Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 272 (2025) 251–260

Procedia Computer Science

www.elsevier.com/locate/procedia

The 16th International Conference on Emerging Ubiquitous Systems and Pervasive Networks (EUSPN 2025)
October 28-30, 2025, Istanbul, Türkiye

Causal Machine Learning Applied to Macroeconomic Analysis: Applications, Challenges, and Perspectives

Jirince K. Biaba[a,b], Christian M. Mulomba[c], Ibsen G. Bazie[a], Boaz N. Nzazi[a,b], Tasho Tashev[d], Witesyavwirwa V. Kambale[e], Kyandoghere Kyamakya[f], Nathanaël M. Kasoro[b], Vinh Ho Tuong[a], Selain K. Kasereka[b,d,*]

[a]*International School, Vietnam National University, Hanoi, Vietnam*
[b]*Department of Mathematics, Statistics and Computer Science, University of Kinshasa, Kinshasa, Democratic Republic of the Congo*
[c]*Department of Global Studies, Kyungsung University, 48434, 309 Suyeong-ro, Nam-gu, Busan, South Korea*
[d]*Department of Information Measurement Systems, Technical University of Sofia, Sofia, Bulgaria*
[e]*Faculty of Information and Communication Technology, Tshwane University of Technology, Pretoria, South Africa*
[f]*Institute of Smart Systems Technologies, University of Klagenfurt, Klagenfurt, Austria*

Abstract

Traditional econometric models such as VAR, VECM, and DSGE often struggle to identify true causal relationships in macroeconomic analysis, particularly in data-scarce and unstable environments common in developing countries. In contrast, CML offers a promising alternative by bridging predictive modeling and causal inference. This review explores CML applications in macroeconomics, with emphasis on challenges faced in low-income regions. Findings reveal fragmented adoption of CML, especially in Africa, where informal economies, limited data, and infrastructure gaps hinder implementation. While tools like DoWhy and CausalML show potential, fewer than 1% of studies incorporate spatiotemporal analysis, a key component for policy evaluation across diverse regions. To address these gaps, the review proposes a contextualized framework built around five pillars: contextual diagnosis, adaptive tool selection, methodological flexibility, collaborative validation, and progressive implementation. This approach reframes regional constraints as opportunities for innovation, encouraging the development of CML systems that are both rigorous and locally relevant. The study highlights the importance of designing CML frameworks that are accessible, culturally adapted, and capable of supporting robust policy evaluation in developing regions. Aligning ML tools with the realities of low-resource environments can unlock new pathways for evidence-based decision-making and inclusive economic development.

Keywords: Causal machine learning; Macroeconomic analysis; Time series; Causal inference; DoWhy; Policy evaluation.

* Corresponding author. Tel.: +243-821-828-964.
E-mail address: selain.kasereka@unikin.ac.cd

10.1016/j.procs.2025.10.203

1. Introduction

Understanding the causal mechanisms underlying macroeconomic dynamics is a fundamental challenge, both for academic research and for public policy-making. Traditional approaches, such as VAR (Vector Auto-Regression), VECM (Vector Error Correction Models) or DSGE (Dynamic Stochastic General Equilibrium) models, are based on strong assumptions of linearity, stationarity and exogeneity, which limit their ability to capture the complexity of real economic systems [1, 2, 3, 4, 5]. Although these models are useful for descriptive and predictive purposes, they struggle to identify genuine cause-and-effect relationships, particularly in contexts marked by structural instability or the availability of purely observational data [6].

The recent emergence of Causal Machine Learning (CML) opens up new perspectives by combining the predictive power of machine learning with the rigour of causal inference [7]. CML draws on two major theoretical foundations: the graph-based structural approach [7] and the potential outcomes framework [8]. These paradigms have given rise to modern tools such as `EconML` and `CausalML` [9, 10], initially designed for fields such as health, marketing, and social policy, but now also explored in macroeconomic analysis [11].

However, the application of CML to macroeconomics remains marginal and fragmented. Open questions persist regarding its validity in time series contexts, robustness to noise and data aggregation, and its ability to adapt to the realities of developing countries [6, 7, 11]. In the latter, the scarcity and heterogeneity of data are major obstacles to rigorous evaluation of public policies [11].

