

CONSTRUÇÃO DE MATRIZES DE INSUMO-PRODUTO A PARTIR DE NOTAS FISCAIS ELETRÔNICAS: UMA ANÁLISE BIBLIOMÉTRICA

Daniel Oliveira dos Santos – dan-@outlook.com

Program of Postgraduate in Systems and Product Engineering – Federal Institute of Bahia

Fabício Leite – fabricao.leite@ufba.br

Program of Postgraduate in Economics – Federal University of Bahia

Rilton Gonçalves Bonfim Primo – rilton@ufba.br

Program of Postgraduate in Industrial Engineering - Federal University of Bahia

Marcelo Santana Silva – marcelosilva@ifba.edu.br

Program of Postgraduate in Systems and Product Engineering – Federal Institute of Bahia

Renato Novais – renato@ifba.edu.br

Program of Postgraduate in Systems and Product Engineering – Federal Institute of Bahia

RESUMO: As matrizes de insumo-produto (MIP) são ferramentas essenciais para a análise econômica, mas sua elaboração tradicional, baseada em censos e pesquisas, é um processo caro e com grande defasagem temporal. Este artigo investiga o estado da arte na literatura científica sobre a construção e utilização de MIPs. Por meio de uma análise bibliométrica, utilizando as bases de dados Scopus e SciELO e o software VOSviewer, foi mapeada a produção científica sobre o tema. Os resultados revelam que a literatura *mainstream* é focada em estudos de aplicação das matrizes, não em métodos de construção inovadores. Especificamente, a abordagem de utilizar dados granulares e em tempo real, como os de notas fiscais eletrônicas (NF-e), é praticamente ausente. No entanto, a análise identifica uma fronteira de pesquisa emergente, com trabalhos pontuais de alto impacto que validam essa nova metodologia. O estudo conclui posicionando a contribuição deste trabalho: a adaptação e aplicação de um framework metodológico baseado em NF-e para o contexto brasileiro, preenchendo uma lacuna significativa e alinhando o país a uma tendência global de modernização das estatísticas econômicas.

Palavras-chave — Análise de insumo-produto, big data, matriz de insumo-produto, nota fiscal eletrônica.

ABSTRACT Input-output models are essential tools for economic analysis, but their traditional development, based on censuses and surveys, is a costly process with significant time lags. This article investigates the state of the art in the scientific literature on their construction and usage. Through a bibliometric analysis using the Scopus and SciELO databases and VOSviewer software, we mapped the scientific production on the topic. The results reveal that the mainstream literature focuses on studies of the application of the matrices, not on innovative construction methods. Specifically, the approach of using granular, real-time data, such as that from electronic invoices (NF-e), is virtually absent. However, the analysis identifies an emerging research frontier, with specific, high-impact studies that validate this new methodology. The study concludes by highlighting the contribution of this work: the adaptation and application of an NF-e-based methodological framework to the Brazilian context, filling a significant gap and aligning the country with a global trend toward modernizing economic statistics.

Keywords — Input-output analysis, big data, input-output matrix, electronic invoice.



1 INTRODUÇÃO

A formulação de políticas públicas e decisões de investimento eficazes depende de uma análise precisa e atualizada das estruturas econômicas (GUILHOTO, 2011). Nesse contexto, a análise de insumo-produto (MIP) se destaca como uma estrutura poderosa para quantificar as interconexões setoriais (BATISTA, 2021; FREITAS et al., 2024). Contudo, a eficácia dessa análise depende criticamente da atualidade dos dados.

O principal desafio reside na grande defasagem temporal das matrizes oficiais. Por exemplo, no estado da Bahia, a matriz mais recente, no ano de 2025, foi publicada em 2020 com dados de 2012 (BAHIA, 2020), comprometendo sua capacidade de refletir a economia atual. Diante do advento de dados administrativos massivos (*Big Data*), como os de notas fiscais eletrônicas (NF-e), surge a oportunidade de modernizar a compilação das MIPs.

