IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 10 July 2025, accepted 27 July 2025, date of publication 7 August 2025, date of current version 15 August 2025.

Digital Object Identifier 10.1109/ACCESS.2025.3596680

SURVEY

Applying Causal Machine Learning to Spatiotemporal Data Analysis: An Investigation of Opportunities and Challenges

CHRISTIAN M. MULOMBA[1], VOGEL M. KIKETA[2,3], DAVID M. KUTANGILA[4], PESCIE H. K. MAMPUYA[4], JUNIOR N. MUKENZE[5], LANDRY M. KASUNZI[6], KYANDOGHERE KYAMAKYA[7], TASHO TASHEV[8], AND SELAIN K. KASEREKA[8,9,10]

[1]Department of Global Studies, Kyungsung University, Nam-gu, Busan 48434, South Korea
[2]School of Engineering, Seoul National University, Seoul 08826, South Korea
[3]Faculty of Economics and Management, University of Kinshasa, Kinshasa, Democratic Republic of the Congo
[4]Faculty of Computer Science, Protestant University of the Congo, Kinshasa, Democratic Republic of the Congo
[5]Department of Advanced Convergence, Handong Global University, Pohang 37554, South Korea
[6]Faculty of Computer Science, Université Nouveaux Horizons, Lubumbashi, Democratic Republic of the Congo
[7]Institute of Smart Systems Technologies, University of Klagenfurt, 9020 Klagenfurt am Wörthersee, Austria
[8]Department of Information Measurement Systems, Technical University of Sofia, 1000 Sofia, Bulgaria
[9]Department of Mathematics, Statistics and Computer Science, University of Kinshasa, Kinshasa, Democratic Republic of the Congo
[10]ABIL Research Center, Kinshasa, Democratic Republic of the Congo

Corresponding authors: Kyandoghere Kyamakya (kyandoghere.kyamakya@aau.at) and Selain K. Kasereka (selain.kasereka@tu-sofia.bg)

This work was supported by European Regional Development Fund within the Operational Program ''Bulgarian National Recovery and Resilience Plan,'' Procedure for Direct Provision of grants ''Establishing of a Network of Research Higher Education Institutions in Bulgaria,'' ''Improving the Research Capacity and Quality to Achieve International Recognition and Resilience of TU-Sofia'' under Project BG-RRP-2.004-0005.

ABSTRACT Traditional spatiotemporal data analysis often relies on predictive models that overlook causal relationships, making it difficult to identify true drivers and formulate effective interventions. To bridge this gap, we review causal machine learning (CML) techniques for spatiotemporal data, aiming to provide robust insights into their unique advantages. Our literature review reveals that fewer than 1% of studies in major databases explicitly integrate CML with spatiotemporal analysis. After rigorous screening, we analyze 51 relevant papers, categorizing their contributions into four key areas (totaling 62 methodological approaches due to multi-category papers): 1) causal effect discovery and estimation (32 approaches), 2) prediction accuracy enhancement (19), 3) pattern recognition limitations (10), and 4) interpretability (1). This distribution highlights a critical research gap, particularly in interpretability and comprehensive frameworks. We further examine unique challenges in spatiotemporal data, such as spatial autocorrelation and temporal dependencies, that complicate causal inference but also present opportunities for innovation. Promising approaches include the synergy of spatiotemporal Granger causality and structural equation modeling with spatial lags, which capture complex interdependencies while preserving interpretability. Future directions include developing interpretable causal models, advancing real-time causal inference in dynamic environments, and addressing computational challenges (scalability, efficiency, and complexity-interpretability trade-offs). We also discuss ethical considerations, such as bias mitigation in causal discovery and societal implications of spatiotemporal causal inference. By synthesizing challenges and opportunities, this work advances the application of CML in spatiotemporal analysis, with implications for climate science, economics, epidemiology, and urban planning.

INDEX TERMS Causal machine learning, spatiotemporal data analysis, synergy methods, ethics.

The associate editor coordinating the review of this manuscript and approving it for publication was Chao Tong.

I. INTRODUCTION

The rapid advancement of technology has led to an exponential growth in spatiotemporal data across diverse domains, including climate science, urban planning, epidemiology, and social sciences. While traditional methods such as Granger causality [1] and structural equation modeling [2] have established foundational approaches for causal inference in spatiotemporal data, they often struggle to uncover the underlying causal mechanisms in increasingly complex, high-dimensional datasets. This limitation significantly constrains the ability to develop effective interventions for critical real-world challenges [3], [4], [5], [6]. Causal machine learning (CML) has emerged as a powerful paradigm to address these limitations. Recent advances in hybrid approaches, particularly neural causal discovery models and graph-based learning, have expanded the frontiers of causal inference in several dimensions, such as in complex dependency modeling, where modern CML methods excel at capturing high-dimensional, non-linear relationships and spatiotemporal interdependencies that challenge traditional approaches [5], in scalability, where techniques like Graph Neural Networks (GNN) enable efficient representation of causal structures in large-scale applications (e.g., urban traffic networks or climate systems), overcoming the computational limitations of classical methods [6], or in data integration, where they facilitate the combination of observational data with domain knowledge, proving particularly valuable in settings with incomplete prior understanding of causal mechanisms [4]. Thanks to these advances, contemporary CML approaches can effectively identify causal structures in dynamic systems characterized by feedback loops and latent confounders (scenarios where traditional constraint-based or score-based methods typically fail), enabling more accurate identification of true causal drivers, supporting targeted interventions, improved decision-making, and deeper understanding of complex system dynamics [7].

Despite these advancements, the integration of causal machine learning with spatiotemporal data analysis remains significantly underexplored. Our review of peer-reviewed literature from reputable academic databases reveals that only a small fraction of studies explicitly address this critical intersection. This gap represents both a challenge and a substantial opportunity for methodological innovation.

This paper provides three main contributions to bridge this research gap by:

1) Providing a comprehensive synthesis of opportunities and challenges in applying causal machine learning to spatiotemporal analysis,
2) Highlighting on the potential of hybrid approaches that combine traditional and emerging methods to yield insights unattainable through conventional techniques alone [4], [5], [8],
3) Analyzing novel methodological combinations to address critical yet understudied aspects, particularly the interpretability of complex spatiotemporal causal models.

Our review advances the field through four principal dimensions:

1) **Theoretical foundations:** We systematically examine the interplay between causal machine learning (including graphical models and counterfactual reasoning) and core spatiotemporal principles (such as autocorrelation and non-stationarity), establishing a unified theoretical framework.
2) **Methodological integration:** We critically analyze both established techniques (e.g., spatiotemporal Granger causality, structural equation modeling with spatial lags) and emerging neural causal approaches, assessing their respective capacities to address challenges like:
   - High-dimensional dependency modeling,
   - Interpretability-scalability trade-offs,
   - Dynamic system representation.
3) **Practical implementation:** We synthesize actionable insights from diverse application domains (including epidemiology and urban mobility), identifying:
   - Common computational bottlenecks,
   - Data quality considerations,
   - Effective implementation strategies.
4) **Future Directions:** We present a forward-looking synthesis that:
   - Examines ethical considerations with emphasis on bias mitigation,
   - Maps emerging research frontiers, including:
     * Real-time causal analysis,
     * Human-in-the-loop frameworks,
     * Cross-domain transfer learning.

This study serves as both a reference and a roadmap for researchers, practitioners, and policymakers working at the intersection of causal inference and spatiotemporal analysis. The paper is organized as follows: Section II establishes foundational concepts; Section III surveys existing literature; Section IV details our review methodology; Section V examines integration approaches; Sections VI-VIII analyze challenges, ethical considerations, and future directions; and Section IX presents concluding remarks.

II. BACKGROUND

A. FUNDAMENTALS OF CAUSAL MACHINE LEARNING

Causal inference and machine learning represent two complementary paradigms in data science. While machine learning focuses on predictive accuracy through pattern recognition in data, causal inference aims to establish cause-and-effect relationships between variables [9], [10]. Applying causal inference analysis involves:

- **Causal effect identification**: Determining the existence of causal relationships between variables while distinguishing them from spurious correlations [11]. This step involves:

  * **Causal discovery**: Techniques for uncovering causal structures from observational data, including:
    - Constraint-based methods (e.g., PC and FCI algorithms),
    - Score-based approaches (e.g., Bayesian networks),
  * **Confounder control**: Addressing confounding variables that may bias observed relationships;
  * **Selection bias mitigation**: Ensuring the sample represents the target population;
- **Causal effect estimation**: Quantifying the magnitude and direction of causal effects [12]. Common approaches include:
  * Randomized Controlled Trials (RCTs)-the gold standard,
  * Observational methods (matching, weighting),
  * Instrumental variable analysis.
- **Counterfactual analysis**: Evaluating hypothetical scenarios to understand potential outcomes under different conditions or interventions [13]. This framework is essential for:
  * Assessing treatment effects,
  * Policy impact evaluation,
  * Robust causal model development.

Machine learning algorithms can complement causal inference by:

- **Pattern recognition in complex data**: Machine learning enhances causal effect identification through:
  * Automated feature extraction from high-dimensional datasets,
  * Dimensionality reduction for improved visualization and analysis,
  * Detection of non-linear relationships prevalent in real-world systems [3], [13].
- **Causal Effect Estimation**: Machine learning provides robust estimation techniques, including:
  * Advanced regression methods controlling for confounders,
  * Structural equation modeling for complex causal networks,
  * Double/debiased machine learning for observational data [3], [13].
- **Counterfactual Simulation**: Machine learning enables:
  * Synthetic data generation for hypothetical scenarios,
  * Treatment effect estimation through simulated interventions,
  * Outcome prediction under alternative conditions [3], [13].

The convergence of causal inference with machine learning presents significant opportunities for addressing complex real-world challenges, particularly in spatiotemporal data analysis, where it enables the extraction of actionable insights from dynamic systems, the identification of space-time-dependent causal mechanisms, or the development of data-driven intervention strategies [3], [12], [14].

B. FUNDAMENTALS OF SPATIOTEMPORAL DATA

1) CHARACTERISTICS AND CHALLENGES

Spatiotemporal data integrates both spatial (geographic) and temporal (time-based) dimensions, capturing phenomena that evolve across space and time. This data type is characterized by two key resolution parameters, namely, the spatial resolution, which is the granularity of geographic information, ranging from high resolution (precise location data) to low resolution (broad area coverage), and the temporal resolution, which is the sampling frequency, including high frequency (continuous monitoring) and the Low-frequency (periodic sampling). Common sources of spatiotemporal data include Geographic Information Systems (GIS), Satellite imagery, GPS trajectories, Sensor networks (Weather stations, traffic monitors), Digital platforms (Geotagged social media, mobile data), or Economic indicators (location-based market data) [16], [17]. It is worth noting that the selection of appropriate resolutions presents analytical challenges, as it directly impacts the validity and interpretability of results [18], [19].

2) UNIQUE CHALLENGES

Spatiotemporal data analysis faces several distinctive challenges, such as:

- **Data heterogeneity**, which involves mixed data types (numerical, categorical, geospatial) and varying measurement scales and formats that require specialized preprocessing pipelines [19];
- **Collection biases** in the process of data collection due to sampling bias in sensor placement or measurement errors in field data, which can induce representativeness concerns [15];
- **Complex dependencies** that they can exhibit, such as spatial autocorrelation (Tobler's First Law) or temporal autocorrelation (time-series dependencies), resulting in complex spatiotemporal interactions [20].

3) IMPORTANCE OF DOMAIN KNOWLEDGE

Domain knowledge is crucial for understanding and addressing the challenges of spatiotemporal data analysis. Expertise in a relevant field can help identify important variables and features, select appropriate data sources and collection methods, interpret results and draw meaningful conclusions, and address domain-specific challenges and limitations [19], [21].

C. TRADITIONAL VS. MODERN APPROACHES TO SPATIOTEMPORAL DATA ANALYSIS

Traditional approaches for spatiotemporal data analysis primarily employ statistical methods such as time series analysis, spatial regression, and geostatistics. Although these techniques have proven effective in many research contexts, they exhibit significant limitations when handling complex, nonlinear relationships or large-scale datasets. Widely used methods like Granger causality and structural equation modeling (SEM), for instance, often struggle with high-dimensional data, intricate temporal dependencies, and spatial challenges of heterogeneity that are increasingly common in contemporary spatiotemporal datasets. These shortcomings underscore the necessity for more advanced analytical techniques capable of addressing the demands of data-intensive research [16]. Modern (or hybrid) approaches, powered by advancements in machine learning and data mining, offer promising solutions to these challenges. Deep learning models, including Recurrent Neural Networks (RNNs), Convolutional Neural Networks (CNNs), and Long Short-Term Memory (LSTM) networks, excel at capturing temporal patterns and spatial features, making them well-suited for spatiotemporal analysis. Graph Neural Networks (GNNs) further enhance this capability by modeling complex spatial relationships and dependencies. Despite their strengths, these methods often lack explicit causal reasoning, limiting their ability to uncover the underlying mechanisms driving observed spatiotemporal phenomena [18], [22]. To bridge this gap, emerging hybrid methodologies are gaining prominence. Neural causal discovery models integrate deep learning with causal inference, enabling the identification of causal relationships in complex spatiotemporal datasets where traditional methods like Granger causality and SEM fall short. Similarly, graph-based learning techniques leverage graph structures to simultaneously capture spatial dependencies and temporal dynamics, offering a more comprehensive analytical framework. Hybrid models further enhance robustness by combining traditional statistical methods with modern machine learning (for example, integrating SEM with GNNs or augmenting spatial regression with RNN-based temporal modeling). These innovations provide a more nuanced and scalable approach to spatiotemporal data analysis [20], [23], [38].

III. AN OVERVIEW OF RESEARCH AND APPLICATIONS

A. CAUSAL MACHINE LEARNING

In [5], a survey of causal machine learning and open problems is provided. The authors cover important topics like the categorization of causal machine learning into five groups of depending on the problem to solve. This paper reviews specific areas of application and provides an overview of other benchmark solutions proposed by causal machine learning applied in computer vision, natural language processing, and graph representation learning. The relevance of considering causality in machine learning is presented in [37] where core concepts of this domain are explained alongside different areas of application. This survey provides a comprehensive review of causal machine learning by introducing frameworks for causalities and describing different methods applied in the domain. Recent developments in causal inference and machine learning are discussed in [4] where the authors review advances in causal inference with relevance to sociology. After a consistent definition and description of the incorporation of machine learning in causal inference, they presented the relevance of including causal machine learning in sociology, believing that this integration will reduce biases during analysis. The importance of causality analysis in scientific research is presented in [38]. In their research, the authors explained the relevance of considering causal machine learning in the field of oceanic science, where deep insight could be provided not only to understand natural phenomena but also to foster the explainability of machine learning (deep learning) algorithms. Their research was applied in China, providing satisfactory results in prediction and interpretability compared to the benchmark. After a review of 20 papers related to causal machine learning, in [8], the authors could highlight that Bayesian networks are commonly used in the context of causality while the propensity score is extensively used in causality research. Furthermore, their study not only systematically examined causal machine learning approaches used so far, but they also considered the categorization of data applied for causal machine learning based on the data type, value, and dimensionality of the data. This review offers a practical guide to the selection of causal machine learning systems. An application of causal machine learning to address fairness is introduced in [6]. After reviewing notions and methods to detect and eliminate algorithmic discrimination, the authors review causality-based fairness in various fields, highlighting advantages and disadvantages of its application. The value added of machine learning to causal inference is presented in [9]. Based on their study, which applies causal machine learning in the field of economics, they conclude that causal machine learning methods are more suitable than traditional methods to capture the effects of the covariates. Furthermore, they also argued that these methods of causal machine learning are useful in settings with many confounders relative to the sample size, making it suitable to treat such cases compared to the traditional approach. An application of causal inference and deep learning to estimate soybean yields using satellite remote sensing data is presented in [7]. Through an empirical study aiming to predict soybean yields using machine learning, the authors noticed that the integration of causality (causal relationship) improves prediction accuracy when dealing with spatiotemporal remote sensing data. Furthermore, they highlighted that this improvement is also the result of their novel framework, which combines a structural causal model with deep learning to provide a unique method to address this problem.