The objective of this study is to provide a comprehensive and critical review of recent developments in the use of CML within macroeconomic analysis. Specifically, the paper aims to:

- Examine the theoretical foundations and methodological frameworks underpinning causal inference in economic time series.
- Review key tools, libraries, and algorithms used in CML, highlighting their relevance to macroeconomic research.
- Explore practical applications of CML in macroeconomics, with particular attention to lessons learned, limitations encountered, and use cases from emerging economies.
- Analyze the strengths and weaknesses of leading CML tools through comparative evaluation.
- Present a context-aware adaptation framework for applying CML in African macroeconomic settings.
- Identify persistent challenges and outline future research directions for extending CML adoption and impact.

The remainder of this paper is organized as follows: Section 2 presents the research methodology used to identify and analyze relevant studies. Section 3 outlines the theoretical foundations of causal inference applied to time series data. Section 4 reviews key tools, libraries, and algorithms employed in Causal Machine Learning. Section 5 examines applications to macroeconomics, discussing practical cases, observed limitations, and key lessons. Section 6 provides a comparative analysis of leading CML tools. Section 7 introduces a contextual adaptation framework for African macroeconomic settings. Section 8 discusses ongoing challenges and outlines future research directions. Finally, Section 9 offers the concluding remarks.

2. Research Methodology

This review adopts a structured and targeted narrative approach aimed at critically synthesizing recent developments in Causal Machine Learning (CML) applied to macroeconomic analysis. The paper draws on solid theoretical foundations and practical applications drawn from real-world contexts, with the goal of identifying the most relevant contributions to understand the potentials, limitations, and prospects of CML applied to time series economic data, especially in low-income African countries, where policy decisions are often based on incomplete, noisy, or infrequent data.

2.1. Database Search and Scope

The literature review was conducted using four major academic databases: *Google Scholar*, *arXiv*, *IEEE Xplore*, and *SpringerLink*, in order to capture the conceptual and applied scope of causal machine learning applied to macroeconomic time series analysis in both general and African contexts. The selected keywords were designed to cover the conceptual and applied scope of the topic: *causal inference*, *causal machine learning*, *time series*, *macroeconomics*, *DoWhy*, *CausalML*, *EconML*, *Granger causality*, *DAG*, *DSGE*, *PCMCI*, *graphical models*, *deep causal learning*, etc.

Table 1 summarizes the approximate number of results returned for key thematic queries in each database:

Table 1. Search results by query and database

Query	Google Scholar	arXiv	SpringerLink	IEEE Xplore
`"Causal Machine Learning" AND macroeconomics AND "time series"`	23,500	129	85	41
`"Causal inference" AND DSGE AND macroeconomics`	11,200	73	60	34
`"Granger causality" AND "deep learning"`	19,800	142	91	58
`"DoWhy" OR "EconML" OR "CausalML"`	10,700	86	45	22
`"Causal learning" AND Africa`	1,240	18	7	5

Given the large volume of articles retrieved, a selective and transparent screening process was necessary. The search focused primarily on *peer-reviewed articles* published after 2018, a period that has seen rapid growth in automated causal reasoning and CML. This temporal threshold reflects the formal emergence of Causal Machine Learning (CML) as a distinct research field, building on foundational work such as the potential outcomes framework [8] and advances in causal graph approaches [1, 7]. These ideas were subsequently expanded through systematic surveys and methodological contributions appearing around 2018-2019 [7]. The aim was to identify methodologically sound studies applying CML to macroeconomic and time series analysis, particularly those relevant to African or low-resource policy contexts.

2.2. Inclusion and exclusion criteria

- **Included:** peer-reviewed *journal* articles from 2018 onward that apply or develop CML methods relevant to macroeconomic or time-series settings. A single foundational methodological exception is retained only for definitions/notation (Rubin, 2007 [8]), and is not counted in the empirical corpus.
- **Excluded:** conference proceedings; preprints and grey literature (e.g., arXiv/SSRN/TechRxiv) when not published in a peer-reviewed journal; books and book chapters; blogs and non-scholarly materials; purely predictive ML papers without causal identification; items not related to macroeconomic time series. When both a preprint and a journal version exist, only the peer-reviewed journal article is kept.

