



## Perspective

## Advancing UN Comtrade for physical trade flow analysis

Wei-Qiang Chen<sup>a,b,c,\*</sup>, Heming Wang<sup>d</sup>, Nan Li<sup>a,b,c</sup>, Peng Wang<sup>a,b</sup><sup>a</sup> Key Lab of Urban Environment and Health, Institute of Urban Environment, Chinese Academy of Sciences, Xiamen 361021, China<sup>b</sup> Xiamen Key Lab of Urban Metabolism, Xiamen 361021, China<sup>c</sup> University of Chinese Academy of Sciences, Beijing 100049, China<sup>d</sup> State Environmental Protection Key Laboratory of Eco-Industry, Northeastern University, Shenyang 110819, China

## ARTICLE INFO

## Keywords:

UN Comtrade

Physical trade

Material flow analysis database

Sustainability science

Natural resources fuel economic development and underpin social well-being. Since the 1970s, global natural resource extraction has tripled (IRP, 2019). Due to differences in natural resource endowments across countries, trade plays an essential role in mitigating regional imbalances in resource availability and affecting human well-being. During the same period, the physical trade increased by about five times, from 2.5 billion tons to 12.7 billion tons (Fig. 1), and fossil fuels are the most traded resources, accounting for around half of the global total. As trade prices are frequently influenced by world events (such as wars and financial crises), studies on trade using physical angles attain increased importance. In addition, studying physical trade can facilitate the analysis of how trade 1) supports global systems of production and consumption to increase global social welfare through resource re-allocation, 2) changes the location and scale of both environmental and social impacts, and 3) carries supply chain risk for critical minerals and commodities.

To account for the above analysis associated with trade, data on physical trade flows are vital. For example, physical trade flows are essential components of economy-wide material flow indicators, such as domestic material consumption (DMC) and physical trade balance (PTB). They are also the basis for tracking anthropogenic flows of substances (Graedel, 2019), such as Al, As, Cu, and Fe. In addition to material flow analysis (MFA), physical trade flows are the cornerstone for the life cycle assessment (LCA) approach to accounting for the embodied environmental impacts of trade.

Among the existing databases that can cover physical trade, such as FAO, IEA, Global Material Flows Database, and Industrial Ecology Data

Commons prototype, the database of UN Comtrade is one of the most widely used. That is because it has the broadest coverage of commodity categories and reporters. It has supplied trade information to policy-makers, business communities, academic institutions, and the general public for more than 50 years. However, it still has some data quality issues like outliers, missing values, and bilateral asymmetries, as we categorized in our follow-up papers:

- Outliers refer to the sampling points outside the data group's main body.
- Missing values refer to the incomplete data values, mainly in the physical commodity trade dataset.
- Bilateral asymmetries mean that importers' records do not match the exporters' records in the same transaction.

Those issues can lead to considerable trade misunderstanding, limiting the application of UN Comtrade. For example, the issue of outliers could cause violent fluctuations of time-series data, leading to a misleading interpretation; the issue of missing values could cause an underestimating of material flows and environmental influences, even resulting in net flow reversals (e.g., shifting from net importers to net exporters). Therefore, this is a significant challenge to address, as the data quality can substantially influence the research reliability and further mislead the resource, environmental, and public policy guidance.

In the few studies where those above issues have been dealt with, valuable experiences have emerged. For example, when analyzing the

\* Corresponding author.

E-mail address: [wqchen@iue.ac.cn](mailto:wqchen@iue.ac.cn) (W.-Q. Chen).<https://doi.org/10.1016/j.resconrec.2022.106520>

Received 2 July 2022; Accepted 3 July 2022

Available online 14 July 2022

0921-3449/© 2022 Elsevier B.V. All rights reserved.