B. SPATIOTEMPORAL DATA ANALYSIS

The ever-growing importance of spatiotemporal data in research is discussed in [15] where the authors introduced novel ways to develop algorithms and technologies to capture, store, and analyze these data. Their claim is that these approaches are opening doors for various analyses suitable to deepen one's understanding of specific domains. They also discuss the future of spatiotemporal data analytics, which offers promising advancements. Challenges and opportunities presented in the domain of spatiotemporal data mining (STDM) are presented in [18] where authors unveiled the relevance of understanding efficient ways to handle these complex data types. The study examines critical challenges, including (1) complex implicit spatiotemporal relationships, (2) interdisciplinarity demands requiring fusion of multi-domain expertise and data, and (3) discretization hurdles arising from scale, time-zone biases, and heterogeneous/dynamic data properties. It further identifies gaps in prior research on such data. A survey on the use of spatiotemporal data applied to traffic prediction is presented in [39] where spatiotemporal data is portrayed as potentially robust to address traffic-related issues when combined with advanced processing techniques like machine learning. After acknowledging the challenge of modeling spatiotemporal dependencies to improve the forecasting capabilities of algorithms despite the large presence of this data type, in [23] they proposed a novel approach to achieve this task based on graph neural networks architecture under the constraint of cross-node federated learning. This technique could help improve prediction accuracy when applied to traffic flow forecasting. A similar approach is presented in [20] where the need to capture dependencies within spatiotemporal data was addressed using graph neural networks to improve prediction accuracy. While this task was complex considering the nature of the data, capturing existing relationships between spatiotemporal features appears to be relevant for advancement in the domain of predictive modeling.

Across diverse domains, causal machine learning and spatiotemporal data analysis have demonstrated their utility in uncovering complex relationships, supporting evidence-based decision-making, and addressing critical societal challenges. While existing studies have yielded valuable insights, they also reveal persistent methodological challenges requiring further investigation. More importantly, current research has not systematically explored the unique opportunities and challenges emerging from the convergence of causal machine learning and spatiotemporal analysis. In Section V, we conduct a comprehensive examination of these underexplored aspects, specifically, the emerging opportunities and challenges at the intersection of these methodologies, alongside with strategic directions for future research.

Table 1 depicts major observations from the samples of applications considered.

The synthesis of studies presented in Table 1 reveals significant advances in the integration oftion of machine learning with causal inference techniques, demonstrating their comparative advantages over traditional machine learning approaches. Incorporation of causal frameworks into spatiotemporal data analysis shows particular promise in overcoming the limitations of conventional pattern recognition methods. However, we could identify several persistent challenges, such as the insufficient theoretical development of the causal ML framework, the limited comparative benchmarking against established methods, existing difficulties in addressing complex confounding structures, and the problematic assumption of sparsity without validation mechanisms. Current research indicates that while these approaches are promising, many implementations remain inadequate and fail to incorporate state-of-the-art methodologies. This underscores the critical need for further theoretical development and empirical validation. Notably, the application of causal ML in spatiotemporal analysis holds special significance for developing regions, where it could substantially enhance decision-making in urban infrastructure planning, agricultural productivity optimization, or disaster preparedness and response systems. These potential applications highlight both the timeliness and importance of our current investigation into this emerging interdisciplinary field.

IV. METHODOLOGY

A. DATA COLLECTION

We retrieved papers from five scientific platforms: Google Scholar, Mendeley, Springer Nature, IEEE, and UCL Library Services. We used keywords such as "causal machine learning", "spatiotemporal data analysis", and the combinations "causal machine learning AND spatiotemporal data analysis" and "causal machine learning & spatiotemporal data analysis". These keywords were enclosed in quotation marks to avoid irrelevant results. The search spanned from 2014 to 2025, reflecting the recent development of causal machine learning. We focused on peer-reviewed papers (reviews and empirical papers), books, or book sections to ensure confidence in the results. This approach helped us retrieve relevant publications in these fields separately and together, showcasing their individual and converged advances.

B. RETRIEVED PUBLICATIONS

Table 2 summarizes the search result explained in Subsection IV-A.

Based on the criteria explained in IV-A, UCL Library Services retrieved the largest number of publications (49,709) for all keywords, followed by IEEE (2,599), Google Scholar (340), Mendeley (216) and Springer Nature (143). These insights highlight the growing importancend wide-ranging applications of causal machine learning and spatiotemporal data analysis. The multifaceted and interdisciplinary application of these methods is making substantial contributions to computer science, environmental studies, and geography. The variety of journals and subjects underscores its broad relevance and applicability across different domains in contemporary research. More details of the retrieved result are provided in Appendix A. Interestingly, when

TABLE 1. Summary of Samples considered, including case studies, key finding, model specification, and limitations.

Author	Case study	Key finding	Model specifications	Metrics	Limitations
S. Ali and J. Wang, 2022 [5]	Review of Causal Machine learning	Taxonomy of Causal ML applications	Not Covered	Not Covered	Highlight problems without possible solutions
Atul Rawal et al. 2022 [37]	Review of Causality frameworks	Challenges in AI/ML causality include lack of data sets, ground truth, standardized definitions, evaluation metrics, balance between causality and performance, and causal DL models	Not covered	Distances between predicted causal graphs and ground truth using observational data, ROC, TPR, Precision, Recall, F-1, F-test, and MSE	Emphasize problems but do not offer solutions.
S. Arti et al., 2020 [8]	Identification of generally used Causal ML approaches	Bayesian Networks are commonly used in the context of causality,and the propensity score is the most extensively used metrics in causality research	Bayesian Networks, Random Forest, XGBoost, SVR, Linear and Logistic Regression, Optimal Discriminant Analysis, Ordinary Least Square Lagrangian, Markov	Confusion Matrix, Accuracy, Recall, Precision, F-1 measures, Standard deviation	Limited insight on Causal ML frameworks
F. Wang et al., 2024 [7]	Soybean Yields estimation	Improved performance in combining deep learning with structural causal model	SCM-GAT based on causal relationships	R^2, RMSE, rRMSE	Limited confounding
A. Baiardi and A. A. Naghi, 2024 [9]	Application of Causal ML on Econometric literature	Superiotity of Causal ML methods to traditional in handling confounders, capturing effects, and conducting robustness checks. Ideal for studies with uncertain covariate influences and for identifying heterogeneity	Double Machine Learning, Causal Forest	Magnitude, standard error	Assumption of sparsity which is not testable
A. Hamd et al., 2022 [18]	Survey on open challenges of spatiotemporal data mining	Huge potential of causality in spatiotemporal data analysis	Not covered	Not covered	Limited insight on causlity and causal ML
C. Meng, S. Rambhatla, and Y. Liu, 2021 [23]	Application of Graph Neural Network for Spatio-Temporal Data Modeling	Improved data modeling and forecasting performance in combining federated learning with graph neural networks	Cross-Node Federated Graph Neural Network (CNFGNN)	RMSE	Limited benchmark
G. Jin et al., 2024 [20]	Predictive learning scenarios in urban computing	Improved data modeling and forecasting performance in combining federated learning with graph neural networks	Spatio-Temporal Graph Neural Network (STGNN)	Not covered	State-of-art for implementation not provided

combining the search terms 'causal machine learning' and/or & 'spatiotemporal data analysis' (with quotation marks used for all searches except those conducted in IEEE), the results were limited. In fact, the IEEE retrieved only 12 publications (3 journals and 9 conference proceedings), and UCL Library Services produced 61 publications (13 books and 48 peer-reviewed papers). This observation suggests that while both fields (causal machine learning and spatiotemporal data analysis) are well-researched individually, their intersection may represent a relatively nascent area of research with significant potential for growth and innovation. This finding highlights the need for further research to bridge this gap and advance the field of causal inference in spatiotemporal contexts. Figure 1 summarizes the overall process.

The results of our systematic search process, illustrated in Figure 1 reveal a limited convergence between causal machine learning and spatiotemporal analysis in peer-reviewed literature. Applying our rigorous selection criteria, we found that fewer than 1% of initially identified publications met our inclusion standards. Although substantial discussion exists in gray literature (blogs, web forums) and conference proceedings, the peer-reviewed scientific literature on this intersection remains remarkably sparse, indicating a nascent field requiring further exploration. To maintain methodological rigor,

TABLE 2. Summary of retrieved publications by keywords and sources.

Keywords	Source	Retrieved publications	Observations
Causal Machine learning	Google Scholar	258	Not specified
	Mendeley	138	155 peer-reviewed papers, 3 books section
	Springer Nature	61	All peer-reviewed articles
	IEEE	388	379 peer-reviewed, 9 books
	UCL Library Services	8,114	Books and peer-reviewed articles from all the resource
Spatio-temporal data analysis	Google Scholar	82	Peer-reviewed articles
	Mendeley	78	71 peer-reviewed articles, 1 book, 6 book section
	Springer Nature	82	Peer-reviewed articles
	IEEE	2,208	2,201 peer-reviewed articles, 7 books
	UCL Library Services	41,534	Books and peer-reviewed articles from all the resource
Causal machine learning AND spatiotemporal data analysis	IEEE	3	12 retrieved (3 peer-reviewed articles, 9 conference proceedings) but only peer-reviewed considered for analysis
	UCL Library Services	61	13 books, 48 peer-reviewed articles

TABLE 3. Summary of the 51 selected publications with key contributions.

Aspects	Key contributions	References
Enhancing prediction accuracy	Causal machine learning techniques improve the prediction accuracy of spatiotemporal data by identifying hidden causal effects within these complex data types	[48], [96]–[99], [101]–[114]
Overcoming pattern recognition limitations	These techniques are combined with existing machine learning and deep learning algorithms to address the limitations of pattern recognition	[96], [97], [99], [112], [114], [115], [117], [118], [147]
Improving interpretability	Causality discovery is utilized to enhance the interpretability of complex machine learning algorithms, such as using graphs to visualize potential connections between features	[119]
Various applications of Causal ML	Causal machine learning techniques are applied to improve causal discovery and effects estimations in domains where spatiotemporal data play a crucial role	[20], [25], [25], [104]–[107], [109], [111], [112], [114], [115], [117], [120]–[123], [125]–[132], [136], [137], [143]–[147]

we exclusively focused on peer-reviewed journal articles, excluding:

- Conference proceedings (including 9 from IEEE),
- Books from UCL Library Services (13 which contained only conference proceedings).

This selection process yielded 51 high-quality peer-reviewed articles for our analysis. Through this comprehensive review, our aim is to provide deep insight into the integration of causal machine learning with spatio-temporal data analysis.

C. ANALYSIS OF THE CONVERGENCE

Through an in-depth review of the retrieved publications, focusing on peer-reviewed papers as the books were conference proceedings, we uncovered the main areas where causal machine learning has been utilized for spatiotemporal data analysis. Table 3 summarizes these findings.

As illustrated in Fig. 2, the methodological focus of surveyed studies reveals distinct research priorities:

- **Causal discovery and effect estimation** (32 studies): The predominant application of causal machine learning techniques, emphasizing the identification and quantification of causal relationships in spatiotemporal data;
- **Prediction accuracy enhancement** (19 studies): Secondary focus on improving predictive performance while maintaining causal validity;
- **Pattern recognition limitations** (10 studies): Approaches addressing the constraints of conventional pattern detection methods in causal contexts;
- **Interpretability improvement** (1 study): The least explored area, focusing on making complex causal models more transparent when applied to spatiotemporal data.

This distribution highlights a significant research gap in developing interpretable causal machine learning frameworks for spatiotemporal analysis.

These studies address multiple dimensions of causal machine learning in spatiotemporal analysis, explaining why we reference 62 methodological contributions while analyzing 51 distinct papers. While these works demonstrate significant advances, they systematically fail to:

1) Fully leverage specialized techniques that address unique spatiotemporal challenges, particularly:
   - Spatiotemporal Granger causality for dynamic systems,
   - Structural equation modeling with spatial lags,

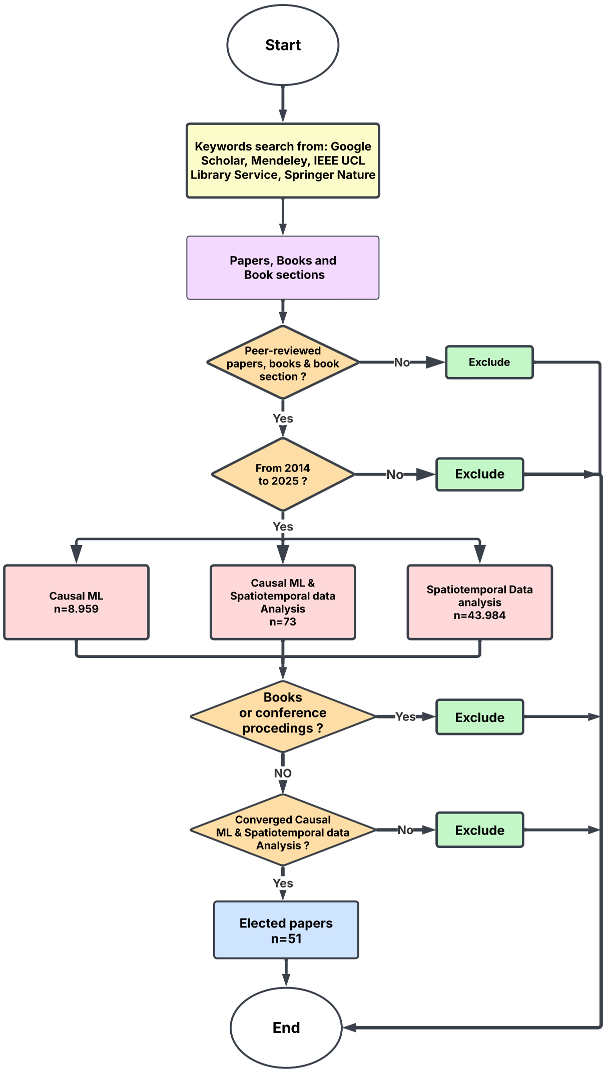


FIGURE 1. Summary of data collection process.

- Causal graph analysis for complex interdependencies.

2) Provide comprehensive guidelines for implementing these methods to solve core challenges:
   - Capturing spatial autocorrelation effects,
   - Modeling temporal dependencies in causal structures,
   - Maintaining interpretability while handling complexity.

This critical gap in the literature reveals an urgent need for methodological research focusing on these specialized approaches. In Section V, we conduct a thorough investigation of these three key techniques, examining their theoretical foundations, implementation challenges, and emerging opportunities to establish clearer guidelines for the field.