2.3. Final Corpus and Categorization

After a cross-reviewed selection (Biaba and Mulomba), a final set of **twelve** peer-reviewed *journal* articles (2018+) was retained based on *scientific quality*, *methodological diversity*, and *contextual relevance*. A single pre-2018 foundational source (Rubin, 2007) is cited solely for definitions/notation and is *not* counted in the empirical corpus. Table 2 groups the twelve articles into four thematic categories, with a brief description and the corresponding references.

Table 2. Final corpus (peer-reviewed journals, 2018+): categories, brief description, and included references (n=12)

Category	**Brief description**	**Included references**
Theoretical Foundations	Structural, counterfactual, and time-series causality	[1, 6, 7, 12]
Libraries & Algorithms	Toolkits (estimation) and discovery methods applicable to time series	[10, 13, 14, 15]
Macroeconomic Applications	Empirical/simulation studies on economic time-series data	[3, 5, 16]
African Context	Low-resource/African policy data and spatiotemporal analysis	[11]

Screening note. Items are grouped thematically; some references may be relevant to multiple categories. Foundational sources (e.g., [8]) are tracked separately and not counted in Table 2.

Example. Mulomba et al. (2025) analyze spatiotemporal macroeconomic shocks in Sub-Saharan Africa and discuss the practical constraints of CML under data sparsity [11].

3. Theoretical foundations of causal inference on time series

This section aims to introduce the core theoretical foundations of causal inference, in order to clarify how Causal Machine Learning (CML) can be meaningfully applied to time series data in macroeconomic contexts, especially in low-income countries. These frameworks are essential to evaluate the causal impact of public interventions such as a rise in interest rates, the launch of subsidy programs, or fiscal reforms, especially when randomization is not feasible.

Causal analysis aims to answer questions such as: What would be the consequences of a hypothetical intervention on a given economic system? However, in a macroeconomic context, data are often purely observational, making it difficult to distinguish between correlation and causality.

Two major theoretical frameworks structure modern causal inference. The first (Eq. (1)), known as the potential outcomes framework, is at the heart of many empirical methods in applied economics. The second paradigm (Eq. (2)), based on structural causal models, provides a powerful tool for reasoning about the causal structure of a system, even in the presence of latent variables.

- The Potential Outcomes Framework, which lays the formal foundations for estimates of the average treatment effect (ATE, ATT) by defining alternative scenarios (treated vs. untreated) for each observed unit. It is represented as shown in Eq. (1)

$$ATE = \mathbb{E}[Y(1) - Y(0)], \tag{1}$$

  where: $Y_i(1)$ is the outcome for unit i if treated, $Y_i(0)$ is the outcome for unit i if not treated, and E is the expectation or average.
  For example, in evaluating the effect of an agricultural subsidy program on rural GDP, ATE-based estimators can compare treated and untreated regions while controlling for confounding variables. However, in African settings where treatment is rarely randomized and many factors are unobserved, this framework must be adapted cautiously.
- The formalism of causal graphical models, which introduces directed acyclic graphs (DAGs) to visually represent causal relationships and identification rules by calculating counterfactuals via the “do-calculus.” Mathematically represented as presented in Eq. (2)

$$P(Y \mid do(X = x)), \tag{2}$$

  where P denotes the probability distribution, Y is the outcome variable of interest, and $do(X = x)$ refers to the *interventional* operation of setting the variable X to the value x. The vertical bar | indicates conditioning, meaning ”given” or ”conditional on”. The entire expression captures the distribution of Y when we actively intervene to set X to x, rather than merely observing instances where X happens to equal x.
  DAG-based models are especially valuable in assessing the effect of complex policy mixes, such as a simultaneous change in taxation and public investment. They allow encoding expert knowledge about economic dependencies. In African macroeconomic data, however, the required conditional independence tests may be unreliable due to noisy or sparse datasets.

In the specific case of economic time series, Granger causality presented in Eq. (3) has long been the standard.

$$Y_t = \sum_{i=1}^{p} \alpha_i Y_{t-i} + \sum_{j=1}^{q} \beta_j X_{t-j} + \epsilon_t, \tag{3}$$

where p and q are the maximum lag orders, α_i and β_j are the regression coefficients, and ϵ_t denotes the white noise error term. Variable X is said to Granger-cause Y if any $\beta_j \neq 0$, typically tested using an F-test.