A utilização de dados de notas fiscais alinha-se a uma visão transformadora para as estatísticas, conhecida como "Economia em Tempo Real" (*Real-Time Economy*). Conforme discutido por Akagi (2024), o aproveitamento desses microdados tem o potencial de aumentar drasticamente a granularidade e a frequência das Contas Nacionais. Embora essa abordagem seja inovadora, não é totalmente inédita. Um trabalho pioneiro de Gao et al. (2019) demonstrou a viabilidade da construção de uma matriz multirregional para a China a partir de dados de VAT (o equivalente à NF-e). Em contraste, outras abordagens que buscam a tempestividade, como a do projeto EMERGING (Huo et al., 2022), ainda dependem da combinação de dados de comércio internacional com estatísticas nacionais, baseando-se em estimativas.

Este artigo, portanto, realiza uma análise bibliométrica para mapear o cenário global da pesquisa sobre MIPs, posicionando a abordagem via NF-e dentro desta literatura e destacando a lacuna específica para sua aplicação no Brasil.

2 METODOLOGIA

A prospecção bibliométrica foi realizada em julho de 2025, utilizando as bases de dados Scopus, de abrangência internacional, e SciELO, que concentra a produção latino-americana. O objetivo da busca foi identificar publicações que tratassem da construção e aplicação de Matrizes de Insumo-Produto (MIPs), com ênfase na detecção de estudos que utilizassem dados fiscais eletrônicos (por exemplo, NF-e) ou dados administrativos massivos. A fim de mapear o cenário científico, a estratégia de busca envolveu a combinação de descritores técnicos e setoriais com operadores booleanos, visando identificar publicações relevantes nos idiomas inglês e português, conforme detalhado na Tabela 1.

LOCAL	TERMOS DE PESQUISA
Scopus	TITLE-ABS-KEY (("input-output matrix" OR "input-output model" OR "input-output analysis" OR "input-output table") AND ("construction" OR "invoice")) AND PUBYEAR > 2019 AND PUBYEAR < 2026 AND (LIMIT-TO (SUBJAREA , "ENVI") OR LIMIT-TO (SUBJAREA , "ECON") OR LIMIT-TO (SUBJAREA , "BUSI")) AND (LIMIT-TO (DOCTYPE , "ar")) AND (LIMIT-TO (EXACTKEYWORD , "Input-output Analysis") OR LIMIT-TO (EXACTKEYWORD , "Input Output Analysis") OR LIMIT-TO (EXACTKEYWORD , "Input-output") OR LIMIT-TO (EXACTKEYWORD , "Input Output Model") OR LIMIT-TO (EXACTKEYWORD , "Input--output Analysis") OR LIMIT-TO (EXACTKEYWORD , "Input-output Modeling") OR LIMIT-TO (EXACTKEYWORD , "Input-output Table")))
SciELO	(matriz insumo-produto) OR (insumo-produto) OR (análise insumo-produto) OR (input-output analysis) OR (input-output)

Tabela 1 – Descritores e operadores booleanos utilizados por base

A busca inicial resultou em 264 artigos na Scopus e 14 artigos na SciELO. Os registros foram exportados em formato **RIS** e gerenciados no software Zotero para remoção automática de duplicatas, seguida de verificação manual.

A seleção final seguiu três etapas:

1. Leitura de título e resumo para exclusão de artigos irrelevantes.
2. Leitura introdutória (primeiras 2–3 páginas) dos artigos potencialmente elegíveis para confirmar a aderência ao escopo.
3. Inclusão apenas de artigos que:
 - a. Utilizassem ou discutissem MIPs como ferramenta central de análise;
 - b. Artigos originais publicados em periódicos.
 - c. Estivessem dentro das áreas de economia, negócios ou meio ambiente.

Critérios de exclusão:

- Artigos de revisão, editoriais, atas de conferência ou capítulos de livros.
- Trabalhos em que MIPs aparecessem apenas como elemento secundário;
- Artigos cujo tema central fugia do escopo de análise econômica, ambiental ou de negócios.

Após a aplicação dos critérios, foram selecionados **201 artigos** para compor o corpus final da análise.

3 RESULTADOS E DISCUSSÃO

A análise dos 201 artigos selecionados permitiu traçar um panorama detalhado da pesquisa sobre matrizes de insumo-produto.