V. INTEGRATION OF CAUSAL MACHINE LEARNING WITH SPATIOTEMPORAL DATA ANALYSIS

The integration of causal machine learning with spatiotemporal analysis remains an emerging research frontier with significant untapped potential. This section provides a focused examination of three pivotal methodologies that address the distinctive challenges of causal inference in spatiotemporal domains:

- **Spatiotemporal Granger Causality**: Extends traditional temporal causality frameworks to incorporate spatial dependencies;
- **Structural Equation Modeling with Spatial Lags**: Integrates spatial autocorrelation into causal pathway analysis;
- **Causal Graph Analysis**: Captures complex interdependencies in networked spatiotemporal systems.

These techniques were systematically selected for their demonstrated capability to account for spatial autocorrelation effects, to model hierarchical temporal dependencies, to maintain interpretability despite system complexity, and to scale to real-world spatiotemporal datasets. For each method, our analysis investigates the theoretical foundations and assumptions, implementation requirements and computational considerations, current limitations, and open challenges and opportunities.

A. SPATIOTEMPORAL GRANGER CAUSALITY

This method extends the traditional Granger Causality to account for both spatial and temporal dependencies, allowing the identification of the causal relationship over time and space [46], [47]. To simplify its understanding, let us consider two time series X_t and Y_t where t represents the time. The spatiotemporal Granger causality from X to Y can be represented as shown in Eq. (1).

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

where:

- a_i and β_j are the coefficients,
- p and q are the number of lags for Y and X respectively,
- ϵ_t is the error term, which represents the portion of the dependent variable that is not explained by the model, which accounts for the randomness and noise in the data, as well as any unobserved factors that might influence the dependent variable but are not included in the model.

The idea is to test whether the coefficients of X (β_j) are significantly different from zero, which means that the past values of X provide information about future values of Y. The statistical significance of the coefficients β_j is typically assessed through:

1) Wald tests: Examining whether the estimated coefficients differ significantly from zero [40], with the test statistic represented in Eq. 2:

$$W = \hat{\beta}^T \Sigma^{-1} \hat{\beta}, \tag{2}$$

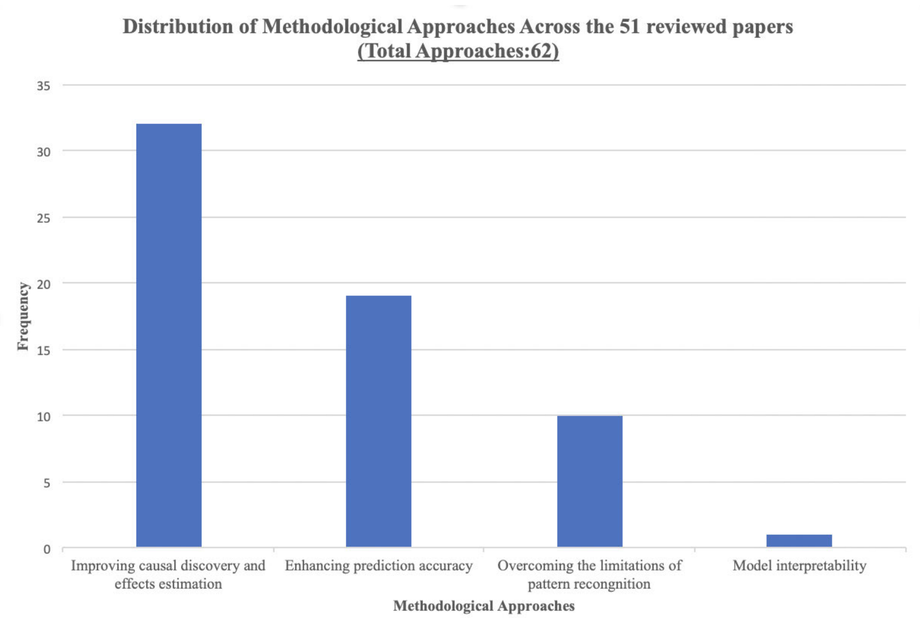


FIGURE 2. Distribution of methodological approaches.

where:
- $\hat{\beta}$ is the vector of estimated regression coefficients (e.g., $\hat{\beta} = [\hat{\beta}_1, \hat{\beta}_2, \ldots, \hat{\beta}_p]^T$),
- Σ is the covariance matrix of $\hat{\beta}$, capturing the variances and covariances of the coefficient estimates,
- Σ^{-1} is the inverse of the covariance matrix, used to standardize the quadratic form,
- W follows a χ^2 distribution with p degrees of freedom (where p = number of coefficients) under the null hypothesis.

2) F-tests: Comparing the residual sum of squares between restricted (without X terms) and unrestricted models [41], calculated as shown in Eq. (3):

$$F = \frac{(RSS_R - RSS_U)/q}{RSS_U/(T - p - q)}, \tag{3}$$

where:
- T is the sample size,
- RSS_R and RSS_U are the restricted and unrestricted residual sums of squares,
- q the number restrictions, and,
- p the number of predictors in the unrestricted model.

3) Likelihood ratio tests: Evaluating the difference in log-likelihoods between nested models with and without the causal terms [42].

These approaches test whether the past values X provide statistically significant information about future values Y beyond what is contained in Y's own history. This rigorous testing framework provides a more accurate and reliable measure of causality, resulting in its extensive application in several fields [43].

B. STRUCTURAL EQUATION MODELING WITH SPATIAL LAGS

This technique integrates spatial dependencies directly into structural equation models, giving a comprehensive view of causal relationships while accounting for spatial autocorrelation [44], [45]. It is represented in Eq. (4) as:

$$Y_i = \rho \sum_{j \neq i} W_{ij} Y_j + \beta X_i + \epsilon_i, \tag{4}$$

where, considering two regions i and j:
- Y_i is the dependent variable for region i,
- ρ is the spatial autoregressive parameter,
- W_{ij} is the spatial weight matrix element between regions i and j,
- X_i is the vector of independent variables for region i,
- β is the vector of coefficients for the independent variables,
- ϵ_i is the error term for the region$i$, which accounts for the discrepancies between unobserved values and the values predicted by the model for each region.

This approach captures the influence of neighboring regions on the dependent variable Y_i through the spatial lag term W_{ij}. This is crucial when the need to uncover causal relationships in consideration of spatial autocorrelation arises.

C. CAUSAL GRAPH

Utilizing graphical models to depict and analyze causal relationships, this technique, like directed acyclic graphs (DAGs), can help visualize and infer causal connections in spatiotemporal data [48]. This can be easily explained by:

$$X \rightarrow Y \quad \text{and} \quad Z \rightarrow Y,$$

three variables X, Y, and Z where X influences Y and both X and Z influence Y. In such a scenario, Y is a function of both X and Z, which can be represented mathematically as shown in Eq. (5):

$$Y = f(X, Z) + \epsilon, \tag{5}$$

where:

- Y is the outcome variable,
- X and Z are the predictor variables,
- f is a function that describes the relationship between Y, X, and Z,
- ϵ is the error term representing the unobserved factors affecting Y.

D. REAL-WORLD APPLICATIONS EXAMPLES

Some research has successfully converged these two domains to overcome existing limitations. Table 4 provides a sample of these examples, highlighting the aspect improved.

In urban planning applications, [96] employed causal machine learning to analyze traffic patterns and optimize public transportation routes. Their causal inference approach successfully predicted traffic congestion in metropolitan areas, leading to more efficient urban development strategies.

Causal machine learning has proven instrumental in epidemiological studies, particularly in analyzing disease transmission dynamics [99]. By employing causal discovery techniques enhanced with machine learning, researchers have successfully identified critical environmental factors influencing disease outbreaks, enabling more targeted and effective public health interventions.

The integration of causal ML with deep learning algorithms has significantly improved the accuracy of spatiotemporal predictions in climate science, as demonstrated by [119]

These case studies demonstrate the capability of causal machine learning to address persistent challenges in complex spatiotemporal systems. The methodology shows particular promise in overcoming the limitations of traditional correlation-based analysis, the difficulties in accounting for spatial and temporal confounders, and existing challenges in deriving actionable insights from high-dimensional datasets.

E. LIMITATIONS OF SPATIOTEMPORAL GRANGER CAUSALITY, STRUCTURAL EQUATION MODELING WITH SPATIAL LAGS, AND CAUSAL GRAPH WHEN APPLIED ON SPATIOTEMPORAL ANALYSIS

While these methods offer powerful tools for analyzing spatiotemporal data, they come with certain limitations:

1) SPATIOTEMPORAL GRANGER CAUSALITY

The application of Granger causality to spatiotemporal analysis faces several inherent constraints. First, the method's foundational assumption of linear relationships between variables often proves inadequate for modeling complex, nonlinear system dynamics prevalent in real-world scenarios. A second critical limitation arises in temporal lag specification (selecting suboptimal time windows) that can generate spurious causal inferences, yet determining appropriate lags remains theoretically and computationally challenging. While the framework effectively captures temporal dependencies, its capacity to represent spatial interactions is often limited, particularly in high-dimensional datasets where cross-location dependencies exhibit complex patterns. Most significantly, the approach remains vulnerable to confounding bias, as unaccounted spatial or temporal confounders may systematically distort inferred causal relationships [36].

2) STRUCTURAL EQUATION MODELING (SEM) WITH SPATIAL LAGS

Developing accurate structural equation models (SEM) for spatiotemporal systems presents several methodological hurdles. First, model specification demands rigorous attention to both spatial dependence structures and temporal dynamics, requiring extensive validation to ensure theoretical fidelity. A critical issue is model identification (the problem of obtaining unique parameter estimates becomes increasingly acute in high-dimensional systems with numerous interacting variables). Measurement error further compounds these challenges, as even minor inaccuracies in observed variables can propagate through the model, systematically biasing causal effect estimates. Additionally, like spatiotemporal Granger causality, conventional SEM implementations typically rely on linearity assumptions that often fail to capture the nonlinear interactions prevalent in real-world systems [95].

3) CAUSAL GRAPHS

First, reliably inferring the correct causal structure from observational data remains problematic, particularly when dealing with confounding variables and intricate interaction patterns. A second major limitation stems from missing data, which frequently introduces bias into causal effect estimates. Perhaps most critically, the potential presence of unmeasured confounders (variables influencing both cause and effect) can systematically distort results. Furthermore, as system complexity increases, these models often become computationally intensive while simultaneously losing interpretability, creating validation difficulties, especially in large-scale applications [48]. These challenges are particularly acute in spatiotemporal contexts where dynamic dependencies and spatial correlations further complicate causal structure learning.

While the foundational assumptions of these causal techniques may not universally hold for complex spatiotemporal systems, they nevertheless establish essential theoretical and computational frameworks for causal analysis. The limitations discussed (ranging from linearity constraints in Granger causality to identification challenges in spatial SEM) highlight critical gaps that, when addressed, could significantly advance the field. Rather than undermining their utility, these challenges demarcate precisely where innovative integrations with causal machine learning could yield transformative improvements. Specifically, they reveal

TABLE 4. Summary of real-world applications of Causal ML to Spatiotemporal data analysis.

Aspects	Key Consideration	Case Study/Example	References
Enhancing Prediction Accuracy	Causal ML improves prediction accuracy by identifying hidden causal effects.	Used in urban planning and health to predict traffic congestion, optimize road networks and improve Parkinson's disease prediction.	[96], [97]
Overcoming Pattern Recognition Limitations	Causal ML addresses limitations of pattern recognition in complex datasets.	Applied in epidemiology to identify causal relationships between environmental factors and disease spread.	[99], [112]
Improving Interpretability	Causal ML enhances interpretability by visualizing connections between features.	Used in climate science to model the impact of CO_2 emissions on global temperatures.	[119]
Improving Causal discovery	Causal ML is applied to improve causal discovery	Optimizing public transportation routes.	[104], [106]
Enhancing Causal effect estimation	Application of Causal to improve effects estimations	Various applications in domains such as urban planning and in predicting climate change impacts	[120]

opportunities to develop novel combination methods that can leverage the strength of each technique to address the limitations mentioned or to consider recent advances in machine learning to architect hybrid models for the same purpose. We systematically explore these research frontiers in Section VI, focusing on methodological innovations that could unlock the full analytical potential of spatiotemporal causal inference.

VI. OPPORTUNITIES AND CHALLENGES

A. OPPORTUNITIES

Among the existing possibilities that can be considered, based on the subjects covered previously, we present some key opportunities to unlock the potential of spatiotemporal data analysis with causal machine learning:

1) COMBINING METHODS

- **Synergy of techniques**: The combination of multiple methods can strengthen the reliability and validity of causal inference results. For instance, using Granger causality to identify potential causal relationships and then employing structural equation modeling (SEM) or causal graphs to model the underlying causal structure can provide a more comprehensive understanding of a system despite its complexity [48].
- **Simulation of combining Granger Causality and SEM**: For two timeseries X_t and Y_t, where X_t Granger-causes Y_t, identifying potential causal relationships (Granger-causality) can be achieved as shown in Eq. (6):

$$Y_t = \alpha + \sum_{i=1}^{p} \beta_i Y_{t-i} + \sum_{j=1}^{q} \gamma_j X_{t-j} + \epsilon_t, \tag{6}$$

where:
  - α, β_i, and γ_j are coefficients,
  - ϵ_t is the error term,
  - p and q are the lag orders for X_t and, Y_t respectively.

Given the Granger-causality model results, the model underlying causal structure, using structural equation modeling (SEM) to account for direct and indirect effects, can be represented as shown in Eqs. (7)–(8):

$$Y_t = \lambda_1 X_t + \lambda_2 Z_t + \delta, \tag{7}$$
$$Z_t = \theta_1 X_t + \epsilon_z. \tag{8}$$

where:
  - λ_1, λ_2, and θ_1 are the SEM coefficients,
  - δ and ϵ_z are the error term vectors for Y_t and Z_t, respectively.

The integration and synergy of Eqs. (6)–(7)–(8) can form the comprehensive model represented in Eq. (9):

$$Y_t = \lambda_1 \left(\alpha + \sum_{j=1}^{q} \gamma_j X_{t-j} \right) + \lambda_2 Z_t + \delta. \tag{9}$$

This integration offers a holistic approach to causal analysis, leveraging the strengths of both methods. Granger causality identifies potential causal relationships based on the temporal order of events, while structural equation modeling (SEM) captures complex causal structures, including both direct and indirect effects. By combining these methods, it is possible to enhance the reliability of causal inferences through the predictive power of past values and the consideration of broader causal pathways. This comprehensive understanding of the system addresses the limitations of each method individually, resulting in a robust analysis of causal relationships. It is important to mention that the integration of Granger causality, structural equation modeling (SEM) with spatial lags, and causal graphs represents a deliberate methodological strategy, where each technique addresses distinct yet complementary dimensions of spatiotemporal causal analysis. This structured combination follows a logical hierarchy of causal inference. Granger causality serves as the foundational temporal test, answering whether variable X systematically precedes variable Y in time. While powerful for establishing predictive temporal ordering (a necessary condition for causality), it cannot alone distinguish true causation from indirect or spurious relationships that might arise from unmeasured confounding variables. Building upon this temporal foundation, SEM with spatial lags introduces crucial spatial context by simultaneously quantifying both time-lagged effects and contemporaneous spatial dependencies. This proves particularly valuable for phenomena like environmental pollution, where effects may spill across geographic boundaries. However, SEM's reliance on predefined linear structural equations risks model misspecification without proper constraints. This is where causal graphs provide essential theoretical grounding, mapping plausible causal pathways based on domain knowledge while excluding impossible or illogical relationships. For instance, they might prevent statistically significant but physically impossible links like rainfall directly affecting stock market prices without an agricultural intermediary mechanism. While causal graphs excel at encoding structural assumptions, they lack the granularity to quantify temporal or spatial effect sizes alone. Together, these methods form a robust analytical chain: Granger establishes temporal precedence, SEM with spatial lags quantifies the multidimensional dependencies, and causal graphs maintain theoretical consistency. This integration systematically addresses the key challenges of spatiotemporal analysis (temporal dynamics, spatial interactions, and structural plausibility) while compensating for each method's individual limitations through their combined strengths.