It is based on the idea that a variable X causes Y if the past values of X improve the prediction of Y. Although practical, this approach remains correlational and does not capture the structural effects of interventions. For instance, it may detect if changes in interest rates precede shifts in inflation but cannot isolate the causal effect of a deliberate policy intervention. Modern extensions, such as *nonlinear causality* or *PCMCI models*, allow a better handling of the complex dynamics of economic systems [1].

Recent advances in causal deep learning make it possible to integrate these frameworks into neural architectures (graph neural nets, auto-encoders, etc.) capable of modeling complex dependencies while maintaining causal interpretability. These different foundations, although complementary, each pose challenges for application in low-income African countries, particularly due to the lack of data granularity, the absence of natural randomization, and the complexity of informal systems.

4. Tools, Libraries, and Algorithms in Causal Machine Learning

The rise of CML over the past decade has led to the development of a rich ecosystem of software libraries and algorithms for estimating causal effects from observational data, including in temporal contexts. These tools whose key features are presented in Table 3 have greatly contributed to democratizing access to automated causal inference, although their adoption in low-income countries remains limited. In addition to these libraries, there are causal discovery algorithms whose principles, strengths, and limitations are presented in Table 4 capable of automatically reconstructing causal graphs from data [6]. In cases of high non-linearity or noisy data, so-called causal deep learning approaches take over. Recent work incorporates architectures such as causal autoencoders or Graph Neural Networks (GNNs), capable of modeling both temporal dependencies and complex structural relationships [7].

Although powerful, these tools still face significant technical barriers in African countries: lack of long-term data, limited computing infrastructure, and lack of expertise in these emerging technologies [11]. It is therefore crucial to promote accessible, well-documented open-source tools that are adapted to local constraints in order to encourage widespread adoption in public economic policy.

Table 3. Main tools and methods in causal machine learning for macroeconomic time series

Tool / Method	Key Features	Reference
DoWhy	Causal graphs, counterfactuals, time series support	[7]
EconML	Meta-learners (S/T/X), Double ML for economics	[9]
CausalML	Treatment modeling, causal forests, neural nets	[10]
PCMCI, FCI, Bayesian nets	Causal discovery in time series, graph inference	[6]
Causal Deep Learning	GNNs, Autoencoders for nonlinearity and noise	[7]
Barriers in Africa	Sparse data, limited infra, need for local adaptation	[11]

Table 4. Comparison of major causal discovery algorithms for time series

Algorithm and Refs.	**Principle**	**Strengths**	**Limitations**
PCMCI (Peter and Clark Momentary Conditional Independence) [15, 13, 14]	Momentary conditional independence	Handles high-dimensional time series, reduces false positives	Requires large sample sizes
FCI (Fast Causal Inference) [12, 17]	Conditional independence with latent confounders	Handles hidden variables, partially directed graphs	Computationally intensive
Bayesian Networks [18, 19]	Probabilistic graphical models	Flexibility, encodes prior knowledge	Sensitive to prior assumptions

5. Applications to macroeconomics: lessons, limitations and practical cases

The application of causal learning in macroeconomics is still in its infancy, despite its obvious potential for evaluating public policies, modelling exogenous shocks and anticipating the effects of economic interventions. Recent work has focused mainly on three areas: structural analysis of DSGE models, detection of causalities in economic aggregates, and policy evaluation based on observational data.

For example, recent approaches extract the causal structure underlying state-space/DSGE-like models from real data, thereby circumventing strong parametric assumptions, using causal discovery tools such as PCMCI and dynamic graphical models [6, 18]. From a more predictive perspective, causal discovery has also been employed to establish temporal dependencies between macroeconomic, environmental, and socio-political variables, paving the way for multi-system modelling [6]. Beyond sectoral studies, Cordoni and Sancetta (2024) develop a high-dimensional time-series causal framework and illustrate macroeconomic use cases, including the impact of supply-side oil shocks on the economy [16].