3.1 CARACTERIZAÇÃO QUANTITATIVA DA LITERATURA

A Figura 1 mostra a distribuição das publicações entre 2020 e 2025. O interesse no tema se mantém robusto, com um pico de 41 artigos em 2024. Uma análise aprofundada das palavras-chave dos artigos deste ano revela os principais motores por trás desse crescimento: um foco intenso na quantificação de emissões de carbono e no uso como ferramenta de planejamento da economia. A queda observada em 2025 deve-se ao fato de a coleta de dados ter sido realizada em julho do mesmo ano, não representando uma tendência de declínio.

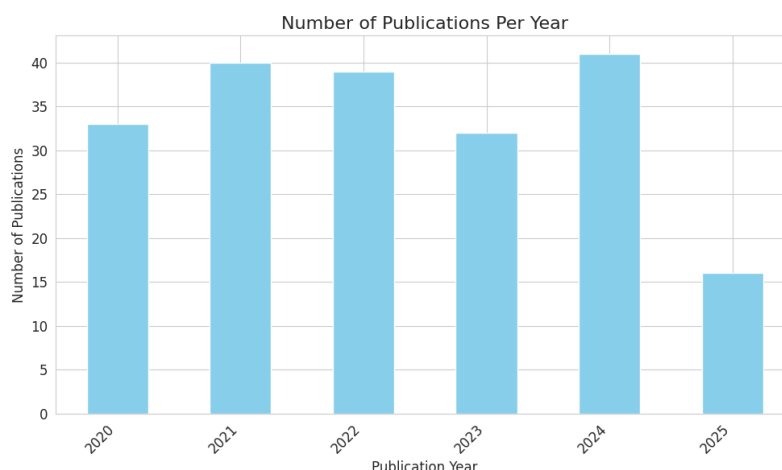


Figura 1: Distribuição temporal das publicações. Fonte: Autores

Em relação ao idioma, há um predomínio absoluto da língua inglesa, o que reflete a internacionalização do debate científico. A presença minoritária de artigos em português e espanhol, como visto na Figura 4, sugere um interesse regionalizado, mas com menor inserção nos periódicos de maior impacto global.

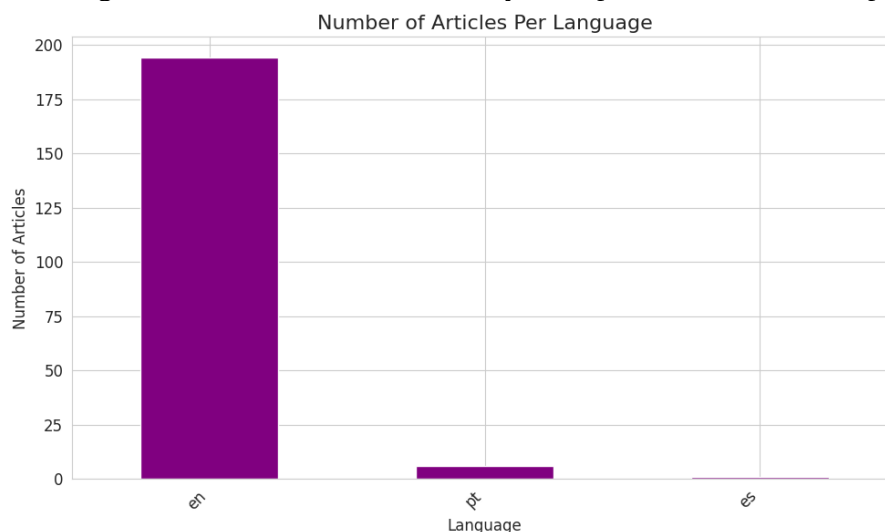




Figura 2: Distribuição das publicações por idioma. Fonte: Autores

3.2 ANÁLISE DAS REDES TEMÁTICAS E DE COLABORAÇÃO

Para aprofundar o entendimento sobre a estrutura do campo, foram geradas análises de co-autoria e coocorrência de palavras-chave com o software VOSviewer.