2) ADVANCED STATISTICAL TECHNIQUES

- **Addressing confounding and selection bias**: Techniques such as propensity score matching and instrumental variable analysis, though well-established in other contexts, are emerging as potential methods to tackle the complexity of spatiotemporal data. Propensity score matching helps balance confounding variables between treated and control groups, simulating randomized experiments. In mathematical terms, it can be represented as shown in Eq. (10):

$$\hat{e}(X) = P(T = 1|X), \tag{10}$$

where:

- $\hat{e}(X)$ is the estimated propensity score,
- T is the treatment indicator, and,
- X is the observed covariates.

By balancing the distribution of observed confounding variables between the treatment and control groups, which reduces bias in estimating the treatment effect, and combining it with instrumental variable analysis, which addresses unobserved confounding by using an instrument (for instance Z), that is correlated with the treatment T but uncorrelated with the outcome Y except through T, can help mitigate the impact of confounding variables and selection bias, leading to more accurate causal estimates [49], [50].

Indeed, instrumental variable analysis involves two stages: The first stage consists of predicting the treatment and is presented in Eq. (11):

$$T = \pi_0 + \pi_1 Z + \pi_2 X + \epsilon, \tag{11}$$

where:

- T is the treatment or intervention variable, which indicates whether or not the subject received treatment,
- π_0 is the intercept term, which represents the baseline level of the treatment when all predictors are zero,
- π_1 is the coefficient for the instrumental variable (Z), which measures the association between the instrumental variable and the treatment,
- Z is the instrumental variable, which is correlated to the treatment T but not directly with the outcome Y except through T,
- π_2 is the coefficient for the observed covariates (X), which represents the effect of the observed covariates on the treatment,
- X is the vector of the observed covariate, and,
- ϵ is the error term.

This is part of the two-stage least squares method used in instrumental variable analysis. The second stage aims to predict the outcome. It is shown in Eq. (12) as:

$$Y = \beta_0 + \beta_1 \hat{T} + \beta_2 X + \eta, \tag{12}$$

where:

- Y is the outcome variable, the dependent variable that is being influenced by treatment,
- β_0 is the intercept term, which represents the baseline level of the outcome when all predictors are zero,
- β_1 is the coefficient for the predicted treatment ($\hat{T}$) which measures the causal effect of the treatment on the outcome,
- $\hat{T}$ is the predicted treatment variable from the first stage. It is the fitted value of the treatment estimated using the instrumental variable Z,
- β_2 is the coefficient for the observed covariates (X) which represents the effect of the observed covariates on the outcome,
- X is the vector of the observed covariates, and,
- η is the error term, which accounts for the variability in the outcome not explained by the treatment and observed covariates.

In the two-stage least squares method, this latter stage equation helps estimate the causal effect of the treatment on the outcome using the predicted treatment ($\hat{T}$) from the first stage. This helps address hidden biases and confounding variables, leading to more accurate causal estimates.

- **Handling complex dependencies**: Advanced statistical methods, such as those used in spatial and temporal econometrics, can account for spatial autocorrelation, temporal dependence, and other complex relationships in spatiotemporal data [51], [52]. For instance, incorporating Spatiotemporal Autoregressive Models (STAR) into causal machine learning for this purpose can be explained as follows:

Having Y_t as the dependent variable representing the outcome for each time period t, and the covariates $X_t = \{X_{t1}, X_{t2}, \ldots, X_{tp}\}$, where X_t represents the covariates at time t. The spatial weight matrix W can be represented as in Eq. (13):

$$W = \begin{pmatrix} W_{11} & W_{12} & \cdots & W_{1n} \\ \vdots & & \ddots & \vdots \\ W_{n1} & W_{n2} & \cdots & W_{nn} \end{pmatrix}, \tag{13}$$

where each element W_{ij} denotes the spatial influence between locations i and j.

The STAR model can be represented as shown in Eq. (14):

$$Y_t = \phi W Y_{t-1} + X_t \beta + \mu T Y_{t-1} + \epsilon_t, \tag{14}$$

where:

- Y_t is the dependent variable at time t,
- ϕ is the spatial autoregressive coefficient,
- WY_{t-1} is the spatially lagged dependent variable from the previous time period,
- X_t is the covariate matrix at time t,
- β is the coefficient vector for covariates,
- μ is the temporal autoregressive coefficient,
- TY_{t-1} is the temporally lagged dependent variable from the previous time period, and,
- ϵ_t is the error term.

Incorporating causal machine learning by considering the propensity score matching (10) and instrumental variable (11) and (12) to address confounding and selection bias will help to provide solid outcomes. The parameters (ϕ, β, μ) can be estimated using the maximum likelihood method shown in Eq. (15):

$$\min_{\phi,\beta,\mu} \sum_t (Y_t - \phi W Y_{t-1} - X_t \beta - \mu T Y_{t-1})^2 . \tag{15}$$

The validation of the model can be achieved by checking the residuals ϵ_t for spatial and temporal dependencies. The analysis of the estimated coefficients represented by $\hat{\phi}, \hat{\beta}, \hat{\mu}$ needs to take place to understand the influence of spatial-temporal dependencies. The insight provided by the STAR model for informed decision-making can help analyze how spatial and temporal factors influence the outcome. Assessing the impact of policies based on $\phi W Y_{t-1}$ and $\mu T Y_{t-1}$ provides a robust way to analyze the impact of the policies and interventions.

3) MACHINE LEARNING TECHNIQUES

- **Causal discovery algorithms**: Advanced algorithms like the Peter and Clark algorithm (PC algorithm), Fast Causal Inference (FCI), Additive Noise Model (ANM), DoWhy, or Causal ML can be used to discover causal relationships from observational data, even in the presence of latent variables and complex interactions [38], [53].
- **To learn complex patterns and dependencies within spatiotemporal data**, deep learning models, such as recurrent neural networks (RNNs) and convolutional neural networks (CNNs), can be used to learn complex patterns and dependencies in spatiotemporal data [55].
- Thanks to dedicated algorithms and machine learning libraries like DoWhy and Causal ML, machine learning techniques can be used to estimate counterfactual outcomes, or counterfactual inference, allowing to assess the impact of interventions [56].

4) DOMAIN EXPERTS

Domain experts provide three essential contributions to spatiotemporal data analysis. First, they decode the underlying physical, social, or environmental processes that generated the observed data (knowledge that is rarely contained in datasets alone). Second, they validate whether causal relationships identified by algorithms align with domain theory. Third, they identify subtle confounding factors (like unmeasured variables or spatial autocorrelation effects) that could distort conclusions [3], [19].

Effective integration of domain expertise requires two complementary strategies, which are:

- Collaborative modeling, which establishes continuous feedback loops where experts and data scientists jointly refine models through successive iterations (for instance, adjusting spatial weighting schemes in disease transmission models based on epidemiological knowledge) [137];
- Expert-informed feature selection, which prioritizes variables with established causal significance in the domain (like soil permeability in flood models) while deprioritizing statistically prominent but theoretically irrelevant features [138], [139].

This fusion of empirical analysis with domain knowledge serves to improve model accuracy by grounding algorithms in real-world mechanisms and enhance practical utility by ensuring outputs align with decision-making frameworks in fields like emergency response or ecosystem management.

B. KEY CHALLENGES IN SPATIOTEMPORAL CAUSAL LEARNING

Beyond the algorithmic assumptions discussed in Section V, the intrinsic properties of spatiotemporal data combined with methodological constraints present significant hurdles for causal machine learning applications. These challenges emerge from both the complex nature of spatiotemporal phenomena and the technical limitations of current analytical approaches.

1) TECHNICAL IMPLEMENTATION CHALLENGES

- **Spatiotemporal encoding:** Developing effective representations for spatiotemporal interactions remains a fundamental challenge. The dynamic coupling of spatial and temporal patterns (which often evolve at different scales) requires specialized encoding approaches. Irregularities such as non-uniform sampling intervals, heterogeneous spatial grids, and asynchronous measurements further complicate the development of unified representations that preserve critical dependencies while remaining computationally tractable [57].

- **Data sparsity and missingness:** Spatiotemporal datasets frequently suffer from incomplete observations across both dimensions. The missingness mechanisms, whether Missing Completely at Random (MCAR), Missing at Random (MAR), or Missing Not at Random (MNAR), each require distinct handling strategies. Geographic and temporal imbalances in data coverage can introduce substantial biases, particularly when machine learning models trained on such data generalize poorly to underrepresented regions or time periods [70], [71].
- **Measurement noise and confounding:** Multiple noise sources degrade spatiotemporal data quality, including:
  - * Instrumentation errors from sensing devices,
  - * Environmental interference factors,
  - * Human annotation inconsistencies.

More critically, unmeasured spatial or temporal confounders that simultaneously affect treatment and outcome variables can systematically distort causal estimates, potentially leading to invalid conclusions [18], [73].

2) METHODOLOGICAL ISSUES

- **Ensuring generalization in spatiotemporal causal models**: Selecting appropriate models for spatiotemporal causal inference requires addressing two fundamental challenges beyond simple sample size augmentation. First, the inherent structural complexity of spatiotemporal data (where spatial autocorrelation violates IID assumptions and temporal dependencies induce non-stationarity) necessitates specialized regularization approaches. Second, the causal specificity demands preservation of identifiable structural relationships beyond mere predictive performance. A model that becomes overly complex may fit noise rather than underlying causal patterns [74], leading to:
  - Spurious causal edges in learned graphs,
  - Biased treatment effect estimates,
  - Poor generalization to new spatiotemporal contexts.

  Effective mitigation strategies must account for these unique aspects:
  - *Structural regularization*: Spatial Graph Laplacian Penalties ($\mathcal{L}_{spatial} = \lambda_s \mathbf{f}^T \mathbf{L}\mathbf{f}$) and temporal smoothness constraints,
  - *Causal-aware validation*: Spatiotemporal block cross-validation to maintain dependency structures [24],
  - *Architecture selection*: Prioritizing interpretable causal architectures over opaque alternatives.

  These approaches, combined with careful hyperparameter optimization, are essential for obtaining reliable causal estimates [3]. Merely increasing sample size cannot resolve the fundamental identifiability challenges posed by spatiotemporal dependencies.
- **Handling biases**: Biases can occur during data collection, analysis, and interpretation of the analysis [75]. Considering the scope of this research, we will focus only on biases related to data collection. Thus, we have the following biases:
  - * **Confounding bias**: This happens when an outside factor (confounder) influences both the independent variable (the supposed cause) and the dependent variable (the supposed effect), leading to a spurious association. In other words, the observed relationship might not be due to the direct link between the two variables but rather to this external factor. Missing to measure or identify this external factor can bias causal estimates. This challenge can be addressed through sensitivity analysis, which will help to assess the impact of the unmeasured confounders on the estimated causal effects [76], [77].
  - * **Selection bias**: It occurs when the sample chosen for study is not representative of the larger population that needs to be understood [78]. This is critical since it can lead to incorrect or skewed conclusions since the sample does not accurately reflect the diversity and characteristics of the entire population. Two major types of biases can be identified in the selection bias. When the participants in the study are not randomly selected, certain groups may be overrepresented or underrepresented. Such selection bias is called non-random sampling, and when participants drop out during the study at different rates across treatment and control groups, we have the attrition bias, which can lead to a biased result of the analysis [79]. To minimize their effects, random sampling methods like simple random sampling, stratified random sampling, cluster random sampling, and adaptive sampling, coupled with some follow-up analysis like kriging, Bayesian inference, and Markov Chain Monte Carlo, can be considered to account for any non-random sampling effects and potential attrition [80].
  - * **Measurement bias**: Also known as information bias, occurs when there are errors in the way data is collected, leading to systematic inaccuracies. These errors have the potential to skew the result of the analysis, making the finding unreliable [81]. There are several types of measurement biases, among which we have the systematic errors, which are inaccurate measurement instruments or procedures in the same way. Social Desirability Bias arises when participants give answers they think are more

socially acceptable, instead of their true thoughts or actions. Observer bias happens when a researcher's expectations or prior knowledge skew the measurement process [82]. These errors can be addressed by providing standardized procedures, blinding (keeping participants and observers unaware of key aspects of the study), or some appropriate training related to the domain [84].

3) INTERPRETABILITY CHALLENGES

The interpretability of causal machine learning models for spatiotemporal data faces five core challenges, which are:

- **Inherent model complexity**: The complex architectures of modern machine learning models make causal interpretation difficult [147],
- **Multidimensional spatiotemporal dependencies**: Capturing and explaining dependencies across both space and time adds layers of complexity to causal analysis [148],
- **Dynamic counterfactual reasoning requirements**: Spatiotemporal systems require counterfactual analysis that evolves over time and space [31],
- **Cross-dimensional feature interactions**: Interactions between spatial, temporal, and other feature dimensions complicate causal attribution [151],
- **Visualization difficulties for causal dynamics**: Representing causal relationships in spatiotemporal systems poses unique visualization challenges [85], [152].

To address these challenges, hybrid interpretability methods emerge as a particularly promising solution. These approaches combine complementary techniques to leverage their respective strengths while mitigating individual limitations. For instance, the combination of causal graphs with Shapley values integration, where the structural analysis presented by causal graphs (DAGs) provides the qualitative framework for understanding relationships between variables and the quantitative assessment through Shapley values, offers precise attribution of feature contributions within this causal structure. This synergy potentially enables both directional (graph) and magnitude (Shapley) interpretation of causal effects. Another notable example is the combination of attention mechanisms with feature importance ranking, which offers a dynamic focus based on the attention weights identified as relevant to spatiotemporal regions [149] and feature validation through importance ranking, which provides statistical grounding for attention patterns [151]. This dual approach could have the potential to reduce false positives in identified causal relationships. These examples could potentially address core challenges by reducing model complexity through structured causal representations, explicitly modeling spatiotemporal dependencies via attention mechanisms, supporting dynamic counterfactuals through graph-based interventions, disentangling cross-dimensional interactions via layered explanations, and enabling visualization through hierarchically organized interpretability outputs.