In terms of public policy evaluation, recent studies emphasize that causal ML can complement traditional econometric approaches, particularly for heterogeneous-effect estimation and robustness in observational settings [7, 9, 10]. This is particularly useful in African countries, where statistical data quality often remains limited. Despite this, very few studies apply these methods to low-income countries. The exploratory work by Mulomba et al. [11] shows that the application of CML to monetary-policy evaluation in Sub-Saharan Africa remains hampered by the lack of long time series, the scarcity of plausible counterfactuals, and the difficulty of modelling economic informality. Nevertheless, a few emerging initiatives aim to fill this gap. Local researchers are adapting tools such as EconML and CausalML to specific contexts, including the impact of agricultural subsidies, conditional transfers, or external debt on poverty indicators. These efforts, although scattered, demonstrate the African community's growing interest in these approaches (Table 5). It is therefore imperative to strengthen local capacities, develop structured datasets and foster collaboration between researchers, public institutions and technical partners in order to make CML a real lever for macroeconomic policy management in Africa.

Table 5. Overview of CML applications in African or low-resource macroeconomic contexts

Reference	Country / Context	Application Focus
Mulomba et al. (2025) [11]	DR Congo, Sub-Saharan Africa	Spatiotemporal policy evaluation with CML on macro time series
Olayungbo (2019) [19]	Nigeria (LMIC)	Bayesian networks for monetary policy and macro performance
Guo et al. (2020) [7]	Global (incl. LMIC considerations)	Survey of causal methods for observational/limited-data settings
Zhao & Liu (2023) [10]	Software/package context	CausalML toolkit applicable in constrained environments

Note: These examples illustrate the emerging but still limited adoption of CML in African contexts. Most studies remain exploratory, highlighting the need for improved data availability, context-aware modeling strategies, and North–South research collaborations.

6. Comparative Analysis of CML Tools

The current ecosystem of causal machine learning tools offers distinct approaches with substantial differences in their philosophies, capabilities, and African context adaptability. Table 6 summarizes some features of two popular frameworks for CML.

It is worth noting that despite these, recent research of Mulomba et al. reveal that less than 1% of research explicitly combines CML and spatiotemporal analysis, highlighting this domain's emerging nature [11]. This rarity is particularly problematic for African contexts where spatiotemporal data often constitute the primary information source for public policy evaluation. Common limitations include assumption violation sensitivity, external generalization problems, and inadequacy to African data constraints. Existing tools generally presuppose robust computational infrastructures and well-structured datasets, conditions rarely met in developing countries.

Table 6. Comparison of Major Causal Inference Frameworks

Framework	Principle	Strengths	Limitations	Refs
DoWhy	Four-step causal analysis (Model, Identify, Estimate, Refute)	Graphical modeling, Robustness checks, Beginner-friendly	Struggles with high-dim data, Limited non-linear support	[7]
CausalML	Uplift modeling for business optimization	CATE estimation, Complex experiments, Targeting optimization	Steep learning curve, Data quality sensitive	[10]

7. Adaptation Framework for Africa

Given identified challenges, we propose a contextualized adaptation framework leveraging insights from the CML application literature in developing countries. The proposed adaptation framework consists of five key components:

(i) *Contextual Diagnosis*: The first step involves evaluating African context-specific constraints. Mulomba et al. identify three main challenge categories: temporal data sparsity, source heterogeneity, and the predominance of uncaptured informal systems [11]. This diagnosis must also consider local technical capabilities and infrastructure availability.
(ii) *Adaptive Tool Selection*: Tool choice must be guided by the diagnosed constraints. For exploratory analyses using limited data, DoWhy provides a more accessible option with built-in robustness tests. For tasks requiring heterogeneous effect estimation under imperfect data conditions, hybrid approaches combining spatiotemporal causal discovery and meta-learners may be more appropriate.
(iii) *Methodological Adaptation*: The framework incorporates techniques tailored to low-resource contexts. These include Bayesian methods for integrating prior knowledge, transfer learning strategies to address data scarcity, and spatiotemporal imputation methods to handle missing data.
(iv) *Collaborative Validation*: The validation process emphasizes input from local experts over solely statistical measures. This collaborative validation, seen in initiatives from Ghana, Senegal, and Rwanda, enhances contextual relevance and promotes adoption by policymakers [11].
(v) *Progressive Implementation*: The framework supports gradual skill building by starting with simple applications using user-friendly tools and progressing toward more advanced methodologies. This phased approach fosters local expertise while delivering actionable results from the outset.