3.2.1 ANÁLISE DE COAUTORIA DE AUTORES

A análise da rede de coautoria (Figura 3) identificou 55 autores com colaborações recorrentes, organizados em múltiplos clusters. A estrutura da rede é caracterizada pela formação de vários grupos de pesquisa consolidados, que frequentemente colaboram internamente. A proeminência de nomes de origem asiática, especialmente chinesa, distribuídos por quase todos os clusters (como Wang, Y. no cluster azul, Li, H. no amarelo e Zhang, Z. no vermelho), indica uma liderança geográfica. A análise dos dados revela que pesquisadores afiliados a instituições chinesas são autores de 80 dos 201 artigos do corpus final. Este volume massivo não é acidental, mas reflete o uso da análise de insumo-produto como uma ferramenta estratégica para o planejamento e gestão de desafios nacionais, como será detalhado na análise temática a seguir.

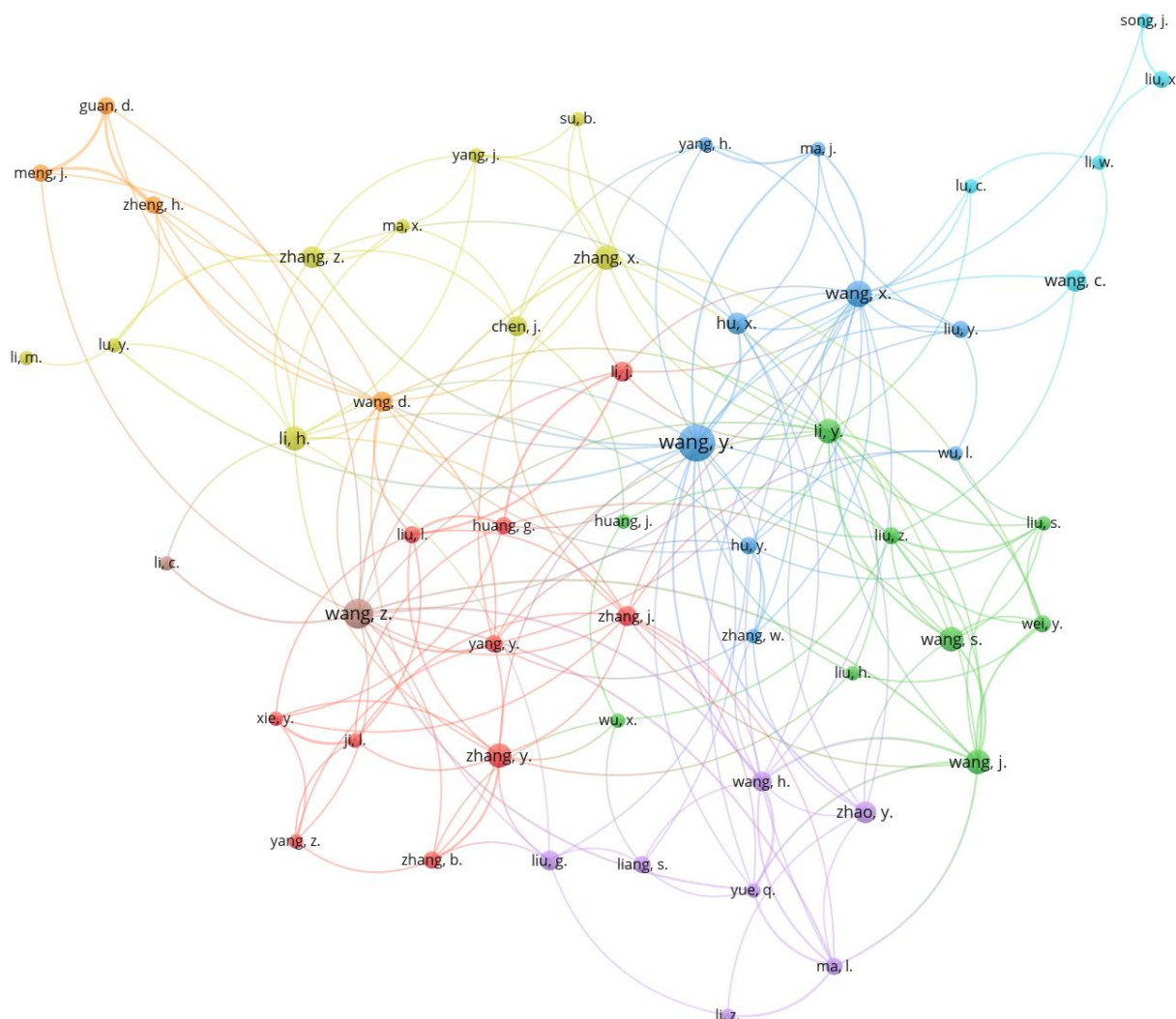


Figura 3: Rede de coautoria da literatura sobre MIPs. Fonte: Autores

3.2.2 ANÁLISE DE COOCORRÊNCIA DE PALAVRAS-CHAVE

A análise de concorrência, apresentada na Figura 4, revelou 55 palavras-chave principais, formando quatro clusters temáticos distintos. Esta análise confirma que a literatura, em geral, está focada na **aplicação** das MIPs, não em sua construção. O termo "*input-output analysis*" funciona como o nó central de toda a rede, conectando os diferentes eixos temáticos:

- **Cluster Vermelho (Impacto Ambiental e Climático):** É o maior e mais denso cluster, conectando o termo central a tópicos como "carbon emission", "emission control", "climate change", "sustainability" e "environmental economics". Isso demonstra que a principal aplicação das MIPs na literatura recente é para estudos de sustentabilidade e impacto ambiental.
- **Cluster Verde (Análise Setorial e Geográfica):** Liderado por "china", "construction industry" e "energy", este grupo foca em análises de setores específicos e cadeias de valor, com um claro enfoque geográfico na economia chinesa, corroborando os achados da análise de coautoria.

-
- The network visualization illustrates the conceptual structure of input-output analysis research. The central node, 'input-output analysis', is the most highly connected. It branches out into several major clusters:
- Environmental and Economic Impact (Red Nodes):** This cluster includes nodes like 'climate change', 'carbon emission', 'emission control', 'greenhouse gases', 'gas emissions', 'life cycle', 'life cycle analysis', 'environmental impact', 'decision making', 'manufacturing', 'sustainable development', 'spatiotemporal analysis', 'environmental economics', 'carbon footprint', 'input-output', 'economics', 'investments', and 'energy utilization'.