The above-mentioned interpretability approaches motivate six critical research questions that bridge theoretical foundations, methodological integration, and practical evaluation. These questions are:

1) To what extent does integrating causal graphs (structural causality) with Shapley values (quantitative attribution) improve both the explanatory depth (e.g., counterfactual reasoning) and attribution accuracy of feature importance in spatiotemporal models, compared to isolated interpretability methods? (*Explanatory synergy*) [31], [150];
2) Can the joint application of causal graphs and Shapley values uncover higher-order interactions or hidden confounders in spatiotemporal data that are missed by conventional feature importance methods (e.g., permutation tests or gradient-based saliency)? (*Latent relationship discovery*) [151];
3) What computational and interpretability trade-offs arise when combining these methods? (*Performance trade-offs*). For instance:
   - Does the increased explanatory power of hybrid approaches come at the cost of higher computational complexity or reduced scalability?
   - How robust are these methods to noise or missing data in spatiotemporal settings?
4) How can attention mechanisms (dynamic focus) and feature importance ranking (global significance) be jointly optimized to resolve conflicts when attention weights and importance scores disagree and provide hierarchical explanations (e.g., spatial attention + temporal Shapley trends) for neural networks processing spatiotemporal data? (*Attention-feature alignment*) [149];
5) What technical architectures (e.g., modular pipelines, end-to-end frameworks) most effectively combine these techniques to handle non-stationary spatiotemporal dependencies and scale to high-dimensional inputs (e.g., satellite imagery, urban mobility graphs)? (*Implementation challenge*);
6) What evaluation frameworks (e.g., user studies, domain expert reviews) can assess whether these hybrid methods genuinely improve model transparency for non-technical stakeholders and enable actionable insights in real-world applications (e.g., climate modeling, epidemiological forecasting)? (*Stakeholder-centric evaluation*) [152].

Addressing these questions systematically will advance interpretability solutions as model complexity increases [3].

4) COMPUTATIONAL CHALLENGES

The computational demands of causal machine learning methods for spatiotemporal analysis present significant implementation barriers. These techniques must process complex dependency structures across both spatial and temporal dimensions while handling high-dimensional datasets, creating substantial requirements for processing power and memory allocation. As demonstrated in Table 5, three key factors (time complexity, memory usage, and scalability) determine their practical viability. Proper evaluation of these metrics enables identification of performance bottlenecks, optimization of resource utilization, and ultimately determines whether these methods can be effectively deployed in real-world scenarios with constrained computational resources [4]. This computational profiling becomes particularly crucial when analyzing large-scale spatiotemporal systems where efficiency directly impacts analytical feasibility.

a: COMPUTATIONAL COMPLEXITY ANALYSIS

Time complexity measures how an algorithm's runtime scales with input size, providing crucial insights into its practical feasibility for growing datasets. Similarly, space complexity quantifies memory requirements during execution, determining hardware constraints for implementation. Together, these complexity metrics form fundamental efficiency measures for computational algorithms [133].

For causal machine learning methods to be viable in real-world applications, they must demonstrate both time and space efficiency while maintaining scalability (the capacity to handle increasing data volumes and complexity without prohibitive performance degradation). Scalability evaluation examines:

- Growth rates of computational time relative to input dimensions,
- Memory consumption patterns across data scales,
- Parallelization potential across distributed systems.

[134]

As shown in Table 5, these considerations have direct implications for:

- Hardware requirements,
- Maximum tractable problem sizes,
- Optimization strategy selection.

b: TRADE-OFFS BETWEEN INTERPRETABILITY AND SCALABILITY

- Interpretable methods Methods such as constraint-based, score-based, and Bayesian networks are highly interpretable and rigorous, making them suitable for domains where transparency is critical (e.g., healthcare, policy-making). However, their low efficiency and scalability limit their applicability to small or medium-sized datasets. These methods are best suited for problems where interpretability is prioritized over computational efficiency.
- Scalable methods Methods like machine learning and deep learning can handle large-scale, high-dimensional datasets and non-linear relationships, making them suitable for complex spatiotemporal problems. However, they are less interpretable and require significant computational resources. They are ideal for applications where scalability and predictive accuracy are more important than interpretability (e.g., climate modeling, urban planning)

c: CHALLENGES WITH HIGH-DIMENSIONAL DATA

Most methods (except deep learning-based approaches) struggle with high-dimensional data due to computational complexity or scalability issues. This suggests employing optimization strategies such as dimensionality reduction or parallel computing to make them feasible for high-dimensional spatiotemporal datasets.

d: SENSITIVITY TO MODEL ASSUMPTIONS

Methods like propensity score matching and Granger causality are sensitive to model misspecification or assumptions (e.g., linearity). A Careful validation and robustness checks are necessary to ensure reliable causal inferences.

e: DOMAIN-SPECIFIC SUITABILITY

Methods like Granger Causality are best suited for time-series data with linear relationships, machine learning (deep learning-based) methods are ideal for complex, non-linear spatiotemporal problems, and Bayesian networks are useful for problems requiring uncertainty quantification. This suggests that the choice of method should align with the specific characteristics and requirements of the problem domain.

f: COMPUTATIONAL RESOURCE REQUIREMENTS

Machine learning (deep learning-based) methods and SCMs require significant computational resources. Using them requires consideration of appropriate resource and infrastructure, such as GPUs, distributed computing, etc.

g: NEED FOR HYBRID APPROACHES

Combining the strengths of different methods, such as what is provided in this study, or using deep learning for feature extraction and constraint-based methods for causal inference, could provide a balance between scalability and interpretability. This approach opens avenues to address the limitations of individual methods.

As highlighted in Table 5, selecting appropriate causal machine learning methods requires careful consideration of interpretability, scalability, and computational demands. No single approach represents a universal solution, necessitating thorough evaluation of each method's trade-offs to ensure informed methodological choices. Addressing computational constraints through optimization and hybrid approaches enables more effective application of causal machine learning to spatiotemporal analysis. For methods challenged by high-dimensional spatiotemporal data, several optimization strategies can enhance efficiency. Dimensionality reduction techniques such as PCA, t-SNE, and autoencoders [58] help preserve essential patterns while reducing variable counts. Approximate inference methods, including variational inference and Monte Carlo sampling [59], offer

TABLE 5. Comparison of Efficiency and Scalability of the Causal ML methods presented.

Method	Efficiency	Scalability	Strengths	Weaknesses	References
Constraint-Based Methods	Low	Low	Interpretable, handles latent confounders	Exponential time complexity, struggles with high-dimensional data	[26]
Score-Based Methods	Low	Low	Flexible, can incorporate domain knowledge	High computational cost, limited scalability	[27]
Propensity Score Matching	Moderate	Moderate	Simple, widely used	Struggles with large datasets, sensitive to model misspecification	[72]
Structural Causal Models	Low	Low	Rigorous, handles complex causal relationships	High computational cost, limited scalability	[28], [141]
Machine Learning-Based Methods	High	High	Scalable, handles non-linear relationships	Requires significant computational resources, less interpretable	[67]
Granger Causality	Moderate	Moderate	Simple, effective for time-series data	Limited to linear relationships, struggles with high-dimensional data	[46], [47],
Bayesian Networks	Low	Low	Handles uncertainty, interpretable	High computational cost, limited scalability	[52], [60]

computationally tractable alternatives for Bayesian network analysis. Stochastic gradient descent [60] provides efficient optimization for deep learning implementations, while parallel computing frameworks like Spark and Dask [61] enable distributed processing. GPU acceleration [62] proves particularly valuable for intensive matrix operations in deep causal models. Specialized causal discovery algorithms, including Fast Causal Inference [63] and optimized constraint-based methods [64], demonstrate improved scalability. Spatiotemporal aggregation approaches [65] reduce dataset dimensionality while maintaining critical patterns, and model simplification strategies [66] focus analysis on key variables and interactions. Incremental learning methods [68] avoid full retraining by updating models with new data, and feature selection techniques like LASSO [69] help identify the most informative variables. Collectively, these strategies enhance the practicality of causal machine learning for complex spatiotemporal applications.

C. SUMMARY OF OPPORTUNITIES AND CHALLENGES

To summarize the opportunities and challenges discussed above, Figure 3 presents a taxonomy organized into key categories.

The taxonomy presented in Figure 3 systematically organizes both opportunities and challenges in applying causal machine learning to spatiotemporal data analysis. On the opportunity side, causal ML demonstrates significant potential for: (1) uncovering non-spurious causal relationships, (2) enhancing model interpretability through structured causal reasoning, and (3) generating actionable insights for critical domains including climate science, urban planning, and epidemiological modeling. Conversely, the framework identifies several persistent challenges, particularly the inherent complexities of spatiotemporal data (spatial autocorrelation, temporal non-stationarity), alongside methodological needs for more robust inference frameworks, rigorous ethical guidelines, and computationally efficient implementations. This structured taxonomy serves as both a research roadmap and a practical guide for implementing causal ML in spatiotemporal contexts.

VII. ETHICAL CONSIDERATIONS

The application of causal machine learning to spatiotemporal analysis introduces significant ethical considerations that demand careful attention. As noted by [18] and [19], spatiotemporal datasets frequently encode societal biases related to race, ethnicity, gender, and socioeconomic status. For example, crime rate data may reflect historical policing biases or neighborhood disparities, potentially leading to models that perpetuate or amplify these inequities through their predictions. Addressing these concerns requires a multi-faceted approach throughout the analytical pipeline. During data selection and preprocessing, researchers must actively identify and mitigate embedded biases. Model development should incorporate techniques like bias detection algorithms [75], fairness-aware machine learning, and counterfactual fairness frameworks [86] to ensure equitable outcomes across demographic groups. Beyond technical solutions, [124] emphasizes the need for broader discussions establishing ethical guidelines for technology deployment. These dialogues should focus on risk assessment, responsible use principles, and mechanisms to promote inclusive outcomes that align with societal values. Such considerations become particularly crucial when causal insights inform policy decisions affecting vulnerable populations.

To enhance the reliability of causal inferences from spatiotemporal data, advanced methods must address biases, confounding, and sensitivity in considering:

- **Bias mitigation**: Techniques like reweighting (adjusting sample weights to balance covariates), matching (pairing treated/control units with similar features), and propensity score adjustments (modeling treatment probability given covariates) can reduce selection bias [100];
- **Missing data handling**: Imputation methods such as multiple imputation (generating plausible values for missing points) and Expectation-Maximization

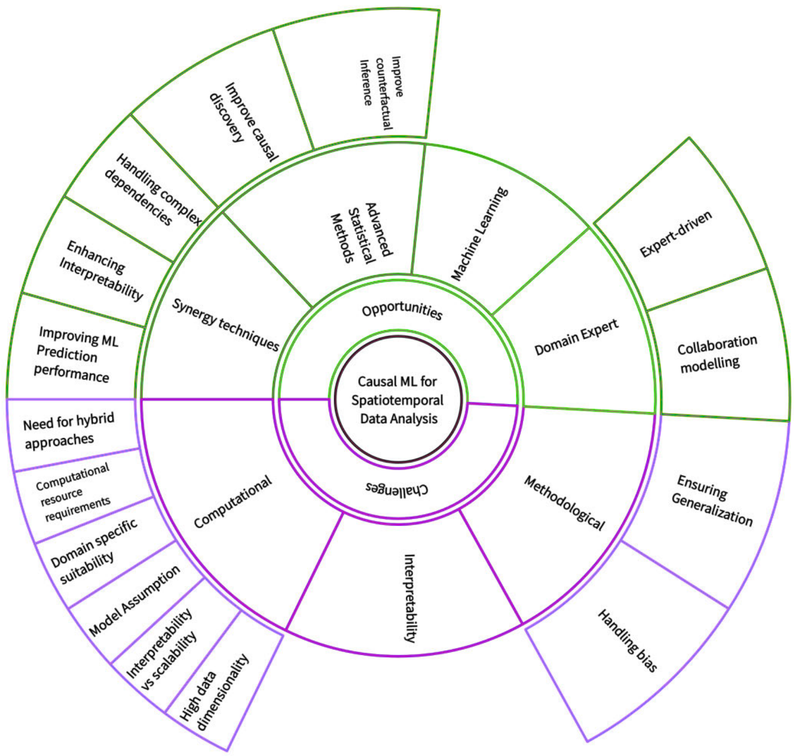


FIGURE 3. Taxonomy of opportunities and challenges.

(iterative parameter estimation) improve dataset completeness [71], [94];

- **Confounding control**: Instrumental variables (affecting outcomes only via treatment) and sensitivity analyses (quantifying unmeasured confounder impacts) strengthen causal validity [3], [141];
- **Robust Sensitivity**: E-values (minimum unmeasured confounding strength needed to nullify effects) and subgroup analyses (heterogeneous effect detection) evaluate inference robustness.

When properly utilized, these methods can improve the reliability and robustness of causal inferences using spatiotemporal data.

Furthermore, the use of ''opaque'' models in sensitive areas like public health or urban planning can have serious societal consequences because their decision-making processes are often opaque. Ensuring transparency requires interpretability techniques such as causal graphs (for visualizing relationships), Shapley values (for quantifying feature contributions), and counterfactual reasoning (for exploring ''what-if'' scenarios). These methods help reveal the underlying mechanisms driving predictions in spatiotemporal contexts, with adaptations for temporal dependencies and spatial correlations. For instance, time-varying factors can be incorporated into causal graphs, while spatial interactions can be addressed in Shapley value calculations. This transparency enables stakeholders to critically evaluate outcomes and ensure alignment with ethical standards [54], [84]. Knowing that spatiotemporal data often contain sensitive location and behavioral information, protecting them requires strong measures, including data anonymization, encryption techniques, and compliance with regulations like GDPR and CCPA [74], [87]. Also, applications like predictive policing require careful consideration of potential biases and discriminatory outcomes. Inclusive discussions with stakeholders are essential to identify risks, ensure responsible use, and promote equitable outcomes [88], [89].

These ethical considerations are crucial for ensuring that causal machine learning applications align with societal values and contribute to positive outcomes.

The integration of ethical considerations into causal machine learning for spatiotemporal analysis constitutes a fundamental requirement for responsible research and deployment. By systematically addressing transparency requirements, privacy protections, and societal impact assessments, researchers can ensure these techniques align with established ethical frameworks while maintaining public trust. This ethical foundation transforms causal ML from a purely technical exercise into a tool for equitable decision-making across sensitive domains, ultimately leading to societally beneficial outcomes that respect individual rights and community values.

VIII. LIMITATIONS, FUTURE DIRECTIONS, AND RESEARCH OPPORTUNITIES

This study uncovers potential and promising research directions and opportunities that can be summarized as follows:

- **Interpretability and Causal explanation methods** [75], [86], [90]: Advancing techniques for explaining causal relationships in spatiotemporal data models, with key objectives such as:

  * Creating benchmarks for evaluating the interpretability of causal machine learning models in spatiotemporal contexts, ensuring standardized comparisons across methods and domains,
  * Leveraging synergy methods (e.g., combining causal discovery algorithms with visualization tools like counterfactual explanations or feature importance maps) to enhance interpretability and make causal analysis accessible to non-experts,
  * Designing user-friendly visualization tools and interactive interfaces tailored for policymakers and community stakeholders, enabling exploration and understanding of causal insights without requiring technical expertise,
  * Investigating human-in-the-loop interpretability to integrate human feedback and domain expertise into the causal machine learning process, ensuring practical and actionable insights in domains such as healthcare and climate science,
  * Exploring emerging techniques in counterfactual inference, including new methods to simulate "what-if" scenarios within spatiotemporal contexts to improve decision-making processes.