The framework's effectiveness relies on transforming African constraints into methodological innovation opportunities, fostering an endogenous CML approach adapted to local realities.

8. Challenges and Future Directions

8.1. Methodological limitations

Despite notable advances in Causal Machine Learning (CML), several methodological limitations continue to hinder its effective deployment in macroeconomic analysis. Causal graph approaches, such as those based on Directed Acyclic Graphs (DAGs), are highly sensitive to violations of causal assumptions particularly the presumed absence of latent variables and feedback loops, which seldom holds true in complex economic systems [1]. Furthermore, many CML models lack external validity; algorithms trained on datasets from Western economies often fail to generalize to structurally distinct contexts such as African economies, introducing significant bias and limiting relevance [7, 11]. Identifiability also presents a major challenge. Persistent endogeneity and unobserved shocks in economic time series complicate efforts to reliably estimate causal effects, even when using sophisticated methods like double machine learning (DML) or meta-learners [9, 7]. Compounding these issues is the underdevelopment of real-time causal inference: most existing models are designed for retrospective analysis, while policymakers increasingly require dynamic tools capable of adapting to rapidly evolving macroeconomic conditions [11].

8.2. *Data limitations in low-income countries*

The deployment of Causal Machine Learning (CML) in low- and middle-income countries, particularly across Africa, is significantly constrained by issues of data availability and quality [7]. Macroeconomic time series data are often sparse, with frequent gaps, irregular updates, and inconsistent indicators. This limits the development of reliable temporal models [11]. Furthermore, the widespread informality of economic activities and the prevalence of untracked transactions introduce hidden confounders that violate core assumptions of many CML techniques [7]. Compounding these challenges is the limited availability of labeled datasets. Most CML frameworks depend on structured panel data with clearly defined treatment and outcome variables, which are rarely found in publicly accessible databases in the region. In addition, data governance in many African nations remains fragmented. Datasets are distributed across multiple institutions, often stored in incompatible formats and with limited accessibility, which hinders reproducibility and regional collaboration [11].

To address these challenges, it is essential to invest in the standardization of data collection protocols, promote open government data, and build partnerships with statistical agencies and academic institutions to generate usable causal datasets.

8.3. *Computational and human capacity constraints*

One of the major obstacles to the adoption of causal learning in African macroeconomic contexts is the weakness of digital infrastructure and the lack of specialised skills. On a technical level, many CML methods require high computing power, particularly when estimating causal forests, deep learning, or generating causal graphs over long time series [6, 7]. However, research institutions and economic administrations in low-income countries rarely have access to high-performance servers, GPUs, or suitable cloud platforms. This limits experimentation and prototyping of advanced causal models. Furthermore, mastering tools such as `DoWhy`, `EconML`, or `CausalML` requires dual expertise: in programming (Python, ML frameworks) and in advanced econometric methods [7, 9, 10]. However, this dual skill set is still rare among economic analysts in African countries, where specialised training in data science applied to economics remains marginal [11]. Added to this are difficulties in accessing technical documentation, which is often only available in English, as well as a weak culture of open science, which hinders the sharing of experiences and reproducible data sets [11].

8.4. *Interpretability and adoption barriers in policymaking*

One of the major challenges of Causal Machine Learning (CML) lies in the limited explainability of models for end users. While some algorithms such as causal trees or structured regressions are relatively transparent, many advanced methods (causal forests, causal neural networks, autoencoders) operate as black boxes, making them difficult to interpret for non-specialist analysts. This opacity undermines trust, especially when results are used to justify sensitive public policy decisions [7]. Moreover, most CML tools do not natively include mechanisms for explanation or visual justification. Users must resort to external techniques (counterfactuals, SHAP values, influence functions), which are not always well understood or available in public administrations of low-income countries [7]. Beyond the technical issue of interpretability, there are deeper barriers to the actual *adoption* of causal results in policymaking. In many African contexts, decision-makers are unfamiliar with empirical causal reasoning, often prefer normative classical methods, or are reluctant to challenge established approaches. Additionally, there is mistrust toward tools developed outside the continent, which are often seen as disconnected from local realities [11]. To overcome these obstacles, it is essential to develop more transparent tools with educational visualizations, and to reinforce collaboration between researchers, analysts, and decision-makers. The co-construction of explainable models with local stakeholders offers a promising path for turning CML into a true instrument for public governance.