 - Carbon and Emissions (Green Nodes):** This cluster includes 'carbon', 'carbon dioxide', 'carbon emissions', 'carbon footprint', 'energy utilization', 'energy use', 'energy', 'construction industry', 'construction sectors', 'input-output modeling', 'input-output analysis', 'input-output model', 'economic growth', 'economic development', 'economic analysis', 'water supply', 'water', 'economic and social effects', 'article', 'industry', 'human', 'decomposition analysis', 'structural decomposition analysis', 'input-output analysis', 'structural path analysis', 'path analysis', 'supply chains', 'supply chain management', 'commerce', 'regression analysis', 'structural path analyse', and 'air pollution'.
 - Energy and Industry (Blue Nodes):** This cluster includes 'energy', 'energy utilization', 'energy use', 'energy', 'construction industry', 'construction sectors', 'input-output modeling', 'input-output analysis', 'input-output model', 'economic growth', 'economic development', 'economic analysis', 'water supply', 'water', 'economic and social effects', 'article', 'industry', 'human', 'decomposition analysis', 'structural decomposition analysis', 'input-output analysis', 'structural path analysis', 'path analysis', 'supply chains', 'supply chain management', 'commerce', 'regression analysis', 'structural path analyse', and 'air pollution'.
 - Supply Chain and Management (Yellow Nodes):** This cluster includes 'supply chain management', 'supply chains', 'path analysis', 'structural path analysis', 'regression analysis', 'structural path analyse', and 'air pollution'.
- The density of the network suggests a highly interdisciplinary and interconnected field of study, with many concepts overlapping and influencing each other.

3.3 A FRONTEIRA DA PESQUISA: POSICIONANDO A LACUNA METODOLÓGICA

No entanto, isso não significa que a ideia seja inexistente. A pesquisa aprofundada revelou estudos de vanguarda que definem a fronteira desta área. O trabalho de Gao et al. (2019) serve como um precursor



metodológico, ao construir uma matriz multirregional a partir de dados de imposto sobre valor adicionado (*Value-Added Tax*, VAT), extraídos de registros eletrônicos de faturas para fins fiscais Paralelamente, Akagi (2024) fornece o arcabouço conceitual, conectando o uso de *e-invoices* à visão de uma "Economia em Tempo Real". Em contraste, o projeto EMERGING de Huo et al. (2022) propõe uma matriz de “quase em tempo real” para economias emergentes, mas baseia-se em estimativas a partir de dados agregados (comércio bilateral e outros dados consolidados), e não em microdados de transações diretas

Portanto, a lacuna identificada por este estudo é mais sutil: não se trata da ausência completa da ideia, mas sim de sua rara aplicação prática, da falta de estudos para a economia brasileira e da necessidade de desenvolver e adaptar a metodologia para as especificidades do sistema de NF-e do Brasil.