- **Causal Reinforcement Learning for Spatiotemporal Policy Learning** [91]: Developing methods to learn optimal policies in reinforcement learning settings that account for both spatial and temporal dimensions, with a focus on real-world applications. Research priorities include:

  * Developing Safe and Fair Causal Reinforcement Learning algorithms to ensure fairness and safety in decision-making across spatial and temporal contexts, particularly for marginalized communities,
  * Creating benchmarks for evaluating the fairness and safety of learned policies in domains like urban planning or public health, ensuring policies do not lead to adverse outcomes,
  * Exploring real-time policy adaptation techniques to dynamically update policies as new data becomes available, enabling continuous improvement in dynamic environments,
  * Addressing biases in spatiotemporal data by incorporating fairness constraints into the reinforcement learning process, ensuring equitable outcomes,
  * Investigating the integration of causal inference with deep reinforcement learning architectures to enhance policy learning in complex spatiotemporal scenarios.

- **Algorithmic advancements for scalable causal inference** [92], [93]: Pushing forward the creation of new algorithms and scaling techniques for large-scale spatiotemporal datasets, with specific objectives such as:

  * Developing scalable causal inference methods using distributed computing and optimization techniques to handle large datasets efficiently, ensuring computational feasibility for real-world applications,
  * Creating benchmarks for evaluating scalability and accuracy of causal inference algorithms in spatiotemporal contexts, enabling standardized comparisons and progress tracking,
  * Innovating real-time causal learning techniques for dynamic environments, enabling continuous updates to causal relationships as new data streams in, with applications in areas like traffic management and disaster response,
  * Addressing challenges in high-dimensional spatiotemporal data, such as feature selection and dimensionality reduction, to improve computational efficiency and model performance,
  * Exploring scalability improvements using advanced parallel computing techniques such as MapReduce frameworks and GPU-accelerated computing to manage computational challenges in large-scale causal inference models.

This research has several limitations that should be acknowledged. First, the keyword usage strategy may have restricted the scope of retrieved papers. Our exact-phrase searches (using quotation marks) required terms to appear in a specific order, potentially excluding relevant studies using synonymous phrasing. Future work could employ broader search terms initially, followed by rigorous relevance filtering. Second, the search rigidity presents a trade-off. While strict criteria improved precision, they may have excluded pertinent studies. For context, preliminary searches without quotation marks returned 158,000 results in Google Scholar, 1,148 in Springer Nature, and 79 in Mendeley - most irrelevant to our focus. A more balanced approach might combine automated searches with manual abstract/title screening to ensure comprehensive coverage without sacrificing relevance.

IX. CONCLUSION

In this paper, we investigate the transformative role of causal machine learning (Causal ML) techniques in spatiotemporal data analysis, emphasizing the often-overlooked importance of causality. Our study underscores how examining causal relationships within features can reveal the true underlying drivers of observed phenomena and enable more effective responses to change. We explore the distinct challenges and opportunities of integrating causal ML into spatiotemporal analysis, along with the ethical considerations necessary to ensure fair and unbiased outcomes. Additionally, we propose promising research directions, such as improving interpretability, optimizing reinforcement learning policies, and developing advanced algorithms for real-time causal inference in dynamic environments. This work highlights the

TABLE 6. Summary of retrieved publications by keywords (Causal Machine Learning) and sources.

Search terms	Source	Journal names	Number retrieved
"Causal Machine Learning"	Mendeley	SSRN Electronic Journal	18
		Labour Economics:	3
		Nature Communications:	3
		Blood	2
		Diabetes	2
		Econometrics Journal	2
		European journal of Operational research	2
		European Review of Agricultural Economics	2
		Health Services Research	2
		IEEE Access	2
	Springer Nature	Machine learning	26
		Artificial intelligence	10
		Computational economics	7
		Data mining and knowledge discovery	7
		Statistical learning	7
		Bayesian network	6
		Predictive medicine	6
		Computational social sciences	5
		Bayesian inference	4
		Data mining	4
		Non-parametric inference	4
		Parametric inference	4
		Attribution theory	3
		Computer vision	3
		Health informatics	3
		Health policy	3

potential of leveraging causal ML to enrich spatiotemporal data analysis, paving the way for future advancements across diverse fields. By providing robust and accurate causal insights, our research aims to facilitate informed decision-making and deliver innovative solutions to complex problems.

X. APPENDIX

DETAILS OF RESULTS RETRIEVED

The distribution of retrieved publications containing the search terms "Causal Machine Learning" and "Spatiotemporal Data Analysis" is detailed in Tables 6–7. Our analysis focuses on results from Mendeley and Springer Nature, as these platforms provided both comprehensive coverage and structured metadata suitable for systematic review. While IEEE Xplore yielded a substantial number of matches (approximately 12,000 results), the volume exceeded practical screening capacity for this study. The selected datasets offer representative insights into current research trends at the intersection of causal inference and spatiotemporal analysis.

For Springer Nature with the search term "Causal Machine Learning", the following are the disciplines and subdisciplines covered:

A. BY DISCIPLINES

1) Computer science (12)
2) Economics (12)
3) Medicine and public health (9)
4) Science, humanities and social sciences, multidisciplinary (8)
5) Engineering (6)
6) Business and management (4)
7) Biomedicine (3)
8) Earth sciences (1)
9) Environment (1)
10) Finance (1)
11) Life sciences (1)
12) Materials science (1)
13) Mathematics (1)
14) Statistics (1)

B. BY SUBDISCIPLINES

1) Artificial intelligence (8)
2) Science, humanities and social sciences, multidisciplinary (8)
3) Computer science, general (7)
4) Econometrics (6)
5) Economics, general (6)
6) Science, multidisciplinary (5)
7) Computer applications in social and behavioral sciences (4)
8) Health informatics (4)

And for the search term "Spatiotemporal Data Analysis",

C. BY DISCIPLINES

1) Computer science (24)
2) Environment (11)
3) Geography (9)
4) Science, humanities and social sciences, multidisciplinary (9)
5) Earth sciences (8)
6) Engineering (7)

TABLE 7. Summary of retrieved publications by keywords (Spatiotemporal data analysis)and sources.

Search terms	Source	Journal names	Number retrieved
âĂIJSpatiotemporal data analysisâĂİ	Mendeley	Land	4
		Journal of Transport Geography	2
		Acta Geotechnica	1
		Advanced Hydroinfo.: ML and Optimization for Water Resources	1
		Aerospce Science and Technology	1
		Annals of Operations Research	1
		Applied Sciences (Switzerland)	1
		Artificial Intelligence Review	1
		BMC Public Health	1
		Circulation	1
	Springer Nature	Geoinformatics	11
		Geographical information system	9
		Multivariate analysis	9
		Data mining	7
		Environmental geography	7
		Regional geography	7
		Machine learning	6
		Time series analysis	6
		Computer vision	5
		Economic geography	5
		Spatial demography	5
		Spatial economics	5
		Urban ecology	5
		Air pollution and air quality	4
		Applied statistics	4
		Big data	4

7) Economics (4)
8) Medicine and public health (3)
9) Business and management (2)
10) Biomedicine (1)
11) Life sciences (1)
12) Mathematics (1)
13) Social sciences (1)
14) Statistics (1)

The results highlight key disciplines and subjects: Machine Learning leads with 26 occurrences, underscoring its centrality in causal machine learning research. Artificial Intelligence follows with 10 publications as a subject and 8 as a subdiscipline, reflecting its prominence. Computational Economics and Data Mining/Knowledge Discovery each have 7 publications. SSRN Electronic Journal stands out with 18 occurrences, indicating its importance as a source for causal machine learning research. Bayesian Networks and Statistical Learning are also significant, showcasing probabilistic approaches. Interdisciplinary applications are evident in Computer Science and Economics (12 occurrences each), as well as Medicine and Public Health (9 occurrences), demonstrating healthcare applications. For spatiotemporal data analysis, Computer Science dominates with 24 occurrences, followed by Environment (11 occurrences) and Geography (9 occurrences), emphasizing its use in ecological and spatial research.

ACKNOWLEDGMENT

The authors express their deepest thanks for the referees' valuable suggestions for revising and improving the manuscript.

CONFLICTS OF INTEREST

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

REFERENCES

[1] X. Lu, L. Su, and H. White, "Granger causality and structural causality in cross-section and panel data," *Econ. Theory*, vol. 33, no. 2, pp. 263–291, Apr. 2017.

[2] H. Lütkepohl, *New Introduction to Multiple Time Series Analysis*. Cham, Switzerland: Springer, 2005.

[3] S. Ali and J. Wang, "Tutorial on causal inference with spatiotemporal data," in *Proc. 1st ACM SIGSPATIAL Int. Workshop Spatiotemporal Causal Anal.*, Atlanta, GA, USA, Oct. 2024, pp. 23–25, doi: 10.1145/3681778.3698786.

[4] J. E. Brand, X. Zhou, and Y. Xie, "Recent developments in causal inference and machine learning," *Annu. Rev. Sociology*, vol. 49, no. 1, pp. 81–110, Jul. 2023, doi: 10.1146/annurev-soc-030420-015345.

[5] J. Kaddour, A. Lynch, Q. Liu, M. J. Kusner, and R. Silva, "Causal machine learning: A survey and open problems," 2022, *arXiv:2206.15475*.

[6] C. Su, G. Yu, J. Wang, Z. Yan, and L. Cui, "A review of causality-based fairness machine learning," *Intell. Robot.*, vol. 2, no. 3, pp. 244–274, 2022, doi: 10.20517/ir.2022.17.

[7] F. Wang, J. Li, D. Peng, Q. Yi, X. Zhang, J. Zheng, and S. Chen, "Estimating soybean yields using causal inference and deep learning approaches with satellite remote sensing data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 17, pp. 14161–14178, 2024, doi: 10.1109/JSTARS.2024.3435699.

[8] S. Arti, I. Hidayah, and S. S. Kusumawardhani, "Research trend of causal machine learning method: A literature review," *Int. J. Informat. Develop.*, vol. 9, no. 2, pp. 111–118, Dec. 2020, doi: 10.14421/ijid.2020.09208.

[9] A. Baiardi and A. A. Naghi, "The value added of machine learning to causal inference: Evidence from revisited studies," *Econometrics J.*, vol. 27, no. 2, pp. 213–234, Jun. 2024, doi: 10.1093/ectj/utae004.

[10] B. Schölkopf, F. Locatello, S. Bauer, N. R. Ke, N. Kalchbrenner, A. Goyal, and Y. Bengio, "Toward causal representation learning," *Proc. IEEE*, vol. 109, no. 5, pp. 612–634, May 2021, doi: 10.1109/JPROC.2021.3058954.

[11] S. Akbari, F. Jamshidi, E. Mokhtarian, M. J. Vowels, J. Etesami, and N. Kiyavash, "Causal effect identification in uncertain causal networks," 2022, *arXiv:2208.04627*.

[12] Z. Chu, J. Huang, R. Li, W. Chu, and S. Li, "Causal effect estimation: Recent advances, challenges, and opportunities," 2023, *arXiv:2302.00848*.

[13] V. Chernozhukov, C. Hansen, N. Kallus, M. Spindler, and V. Syrgkanis, "Applied causal inference powered by ML and AI," 2024, *arXiv:2403.02467*.

[14] S. Feuerriegel, D. Frauen, V. Melnychuk, J. Schweisthal, K. Hess, A. Curth, S. Bauer, N. Kilbertus, I. S. Kohane, and M. van der Schaar, "Causal machine learning for predicting treatment outcomes," *Nature Med.*, vol. 30, no. 4, pp. 958–968, Apr. 2024, doi: 10.1038/s41591-024-02902-1.

[15] M. M. Alam, L. Torgo, and A. Bifet, "A survey on spatio-temporal data analytics systems," *ACM Comput. Surveys*, vol. 54, no. 10s, pp. 1–38, Jan. 2022, doi: 10.1145/3507904.

[16] C. D. Tilakaratne and L. Liyanage-Hansen, "A review of strengths and weaknesses of SpatioTemporal data analysis techniques," in *Proc. Int. Conf. Mach. Learn. Data Eng. (iCMLDE)*, Dec. 2018, pp. 53–55, doi: 10.1109/iCMLDE.2018.00020.

[17] Z. Dong and C. Guo, "A literature review of spatio-temporal data analysis," *J. Phys., Conf. Ser.*, vol. 1792, no. 1, Feb. 2021, Art. no. 012056, doi: 10.1088/1742-6596/1792/1/012056.

[18] A. Hamdi, K. Shaban, A. Erradi, A. Mohamed, S. K. Rumi, and F. D. Salim, "Spatiotemporal data mining: A survey on challenges and open problems," *Artif. Intell. Rev.*, vol. 55, no. 2, pp. 1441–1488, Feb. 2022, doi: 10.1007/s10462-021-09994-y.

[19] I. Ahmed and A. S. Raihan, "Spatiotemporal data analysis: A review of techniques," in *Multimodal and Tensor Data Analytics for Industrial Systems Improvement*, vol. 211. Cham, Switzerland: Springer, 2024.

[20] G. Jin, Y. Liang, Y. Fang, Z. Shao, J. Huang, J. Zhang, and Y. Zheng, "Spatio-temporal graph neural networks for predictive learning in urban computing: A survey," *IEEE Trans. Knowl. Data Eng.*, vol. 36, no. 10, pp. 5388–5408, Oct. 2024, doi: 10.1109/TKDE.2023.3333824.

[21] C. Yang, K. Clarke, S. Shekhar, and C. V. Tao, "Big spatiotemporal data analytics: A research and innovation frontier," *Int. J. Geographical Inf. Sci.*, vol. 34, no. 6, pp. 1075–1088, Jun. 2020, doi: 10.1080/13658816.2019.1698743.

[22] S. Shekhar, Z. Jiang, R. Ali, E. Eftelioglu, X. Tang, V. Gunturi, and X. Zhou, "Spatiotemporal data mining: A computational perspective," *ISPRS Int. J. Geo-Inf.*, vol. 4, no. 4, pp. 2306–2338, Oct. 2015, doi: 10.3390/ijgi4042306.

[23] C. Meng, S. Rambhatla, and Y. Liu, "Cross-node federated graph neural network for spatio-temporal data modeling," in *Proc. 27th ACM SIGKDD Conf. Knowl. Discovery Data Mining*, Singapore, Aug. 2021, pp. 1202–1211, doi: 10.1145/3447548.3467371.

[24] C. Bergmeir and J. M. Benítez, "On the use of cross-validation for time series predictor evaluation," *Inf. Sci.*, vol. 191, pp. 192–213, May 2012, doi: 10.1016/j.ins.2011.12.028.

[25] J. Vanoli, A. De La Cruz Libardi, F. Sera, M. Stafoggia, P. Masselot, M. N. Mistry, S. Rajagopalan, J. K. Quint, C. F. S. Ng, L. Madaniyazi, and A. Gasparrini, "Long-term associations between time-varying exposure to ambient $PM_{2.5}$ and mortality: An analysis of the U.K. biobank," *Epidemiology*, vol. 36, no. 1, pp. 1–10, Jan. 2025, doi: 10.1097/EDE.0000000000001796.

[26] R. Caferra, A. Leitsch, and N. Peltier, "Constraint-based methods," in *Automated Model Building*. Dordrecht, The Netherlands: Springer, 2004, pp. 151–232.

[27] Y. Song, J. Sohl-Dickstein, D. P. Kingma, A. Kumar, S. Ermon, and B. Poole, "Score-based generative modeling through stochastic differential equations," 2020, *arXiv:2011.13456*.

[28] G. Lacerda, P. L. Spirtes, J. Ramsey, and P. O. Hoyer, "Discovering cyclic causal models by independent components analysis," 2012, *arXiv:1206.3273*.