8.5. Toward inclusive and localized CML frameworks

The effectiveness of Causal Machine Learning (CML) as a tool for economic governance strongly depends on its ability to adapt to local contexts. However, most existing libraries have been designed in Western environments, with institutional, linguistic, and technological assumptions that do not align with the realities of African countries. For instance, few tools offer multilingual interfaces, user-friendly modules for non-programmers, or visualizations adapted to low digital literacy settings. Standard workflows (such as those of `DoWhy`, `EconML`, or `CausalML`) often require well-structured datasets, which are either unavailable or poorly formatted in many local contexts [11, 9, 10]. This significantly limits the appropriation of these technologies by national analysts and planners. Moreover, the research priorities embedded in these tools do not always reflect the urgent challenges faced by countries in the Global South: market instability in agriculture, energy vulnerability, informal labor, dependency on foreign aid [11]. There is a critical need to develop "context-aware" CML libraries and frameworks, tailored to local constraints, priorities, and data cultures. This also means promoting "low-tech" approaches: lightweight solutions, documentation in local languages, simplified visual outputs, and interoperability with common tools such as Excel [7]. These adaptations would greatly enhance the practical use of CML in the design and evaluation of public policies in Africa.

8.6. Research and policy integration: the need for African-led initiatives

Beyond technical and methodological challenges, a major issue lies in the effective integration of CML research into public economic policymaking. Too often, innovations in causal modeling remain confined to academic environments or foreign research labs, without any concrete impact on local decision-making processes.

In many African countries, economic planning still relies on normative approaches or imported models that are rarely adapted to the local social and institutional context. The absence of bridges between researchers, statistical agencies, and public decision-makers contributes to the underutilization of locally produced knowledge [11]. It is therefore crucial to strengthen African scientific sovereignty in causal analysis. This requires funding projects led by African institutions, integrating CML into national planning and evaluation agencies, and establishing interdisciplinary research centers capable of combining technical expertise, field knowledge, and dialogue with policymakers.

Initiatives observed in Ghana, Senegal, and Rwanda in the fields of education, health, and agriculture demonstrate that locally anchored approaches enable not only better appropriation of tools but also greater trust in results. Promoting these models of endogenous scientific governance is essential to making CML a genuine lever for transforming economic policy in Africa.

9. Concluding Remarks

Causal Machine Learning (CML) offers a promising avenue for macroeconomic analysis by bridging statistical inference and data-driven modeling. Although its potential is considerable, practical applications remain limited due to methodological constraints such as identifiability challenges and sensitivity to assumptions. These limitations are compounded in African contexts by structural issues including data scarcity, informality, and infrastructural gaps. Despite these obstacles, CML could play a pivotal role in strengthening evidence-based decision-making in developing economies. Realizing this vision requires contextualizing tools to local realities through "context-aware" frameworks that accommodate instability and incomplete data, alongside investments in technical training and computing resources. Innovations should emphasize low-tech solutions with intuitive interfaces and documentation in regional languages to encourage broad adoption. Furthermore, drawing on successful collaborative models from various African nations may support the integration of CML into institutional processes. The path forward entails cultivating genuine scientific sovereignty by supporting local research initiatives and embedding causal modeling into academic programs. In doing so, CML can transition from a promising theory into a practical instrument for inclusive and adaptive economic governance.

Acknowledgments

This work has been realized with financial support by the European Regional Development Fund within the Operational Programme “Bulgarian national recovery and resilience plan”, Procedure for direct provision of grants “Establishing of a network of research higher education institutions in Bulgaria”, under the Project BG-RRP-2.004-0005 “Improving the research capacity and quality to achieve international recognition and resilience of TU-Sofia”.

References

[1] Ali Shojaie and Emily B. Fox. Granger causality: A review and recent advances. *Annual Review of Statistics and Its Application*, 9:195–220, January 2022.

