems Research* **27**, 238-256 (2015).

24 Liu, W., Tang, Z., Han, M., Li, F. & Liu, H. The 2012 China multi-regional input-output table of 31 provincial units. *The State Council of the People's Republic of China* (2018).

25 Liu, W., Li, X., Liu, H., Tang, Z. & Guan, D. Estimating inter-regional trade flows in China: A sector-specific statistical model. *Journal of Geographical Sciences* **25**, 1247-1263 (2015).

26 Liu, W., Tang, Z., Chen, J. & Yang, B. China's interregional input-output tables between 30 provinces in 2010. *China Statistics Press* (2014).

27 Liu, W. *et al.* Theories and practice of constructing China's interregional input-output tables between 30 provinces in 2007. *China Statistics Press* (2012).

28 Mi, Z. *et al.* A multi-regional input-output table mapping China's economic outputs and interdependencies in 2012. *Scientific data* **5**, 1-12 (2018).

29 Zhao, Q. *et al.* An inter-regional input-output table series of China from 1987–2017 with integrated carbon emission data. *Scientific Data* **11**, 1335 (2024).

30 Zheng, H. *et al.* Chinese provincial multi-regional input-output database for 2012, 2015, and 2017. *Scientific data* **8**, 244 (2021).

31 Wang, Y. *et al.* An inter-city input-output database distinguishing firm ownership in the Greater China area during 2002–2017. *Scientific data* **12**, 723 (2025).

32 Zheng, H. *et al.* Linking city - level input – output table to urban energy footprint: Construction framework and application. *Journal of Industrial Ecology* **23**, 781-795 (2019).

33 Zheng, H. *et al.* Entropy-based Chinese city-level MRIO table framework. *Economic Systems Research* **34**, 519-544 (2022).

34 Li, J. *et al.* Consumption - based carbon emissions of 85 federal entities in Russia. *Earth's Future* **12**, e2023EF004323 (2024).

35 Zheng, H. *et al.* Leveraging opportunity of low carbon transition by super-emitter cities in China. *Science Bulletin* **68**, 2456-2466 (2023).

36 Wang, Y., Geschke, A. & Lenzen, M. Constructing a time series of nested multiregion input–output tables. *International Regional Science Review* **40**, 476-499 (2017).

37 Jahn, M. Extending the FLQ formula: a location quotient-based interregional input–output framework. *Regional Studies* **51**, 1518-1529 (2017).

38 Sargent, A. L., Ramos, P. N. & Hewings, G. J. Inter-regional trade flow estimation through non-survey models: An empirical assessment. *Economic Systems Research* **24**, 173-193 (2012).

39 NBSC. *China regional input-output table (in Chinese: 中国地区投入产出表-2017)*. (China Statistical Publishing House, 2020).

40 Bekker, J. & Aldrich, C. The cross-entropy method in multi-objective optimisation: An assessment. *European Journal of Operational Research* **211**, 112-121 (2011).

41 Batty, M. Reilly's challenge: new laws of retail gravitation which define systems of central places. *Environment and Planning A: Economy and Space* **10**, 185-219 (1978).

42 Yamada, M. Construction of a multi-regional input-output table for Nagoya metropolitan area, Japan. *Journal of Economic Structures* **4**, 1-18 (2015).

43 Anderson, J. E. The gravity model. *J Annu. Rev. Econ.* **3**, 133-160 (2011).

44 Wei, W. *et al.* A 2015 inventory of embodied carbon emissions for Chinese power transmission infrastructure projects. *Scientific Data* **7**, 318 (2020).

45 Többen, J. On the simultaneous estimation of physical and monetary commodity flows. *Economic Systems Research* **29**, 1-24 (2017).

46 Li, J. *et al.* *China's Provincial Multi-Regional Input-Output Database for 2018 and 2020*, < <https://doi.org/10.6084/m9.figshare.29927291>> (2025).

Author Contributions

H.Z. designed the study and led the project. J.L. conducted the modelling. J.L., Z.Z., D.L., Q.O., P.T., D.G., and H.Z. contributed to the writing. J.L., Z.Z., and H.Z. collected the raw data.

Competing Interests

The authors declare no competing interests.

Acknowledgements

We sincerely acknowledge the support of the Carbon Neutrality and Energy System Transformation programme and the anonymous reviewers.

Funding

This research was funded by the National Natural Science Foundation of China (72242105 and 72361137002), the National Key R&D Program of China (2023YFE0113000), and the Energy Economics Submission Fee Fund (0140-9883 / © 2025 Elsevier B.V.). Additional funding was provided by the European Union under grant agreement no. 101137905 (PANTHEON) and the Research Grants Council of the Hong Kong Special Administrative Region, China (AoE/P-601/23-N).