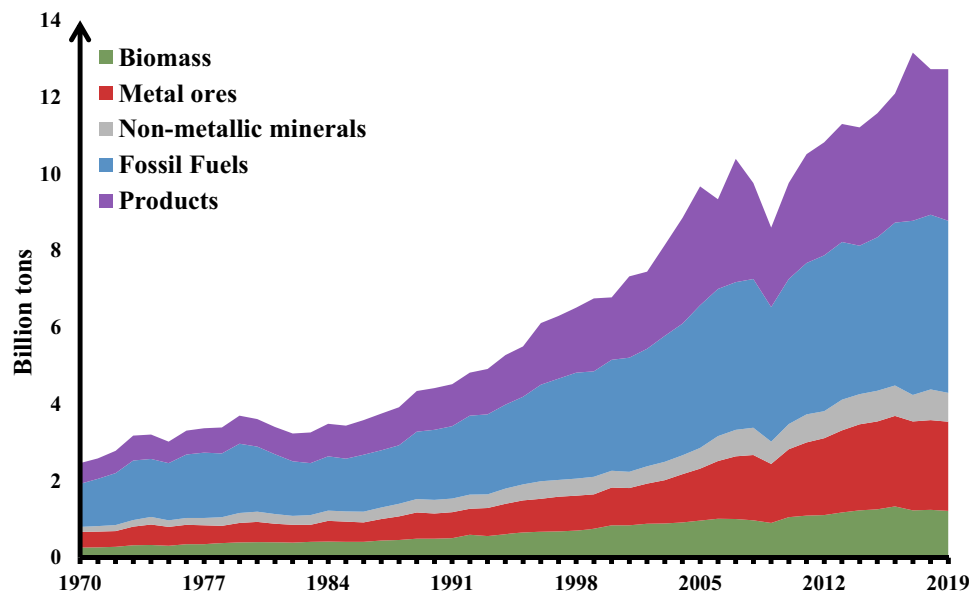


Fig. 1. Trends of global physical trade by five main material categories, 1970 – 2019, million tons. Data source: revised based on UN Comtrade dataset.

European Union's trade policies, Damerval (2012) used three methods to detect outliers: wavelets, Kalman filter, and forward search methods. (Dittrich and Bringezu, 2010) used the average global unit price (for a specific commodity in a particular year) to estimate missing weight data. To solve bilateral asymmetries, Gehlhar (1996) introduced a quality index to assess the reliability of each reporter and adopted the data from reporters with a higher quality index; United Nations harmonized the data by setting up a model to address the issues that could lead to bilateral asymmetries, such as the trading system, monetary valuation, partner attribution, etc.

Despite these successes, there are still many problems to be solved. For instance, some usual and real data might be identified as outliers when detecting outliers. When estimating the missing weight data by the average global unit price, people assumed all the reporters make deals for this commodity at the same price. Thus, this method is sensitive to outliers and some commodities with significant unit price variance, like HS 711,100 (Metals, clad with platinum, semi-manufactured). In this case, this method may over/underestimate the missing weight of trade commodities. Furthermore, how to choose a more appropriate method in a particular situation is also a challenge.

The three papers that follow aim to address the above issues for improving the data quality and constructing a public physical trade flow database. Paper I (Chen et al., 2022) comprehensively reviews data statistics criteria and preprocessing procedures and discusses the three above-mentioned data quality issues in UN Comtrade. Paper II (Jiang et al., 2022) tries to address the issue of outliers by developing a framework of outlier detection and handling and applies the methods to all the records of UN Comtrade during 1988–2019. Finally, Paper III (Zhang et al., 2022) identifies different types of missing values in the UN Comtrade database and modifies the missing physical data and outliers by a framework integrating method selection and data estimation. The physical trade flow database can be accessed via the following website: [www.macycle.org/improved-un-comtrade-data/](http://www.macycle.org/improved-un-comtrade-data/).

Overall, this series of articles aims to provide a higher-quality physical trade flow database, hoping that it will provide data support

for sustainability science and other disciplines using physical trade data. Extended research could be carried out from the aspects of methods (MFA, LCA, complex network, etc.), indicators (PTB, average shortest path length, environmental impact indicators, etc.), and products (integrated products or critical products).

#### Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### Data Availability

Data will be made available on request.

#### Acknowledgements

This research was supported by the National Natural Science Foundation of China (No. 71961147003, No. 52070178, No. 52170184, No.41871204, and No.52070034).

#### References

- Damerval, C., 2012. Identification of outliers in simultaneous time series using wavelets and forward search. <https://doi.org/10.2788/47188>.
- Dittrich, M., Bringezu, S., 2010. The physical dimension of international trade. Part 1: direct global flows between 1962 and 2005. *Ecol. Econ.* 69, 1838–1847. <https://doi.org/10.1016/j.ecolecon.2010.04.023>.
- Gehlhar, M.J., 1996. Reconciling bilateral trade data for use in GTAP. *GTAP Tech. Pap.* 1–31. No. 10.
- Graedel, T.E., 2019. Material flow analysis from origin to evolution. *Environ. Sci. Technol.* 53, 12188–12196. <https://doi.org/10.1021/acs.est.9b03413>.
- IRP, 2019. Global resources outlook 2019: natural resources for the future we want: summary for policy makers. Nairobi, Kenya.