4 CONCLUSÃO

A análise bibliométrica demonstrou que a pesquisa com matrizes de insumo-produto, embora robusta, concentra-se predominantemente na aplicação de matrizes já existentes, construídas por métodos tradicionais. A abordagem inovadora de utilizar dados fiscais massivos, como os da NF-e, representa uma fronteira de pesquisa emergente e pouco explorada.

Este estudo posiciona a proposta de um framework baseado em NF-e não como uma ideia isolada, mas como um passo necessário para alinhar a Bahia a essa vanguarda internacional. A principal contribuição é, portanto, a identificação e a proposição de um caminho para preencher uma lacuna metodológica crucial, adaptando uma abordagem de ponta para a realidade nacional. A implementação de tal metodologia tem o potencial de fornecer aos gestores públicos e pesquisadores uma ferramenta analítica com frequência e granularidade muito superiores, superando as limitações do paradigma atual.

Como trabalhos futuros, sugere-se a aplicação do framework metodológico aqui discutido para a construção de uma matriz para o estado da Bahia, validando sua eficácia e detalhando os desafios técnicos e conceituais da sua implementação no contexto brasileiro.

AGRADECIMENTOS

Este trabalho é parcialmente suportado pelo CNPq (Processo 101868/2024-3).

REFERÊNCIAS

GUILHOTO, J. J. M. Input-Output Analysis: Theory and Foundations (Análise de Insumo-Produto: Teoria e Fundamentos). SSRN Electronic Journal, 2011. ISSN 1556-5068. Disponível em: <<http://www.ssrn.com/abstract=1900073>>.

BATISTA, A. R. D. A. UM GUIA PRÁTICO PARA INFERIR A MATRIZ INSUMO-PRODUTO BRASILEIRA : VERSÃO SISTEMA DE CONTAS NACIONAIS 2010 [A PRACTICAL GUIDE TO INFERRING THE BRAZILIAN INPUT-OUTPUT MATRIX: 2010's NATIONAL ACCOUNTING SYSTEM VERSION]. 2021. Disponível em: <<http://rgdoi.net/10.13140/RG.2.2.18891.95528/1>>.

FREITAS, F. N. P. d. et al. Produto Riscos Fiscais: Desenvolvimento de Matriz de Avaliação da Estrutura Produtiva do Rio Grande do Sul. Rio de Janeiro, 2024. Processo nº 21/1400-0003061-9.

SANTANA, A. C. de. Cartilha sobre matriz de insumo produto: estruturação e análise dos resultados da matriz de insumo-produto dos estados da região Norte (1999). Belém, 2005. 27 p.

IEL/SC, I. E. L. D. S. C. Matriz Insumo Produto De Santa Catarina: Indústria Em Perspectiva. FLORIANOPOLIS, SC: Iel/Sc, 2022. ISBN 9786599900501.



BAHIA, G. do Estado da. Relatório da Matriz de Insumo-Produto. 2020. 34 p. Disponível em: <https://sei.ba.gov.br/images/publicacoes/download/relatorios/relatorio_matriz_2020.pdf>.

AKAGI, K. What does e-invoice data bring to SNA and real-time economy?. *Journal of Economic Structures*, v. 13, n. 15, 2024.

GAO, Y.; LI, M.; LU, Y. What Can Be Learned from Billions of Invoices? The Construction and Application of China's Multiregional Input-Output Table Based on Big Data from the Value-Added Tax. *Emerging Markets Finance and Trade*, 2019.

HUO, J. et al. Full-scale, near real-time multi-regional input-output table for the global emerging economies (EMERGING). *Journal of Industrial Ecology*, v. 26, p. 1218-1232, 2022.