[2] Sarit Maitra. Vector auto regression and vector error correction model. In *A Practical Guide to Static and Dynamic Econometric Modelling: Examples and Analysis with Python Code Embedded*, pages 171–201. Springer, 2025.

[3] Xiaohong Chen, Paul Wohlfarth, and Ron P Smith. China’s money demand in a cointegrating vector error correction model. *Journal of Asian Economics*, 75:101338, 2021.

[4] Jiangbo Xue. *A structural forecasting model for the Chinese macroeconomy*. Hong Kong University of Science and Technology (Hong Kong), 2009.

[5] Md Tanvir Rahman Mazumder, Md Shahadat Hossain Shourov, Iftekhar Rasul, Sonia Akter, and Md Kauser Miah. The impact of macroeconomic factors on the us market: A data science perspective. *Journal of Economics, Finance and Accounting Studies*, 7(2):208–219, 2025.

[6] Jakob Runge, Peer Nowack, Marlene Kretschmer, Seth Flaxman, and Dino Sejdinovic. Inferring causation from time series in earth system sciences. *Nature Communications*, 10:2553, May 2019.

[7] Ruocheng Guo, Lu Cheng, Jundong Li, P Richard Hahn, and Huan Liu. A survey of learning causality with data: Problems and methods. *ACM Computing Surveys (CSUR)*, 53(4):1–37, 2020.

[8] Donald B. Rubin. The design versus the analysis of observational studies for causal effects. *Statistics in Medicine*, 26(1):20–36, January 2007.

[9] Miruna Oprescu, Vasilis Syrgkanis, Keith Battocchi, Maggie Hei, and Greg Lewis. Econml: A machine learning library for estimating heterogeneous treatment effects. In *33rd Conference on Neural Information Processing Systems*, page 6, 2019.

[10] Yang Zhao and Qing Liu. Causal ml: Python package for causal inference machine learning. *SoftwareX*, 21:101294, 2023.

[11] Christian M. Mulomba, Vogel M. Kiketa, David M. Kutangila, Pescie H. K. Mampuya, N. Mukenze, Landry M. Kasunzi, Kyandoghere Kyamakya, Tasho Tashev, and Selain K. Kasereka. Applying causal machine learning to spatiotemporal data analysis: An investigation of opportunities and challenges. *IEEE Access*, 2025.

[12] Eric V Strobl, Shyam Visweswaran, and Peter L Spirtes. Fast causal inference with non-random missingness by test-wise deletion. *International journal of data science and analytics*, 6(1):47–62, 2018.

[13] Christopher Krich, Jakob Runge, Diego G Miralles, Mirco Migliavacca, Oscar Perez-Priego, Tarek El-Madany, Arnaud Carrara, and Miguel D Mahecha. Estimating causal networks in biosphere–atmosphere interaction with the pcmci approach. *Biogeosciences*, 17(4):1033–1061, 2020.

[14] Lykke Silfwerbrand, Yasuharu Koike, Pär Nyström, and Malin Gingnell. Directed causal effect with pcmci in hyperscanning eeg time series. *Frontiers in Neuroscience*, 18:1305918, 2024.

[15] Zhe Wang, Xiaogang Shi, Longcang Shu, Xiaoran Yin, Keke Zhou, and Pengcheng Xu. Quantifying climate factor contributions to groundwater level changes under different soil freezing-thawing states with the wt-pcmci model. *Journal of Hydrology*, 656:132997, 2025.

[16] Francesco Cordoni and Alessio Sancetta. Consistent causal inference for high-dimensional time series. *Journal of Econometrics*, 246(1-2):105902, 2024.

[17] Doris Entner and Patrik O Hoyer. On causal discovery from time series data using fci. *Probabilistic graphical models*, 16, 2010.

[18] Yue Liu, Yijing Wang, Guihuan Zheng, Jue Wang, and Kun Guo. The dynamical relationship between capital market and macroeconomy: based on dynamic bayesian network. *Procedia Computer Science*, 162:46–52, 2019.

[19] David Oluseun Olayungbo. Bayesian graphical model application for monetary policy and macroeconomic performance in nigeria. In *Bayesian Networks-Advances and Novel Applications*. IntechOpen, 2019.