[29] D. Geiger, D. Heckerman, H. King, and C. Meek, "Identifiability and separability of exponential families and Bayesian networks," *Ann. Statist.*, vol. 40, no. 2, pp. 675–697, 1990, doi: 10.1214/aos/1341304192. [Online]. Available: https://projecteuclid.org/euclid.aos/1341304192

[30] P. R. Rosenbaum and D. B. Rubin, "The central role of the propensity score in observational studies for causal effects," *Biometrika*, vol. 70, no. 1, pp. 41–55, 1983, doi: 10.1093/biomet/70.1.41. [Online]. Available: https://academic.oup.com/biomet/article/70/1/41/245760

[31] J. Pearl, *Causality: Models, Reasoning, and Inference*, 2nd ed., Cambridge, U.K.: Cambridge Univ. Press, 2009, doi: 10.1017/CBO9780511803161.

[32] M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler, and S. Hochreiter, "GANs trained by a two time-scale update rule converge to a local Nash equilibrium," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 1–5.

[33] C. W. J. Granger, "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica*, vol. 37, no. 3, pp. 424–438, Aug. 1969, doi: 10.2307/1912791.

[34] M. R. Kosorok and E. A. Moodie, "Direct learning and optimization methods for treatment decisions," in *Mathematical and Statistical Methods for Actuarial Sciences and Finance*. Cham, Switzerland: Springer, 2019, doi: 10.1007/978-3-030-17259-63.

[35] C. Soares and L. Torgo, "Discovery science," in *Proc. 24th Int. Conf. Discovery Sci.*, 2021, pp. 44–54, doi: 10.1007/978-3-030-88942-5.

[36] A. Shojaie and E. B. Fox, "Granger causality: A review and recent advances," *Annu. Rev. Statist. Its Appl.*, vol. 9, no. 1, pp. 289–319, Mar. 2022, doi: 10.1146/annurev-statistics-040120-010930.

[37] A. Rawal, A. Raglin, D. B. Rawat, and B. M. Sadler, "Causality and machine learning review," Army Research Laboratory, Adelphi, MD, USA, Tech. Rep. 188–704, Oct. 2022. [Online]. Available: https://apps.dtic.mil/sti/pdfs/AD1182780.pdf

[38] X. S. Liang, D. Chen, and R. Zhang, "Quantitative causality, causality-aided discovery, and causal machine learning," *Ocean-Land-Atmosphere Res.*, vol. 2, p. 0026, Jan. 2023, doi: 10.34133/olar.0026.

[39] H. Yuan and G. Li, "A survey of traffic prediction: From spatio-temporal data to intelligent transportation," *Data Sci. Eng.*, vol. 6, no. 1, pp. 63–85, Mar. 2021, doi: 10.1007/s41019-020-00151-z.

[40] A. Wald, "Tests of statistical hypotheses concerning several parameters when the number of observations is large," *Trans. Amer. Math. Soc.*, vol. 54, no. 3, pp. 426–482, Jan. 1943, doi: 10.1090/s0002-9947-1943-0012401-3.

[41] S. Lee and Y. Xia, "Modern ANOVA: A unified approach with machine learning extensions," *J. Comput. Graph. Statist.*, vol. 32, no. 1, pp. 1–17, 2023.

[42] S. Vansteelandt and O. Dukes, "Assumption-lean inference for generalised linear model parameters," 2020, *arXiv:2006.08402*.

[43] M. Eichler, "Granger causality and path diagrams for multivariate time series," *J. Econometrics*, vol. 137, no. 2, pp. 334–353, Apr. 2007, doi: 10.1016/j.jeconom.2005.06.032.

[44] J. Oud and H. Folmer. (2008). *A Structural Equation Approach to Spatial Dependence Models*. [Online]. Available: https://www.academia.edu/13125706/A_Structural_Equation_Approach_to_Models_with_Spatial_Dependence?form=MG0AV3

[45] M. Moshagen and M. Bader, "SemPower: General power analysis for structural equation models," *Behav. Res. Methods*, vol. 56, no. 4, pp. 2901–2922, Nov. 2023, doi: 10.3758/s13428-023-02254-7.

[46] Q. Luo, W. Lu, W. Cheng, P. A. Valdes-Sosa, X. Wen, M. Ding, and J. Feng, "Spatio-temporal Granger causality: A new framework," 2012, *arXiv:1210.3889*.

[47] N. Pavasant, T. Morita, M. Numao, and K.-I. Fukui, "Granger causality-based cluster sequence mining for spatio-temporal causal relation mining," *Int. J. Data Sci. Analytics*, vol. 17, no. 3, pp. 275–288, Apr. 2024, doi: 10.1007/s41060-023-00411-x.

[48] Y. Li, D. Yu, Z. Liu, M. Zhang, X. Gong, and L. Zhao, "Graph neural network for spatiotemporal data: Methods and applications," 2023, *arXiv:2306.00012*.

[49] M. L. Davis, B. Neelon, P. J. Nietert, L. F. Burgette, K. J. Hunt, A. B. Lawson, and L. E. Egede, "Propensity score matching for multilevel spatial data: Accounting for geographic confounding in health disparity studies," *Int. J. Health Geographics*, vol. 20, no. 1, p. 10, Dec. 2021, doi: 10.1186/s12942-021-00265-1.

[50] G. Papadogeorgou, C. Choirat, and C. M. Zigler, "Adjusting for unmeasured spatial confounding with distance adjusted propensity score matching," *Biostatistics*, vol. 20, no. 2, pp. 256–272, Apr. 2019, doi: 10.1093/biostatistics/kxx074.

[51] S. J. Cook, J. C. Hays, and R. J. Franzese, "STADL up! The spatiotemporal autoregressive distributed lag model for TSCS data analysis," *Amer. Political Sci. Rev.*, vol. 117, no. 1, pp. 59–79, Feb. 2023, doi: 10.1017/s0003055422000272.

[52] M. L. Nascimento, K. C. M. Gonçalves, and M. J. Mendonça, "Spatio-temporal instrumental variables regression with missing data: A Bayesian approach," *Comput. Econ.*, vol. 62, no. 1, pp. 29–47, Jun. 2023, doi: 10.1007/s10614-022-10269-z.

[53] A. Zanga and F. Stella, "A survey on causal discovery: Theory and practice," 2023, *arXiv:2305.10032*.
[54] F. Mumuni and A. Mumuni, "Explainable artificial intelligence (XAI): From inherent explainability to large language models," 2025, *arXiv:2501.09967*.
[55] S. Wang, J. Cao, and P. S. Yu, "Deep learning for spatio-temporal data mining: A survey," 2019, *arXiv:1906.04928*.
[56] (2022). *Tutorial on Causal Inference and Its Connections to Machine Learning (Using DoWhy+EconML)*. [Online]. Available: https://www.pywhy.org/dowhy/v0.12/example/tutorial-causalinference-machinelearning-using-dowhy-econml.html?form=MG0AV3
[57] P. Madhavi and S. K. P. Supreethi, "A metaphorical analysis of different encoding techniques for spatial–temporal data," *Int. J. Intell. Syst. Appl. Eng.*, vol. 11, no. 11, pp. 302–308, 2023. [Online]. Available: https://ijisae.org/index.php/IJISAE/article/view/3473/2069
[58] C. O. S. Sorzano, J. Vargas, and A. Pascual Montano, "A survey of dimensionality reduction techniques," 2014, *arXiv:1403.2877*.
[59] M. Opper and O. Winther, "Expectation consistent approximate inference," *J. Mach. Learn. Res.*, vol. 6, no. 73, pp. 2177–2204, 2005. [Online]. Available: https://jmlr.org/papers/volume6/opper05a/opper05a.pdf
[60] S. Mandt, M. D. Hoffman, and D. M. Blei, "Stochastic gradient descent as approximate Bayesian inference," *J. Mach. Learn. Res.*, vol. 18, no. 134, pp. 1–35, 2017. [Online]. Available: https://jmlr.org/papers/volume18/17-214/17-214.pdf
[61] M. Zaharia, M. Chowdhury, T. Das, A. Dave, J. Ma, M. McCauley, M. J. Franklin, S. Shenker, and I. Stoica, "Apache spark: A unified engine for big data processing," *Commun. ACM*, vol. 59, no. 11, pp. 56–65, Nov. 2016.
[62] R. Raina, A. Madhavan, and A. Y. Ng, "Large-scale deep unsupervised learning using graphics processors," in *Proc. 26th Annu. Int. Conf. Mach. Learn.*, Jun. 2009, pp. 873–880. [Online]. Available: https://paperswithcode.com/paper/simplifying-probabilistic-expressions-in
[63] P. Spirtes, C. Glymour, and R. Scheines, *Causation, Prediction, and Search*. Cambridge, MA, USA: MIT Press, 2000. [Online]. Available: https://jmlr.org/papers/volume24/21-1463/21-1463.pdf
[64] I. Tsamardinos, L. E. Brown, and C. F. Aliferis, "The max-min hill-climbing Bayesian network structure learning algorithm," *Mach. Learn.*, vol. 65, no. 1, pp. 31–78, 2006.
[65] V. Gupta, S. Y. Philip, and V. H. Tuulos, "Spatiotemporal aggregation for efficient and accurate sentiment analysis," in *Proc. 21st ACM Int. Conf. Inf. Knowl. Manage.*, 2012, pp. 2061–2064.
[66] J. Peters, P. Bühlmann, and N. Meinshausen, "Causal inference using conditional independence tests," *J. Roy. Stat. Soc. B, Stat. Methodol.*, vol. 76, no. 4, pp. 617–646, 2014. [Online]. Available: https://jmlr.org/papers/volume6/opper05a/opper05a.pdf
[67] I. H. Sarker, "Machine learning: Algorithms, real-world applications and research directions," *Social Netw. Comput. Sci.*, vol. 2, no. 3, p. 160, May 2021. [Online]. Available: https://link.springer.com/article/10.1007/s42979-021-00592-x
[68] Z. Li and D. Hoiem, "Learning without forgetting," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 12, pp. 2935–2947, Dec. 2017.
[69] R. Tibshirani, "Regression shrinkage and selection via the lasso," *J. Roy. Stat. Soc. B, Stat. Methodol.*, vol. 58, no. 1, pp. 267–288, Jan. 1996. [Online]. Available: https://hastie.su.domains/Papers/LASSO.pdf
[70] (2018). *How to Deal With Missing Data*. [Online]. Available: https://u.osu.edu/missingdata/mechanisms-of-missingness/?form=MG0AV3
[71] I. Marisca, A. Cini, and C. Alippi, "Learning to reconstruct missing data from spatiotemporal graphs with sparse observations," 2022, *arXiv:2205.13479*.
[72] A. Abadie and G. W. Imbens. (2009). *Matching on the Estimated Propensity Score*. [Online]. Available: https://scholar.harvard.edu/files/imbens/files/matching.pdf
[73] (2025). *What is Noise in Data and Its Relation to Overfitting*. [Online]. Available: https://futuremachinelearning.org/what-is-noise-in-data-and-its-relation-to-overfitting/?form=MG0AV3
[74] B. C. M. Fung, K. Wang, R. Chen, and P. S. Yu, "Privacy-preserving data publishing: A survey of recent developments," *ACM Comput. Surveys*, vol. 42, no. 4, pp. 1–53, Jun. 2010, doi: 10.1145/1749603.1749605.
[75] N. Mehrabi, F. Morstatter, N. Saxena, K. Lerman, and A. Galstyan, "A survey on bias and fairness in machine learning," 2019, *arXiv:1908.09635*.
[76] L. D'Agostino McGowan, "Sensitivity analyses for unmeasured confounders," *Current Epidemiology Rep.*, vol. 9, no. 4, pp. 361–375, Nov. 2022, doi: 10.1007/s40471-022-00308-6.
[77] X. Zhu, Y. Zhang, F. Feng, X. Yang, D. Wang, and X. He, "Mitigating hidden confounding effects for causal recommendation," *IEEE Trans. Knowl. Data Eng.*, vol. 36, no. 9, pp. 4794–4805, Sep. 2024, doi: 10.1109/TKDE.2024.3378482.
[78] M. A. Hernán and S. Monge, "Selection bias due to conditioning on a collider," *BMJ*, vol. 2023, p. 1135, Jun. 2023, doi: 10.1136/bmj.p1135.
[79] Y. Wang and L. Singh, "Analyzing the impact of missing values and selection bias on fairness," *Int. J. Data Sci. Anal.*, vol. 12, no. 2, pp. 101–119, Aug. 2021, doi: 10.1007/s41060-021-00259-z.
[80] T. Molina-Millán and K. Macours, "Attrition in randomized controlled trials: Using tracking information to correct bias," *Econ. Develop. Cultural Change*, vol. 73, no. 2, pp. 811–834, Jan. 2025, doi: 10.1086/730612.
[81] T. Li, C. Sbarufatti, F. Cadini, J. Chen, and S. Yuan, "Particle filter-based hybrid damage prognosis considering measurement bias," *Struct. Control Health Monitor.*, vol. 29, no. 4, pp. 1–18, Apr. 2022, doi: 10.1002/stc.2914.
[82] M. Kuroki and J. Pearl, "Measurement bias and effect restoration in causal inference," *Biometrika*, vol. 101, no. 2, pp. 423–437, Jun. 2014, doi: 10.1093/biomet/ast066.
[83] C. A. Nobre, G. Sampaio, L. S. Borma, J. C. Castilla-Rubio, J. S. Silva, and M. Cardoso, "Land-use and climate change risks in the Amazon and the need of a novel sustainable development paradigm," *Proc. Nat. Acad. Sci.*, vol. 113, no. 39, pp. 10759–10768, Sep. 2016, doi: 10.1073/pnas.1605516113.
[84] J. Pearl, "On measurement bias in causal inference," 2012, *arXiv:1203.3504*.
[85] R. Moraffah, M. Karami, R. Guo, A. Raglin, and H. Liu, "Causal interpretability for machine learning–problems, methods and evaluation," *ACM SIGKDD Explor. Newslett.*, vol. 22, no. 1, pp. 18–33, May 2020, doi: 10.1145/3400051.3400058.
[86] A. Rahmattalabi and A. Xiang, "Promises and challenges of causality for ethical machine learning," 2022, *arXiv:2201.10683*.
[87] I. S. Rubinstein and N. Good, "The trouble with article 25 (and how to fix it): The future of data protection by design and default," *Int. Data Privacy Law*, vol. 10, no. 1, pp. 37–56, Feb. 2020, doi: 10.1093/idpl/ipz019.
[88] K. Lum and W. Isaac, "To predict and serve?" *Significance*, vol. 13, no. 5, pp. 14–19, Oct. 2016, doi: 10.1111/j.1740-9713.2016.00960.x.
[89] B. D. Mittelstadt, P. Allo, M. Taddeo, S. Wachter, and L. Floridi, "The ethics of algorithms: Mapping the debate," *Big Data Soc.*, vol. 3, no. 2, Dec. 2016, Art. no. 2053951716679679, doi: 10.1177/2053951716679679.
[90] N. Burkart and M. F. Huber, "A survey on the explainability of supervised machine learning," *J. Artif. Intell. Res.*, vol. 70, pp. 245–317, Jan. 2021, doi: 10.1613/jair.1.12228. [Online]. Available: https://jair.org/index.php/jair/article/view/12228
[91] Y. Zeng, R. Cai, F. Sun, L. Huang, and Z. Hao, "A survey on causal reinforcement learning," 2023, *arXiv:2302.05209*.
[92] K. Liang and J. Lin, "Advances in causal inference methods for biological network analysis," *Comput. Mol. Biol.*, vol. 14, no. 2, pp. 84–94, Apr. 2024, doi: 10.5376/cmb.2024.14.0010. [Online]. Available: http://bioscipublisher.com/index.php/cmb/article/view/3925
[93] Z. Yu, J. Ruan, and D. Xing, "Learning causal dynamics models in object-oriented environments," 2024, *arXiv:2405.12615*.
[94] I. Marisca, C. Alippi, and F. Maria Bianchi, "Graph-based forecasting with missing data through spatiotemporal downsampling," 2024, *arXiv:2402.10634*.
[95] X. Liu, M. M. Wall, and J. S. Hodges, "Generalized spatial structural equation models," *Biostatistics*, vol. 6, no. 4, pp. 539–557, Oct. 2005, doi: 10.1093/biostatistics/kxi026.
[96] X. Xu, H. Mao, Y. Zhao, and X. Lü, "An urban traffic flow fusion network based on a causal spatiotemporal graph convolution network," *Appl. Sci.*, vol. 12, no. 14, p. 7010, Jul. 2022, doi: 10.3390/app12147010.
[97] B. Muñoz-Ospina, D. Alvarez-Garcia, H. J. C. Clavijo-Moran, J. A. Valderrama-Chaparro, M. García-Peña, C. A. Herrán, C. C. Urcuqui, A. Navarro-Cadavid, and J. Orozco, "Machine learning classifiers to evaluate data from gait analysis with depth cameras in patients with Parkinson's disease," *Frontiers Human Neurosci.*, vol. 16, May 2022, Art. no. 826376, doi: 10.3389/fnhum.2022.826376.

[98] G.-L. Huang, K. Deng, and J. He, "Cognitive traffic anomaly prediction from GPS trajectories using visible outlier indexes and meshed spatiotemporal neighborhoods," *Cognit. Comput.*, vol. 12, no. 5, pp. 967–978, Sep. 2020, doi: 10.1007/s12559-020-09735-3.

[99] S. S. Prasetiyowati and Y. Sibaroni, "Unlocking the potential of naive Bayes for spatio temporal classification: A novel approach to feature expansion," *J. Big Data*, vol. 11, no. 1, pp. 27–106, Aug. 2024, doi: 10.1186/s40537-024-00958-x.

[100] C. Flynn, A. Guha, S. Majumdar, D. Srivastava, and Z. Zhou, "Towards algorithmic fairness in space-time: Filling in black holes," 2022, *arXiv:2211.04568*.

[101] H. Zang, J. Zhu, and Q. Gao, "Deep learning architecture for flight flow spatiotemporal prediction in airport network," *Electronics*, vol. 11, no. 23, p. 4058, Dec. 2022, doi: 10.3390/electronics11234058.

[102] M. T. Kashifi, "Robust spatiotemporal crash risk prediction with gated recurrent convolution network and interpretable insights from Shapley additive explanations," *Eng. Appl. Artif. Intell.*, vol. 127, Jan. 2024, Art. no. 107379, doi: 10.1016/j.engappai.2023.107379.

[103] X. Xie, Z. Wang, M. Xu, and N. Xu, "Daily PM2.5 concentration prediction based on variational modal decomposition and deep learning for multi-site temporal and spatial fusion of meteorological factors," *Environ. Monitor. Assessment*, vol. 196, no. 9, p. 859, Sep. 2024, doi: 10.1007/s10661-024-13005-2.

[104] S. Han, Y. Park, M. Lee, J. An, and D. Lee, "Enhancing spatiotemporal traffic prediction through urban human activity analysis," 2023, *arXiv:2308.10282*.

[105] X. Wen, Y. Xie, L. Jiang, Z. Pu, and T. Ge, "Applications of machine learning methods in traffic crash severity modelling: Current status and future directions," *Transp. Rev.*, vol. 41, no. 6, pp. 855–879, Nov. 2021, doi: 10.1080/01441647.2021.1954108.

[106] M. Pang, E. Du, and C. Zheng, "Contaminant transport modeling and source attribution with attention-based graph neural network," *Water Resources Res.*, vol. 60, no. 6, pp. 1–25, Jun. 2024, doi: 10.1029/2023wr035278.

[107] W. Du, S. Chen, Z. Li, X. Cao, and Y. Lv, "A spatial–temporal approach for multi-airport traffic flow prediction through causality graphs," *IEEE Trans. Intell. Transp. Syst.*, vol. 25, no. 1, pp. 532–544, Jan. 2024, doi: 10.1109/TITS.2023.3308903.

[108] Y. M. Hwang, S. Park, H. O. Lee, S.-K. Ko, and B.-T. Lee, "Deep learning for human activity recognition based on causality feature extraction," *IEEE Access*, vol. 9, pp. 112257–112275, 2021, doi: 10.1109/ACCESS.2021.3103211.

[109] M. Doborjeh, Z. Doborjeh, A. Merkin, R. Krishnamurthi, R. Enayatollahi, V. Feigin, and N. Kasabov, "Personalized spiking neural network models of clinical and environmental factors to predict stroke," *Cognit. Comput.*, vol. 14, no. 6, pp. 2187–2202, Nov. 2022, doi: 10.1007/s12559-021-09975-x.

[110] W. Wu and Y. Kang, "Ensemble empirical mode decomposition Granger causality test dynamic graph attention transformer network: Integrating transformer and graph neural network models for multi-sensor cross-temporal granularity water demand forecasting," *Appl. Sci.*, vol. 14, no. 8, p. 3428, Apr. 2024, doi: 10.3390/app14083428.

[111] T. Shi, P. Wang, X. Qi, J. Yang, R. He, J. Yang, and Y. Han, "CPT-DF: Congestion prediction on toll-gates using deep learning and fuzzy evaluation for freeway network in China," *J. Adv. Transp.*, vol. 2023, pp. 1–16, Apr. 2023, doi: 10.1155/2023/2941035.

[112] M. Jiang and Z. Liu, "Traffic flow prediction based on dynamic graph spatial–temporal neural network," *Mathematics*, vol. 11, no. 11, p. 2528, May 2023, doi: 10.3390/math11112528.

[113] K. Zhang, L. Zheng, Z. Liu, and N. Jia, "A deep learning based multitask model for network-wide traffic speed prediction," *Neurocomputing*, vol. 396, pp. 438–450, Jul. 2020, doi: 10.1016/j.neucom.2018.10.097.

[114] Y. Zhou and X. Yu, "Multi-graph spatial–temporal synchronous network for student performance prediction," *IEEE Access*, vol. 12, pp. 142306–142319, 2024, doi: 10.1109/ACCESS.2024.3471681.

[115] X.-A. Tibau, C. Reimers, A. Gerhardus, J. Denzler, V. Eyring, and J. Runge, "A spatiotemporal stochastic climate model for benchmarking causal discovery methods for teleconnections," *Environ. Data Sci.*, vol. 1, p. e12, Mar. 2022, doi: 10.1017/eds.2022.11.

[116] P. Li, Y. Yu, D. Huang, Z.-H. Wang, and A. Sharma, "Regional heatwave prediction using graph neural network and weather station data," *Geophys. Res. Lett.*, vol. 50, no. 7, pp. 1–11, Apr. 2023, doi: 10.1029/2023gl103405.

[117] N. Easaw, W. S. Lee, P. S. Lohiya, S. Jalan, and P. Pradhan, "Estimation of correlation matrices from limited time series data using machine learning," *J. Comput. Sci.*, vol. 71, Jul. 2023, Art. no. 102053, doi: 10.1016/j.jocs.2023.102053.

[118] L. Wang, S. Shang, and Z. Wu, "Research on indoor 3D positioning algorithm based on WiFi fingerprint," *Sensors*, vol. 23, no. 1, p. 153, Dec. 2022, doi: 10.3390/s23010153.

[119] S. Hong, Y. Choi, and J.-J. Jeon, "Interpretable water level forecaster with spatiotemporal causal attention mechanisms," *Int. J. Forecasting*, vol. 41, no. 3, pp. 1037–1054, Jul. 2025, doi: 10.1016/j.ijforecast.2024.10.003.

[120] X. Liu, S. Qian, H.-H. Teo, and W. Ma, "Estimating and mitigating the congestion effect of curbside pick-ups and drop-offs: A causal inference approach," *Transp. Sci.*, vol. 58, no. 2, pp. 355–376, Mar. 2024, doi: 10.1287/trsc.2022.0195.

[121] Y. Liu, Y.-S. Wei, H. Yan, G.-B. Li, and L. Lin, "Causal reasoning meets visual representation learning: A prospective study," *Mach. Intell. Res.*, vol. 19, no. 6, pp. 485–511, Dec. 2022, doi: 10.1007/s11633-022-1362-z.

[122] C. R. DeVol, S. R. Shrivastav, A. M. Spomer, K. F. Bjornson, D. Roge, C. T. Moritz, and K. M. Steele, "Effects of interval treadmill training on spatiotemporal parameters in children with cerebral palsy: A machine learning approach," *J. Biomechanics*, vol. 177, Dec. 2024, Art. no. 112397, doi: 10.1016/j.jbiomech.2024.112397.

[123] M. BakhshiGanje, S. Mahmoodi, K. Ahmadi, and M. Mirabolfathy, "Potential distribution of biscogniauxia mediterranea and obolarina persica causal agents of oak charcoal disease in Iran's zagros forests," *Sci. Rep.*, vol. 14, no. 1, p. 7784, Apr. 2024, doi: 10.1038/s41598-024-57298-2.

[124] S. Panda, "Ethical use of technology in organizations for creating social values: A review," *SSRN Electron. J.*, pp. 1–5, Sep. 2023. [Online]. Available: https://papers.ssrn.com/sol3/papers.cfm?abstract=4557054

[125] Y. Fu, C. Zhang, C. Li, M. Zhen, W. Chen, Y. Ji, and H. Hua, "An attention-guided spatio-temporal convolutional network (AG-STCN) for spatio-temporal characterization analysis," *Electronics*, vol. 13, no. 24, p. 4889, Dec. 2024, doi: 10.3390/electronics13244889.

[126] C. Liu, K. Gwn Lore, Z. Jiang, and S. Sarkar, "Root-cause analysis for time-series anomalies via spatiotemporal graphical modeling in distributed complex systems," 2018, *arXiv:1805.12296*.

[127] X. Ruan, C. Wülfing, and R. F. Murphy, "Image-based spatiotemporal causality inference for protein signaling networks," *Bioinformatics*, vol. 33, no. 14, pp. i217–i224, Jul. 2017, doi: 10.1093/bioinformatics/btx258.

[128] D. Prabhakaran et al., "Health effects of selected environmental exposomes across the life courSe in Indian populations using longitudinal cohort studies: GEOHealth HEALS study protocol," *BMJ Open*, vol. 14, no. 10, Oct. 2024, Art. no. e087445, doi: 10.1136/bmjopen-2024-087445.

[129] A. Rupe, N. Kumar, V. Epifanov, K. Kashinath, O. Pavlyk, F. Schlimbach, M. Patwary, S. Maidanov, V. Lee, and J. P. Crutchfield, "DisCo: Physics-based unsupervised discovery of coherent structures in spatiotemporal systems," 2019, *arXiv:1909.11822*.

[130] T. Chauhan, V. Chandel, and S. Ghosh, "Global land drought hubs confounded by teleconnection hotspots in equatorial oceans," *NPJ Climate Atmos. Sci.*, vol. 7, no. 1, pp. 11–15, Jan. 2024, doi: 10.1038/s41612-023-00558-1.

[131] Y. Yuan, K. Wang, Z. Wang, H. Zheng, Z. Ma, R. Liu, K. Hu, Z. Yang, and Y. Zhang, "Ambient ozone exposure and depression among middle-aged and older adults: Nationwide longitudinal evidence in China," *Int. J. Hygiene Environ. Health*, vol. 251, Jun. 2023, Art. no. 114185, doi: 10.1016/j.ijheh.2023.114185.

[132] R. Pagany, "Wildlife-vehicle collisions–Influencing factors, data collection and research methods," *Biol. Conservation*, vol. 251, Nov. 2020, Art. no. 108758, doi: 10.1016/j.biocon.2020.108758.

[133] N. Sai Krishna Mohan, "Understanding time and space complexity in algorithms," *Int. J. Sci. Res.*, vol. 8, no. 10, pp. 1893–1895, Oct. 2019. [Online]. Available: https://www.ijsr.net/getabstract.php?paperid=SR24923134130

[134] P. Upadhyaya, K. Zhang, C. Li, X. Jiang, and Y. Kim, "Scalable causal structure learning: Scoping review of traditional and deep learning algorithms and new opportunities in biomedicine," *JMIR Med. Informat.*, vol. 11, Jan. 2023, Art. no. e38266, doi: 10.2196/38266. [Online]. Available: https://medinform.jmir.org/2023/1/e38266

[135] B. Dong, L. Khan, M. Smith, J. Trevino, B. Zhao, G. L. Hamer, U. A. Lopez-Lemus, A. A. Molina, J. Lubinda, U.-S.-D. T. Nguyen, and U. Haque, "Spatio-temporal dynamics of three diseases caused by aedes-borne arboviruses in Mexico," *Commun. Med.*, vol. 2, no. 1, pp. 11–134, Oct. 2022, doi: 10.1038/s43856-022-00192-7.

[136] A. Uthamacumaran and H. Zenil, "A review of mathematical and computational methods in cancer dynamics," 2022, *arXiv:2201.02055*.

[137] S. Wang, X. Huang, M. Zhang, S. Bao, L. Liu, X. Fu, T. Zhang, Y. Song, P. Kedron, J. Wilson, X. Ye, C. Yang, and W. Guan, "Open science 2.0: Revolutionizing spatiotemporal data sharing and collaboration," *Int. J. Digit. Earth*, vol. 5, no. 1, pp. 123–137, Jan. 2025.

[138] D. Pavlyuk, "Feature selection and extraction in spatiotemporal traffic forecasting: A systematic literature review," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 4, pp. 2548–2562, Apr. 2021.

[139] C.-M. Forke and M. Tropmann-Frick, "Feature engineering techniques and spatio-temporal data processing," in *Proc. Int. Conf. Big Data Smart Comput.*, vol. 21, 2021, pp. 237–244.

[140] T. S. Richardson, "A discovery algorithm for directed cyclic graphs," 2013, *arXiv:1302.3599*.

[141] R. Hatami, "A review of the techniques used to control confounding bias and how spatiotemporal variation can be controlled in environmental impact studies," *Water, Air, Soil Pollut.*, vol. 230, no. 6, p. 132, Jun. 2019, doi: 10.1007/s11270-019-4150-9. [Online]. Available: http://link.springer.com/10.1007/s11270-019-4150-9

[142] N. Mankovich, H. Durand, E. Diaz, G. Varando, and G. Camps-Valls, "Recovering latent confounders from high-dimensional proxy variables," 2024, *arXiv:2403.14228*.

[143] Q. Tan, Y. Liu, and J. Liu, "Mutually adaptable learning," *IEEE Trans. Emerg. Topics Comput. Intell.*, vol. 8, no. 1, pp. 240–254, Feb. 2024, doi: 10.1109/TETCI.2023.3300183.

[144] C. Liu, K. Gwn Lore, and S. Sarkar, "Data-driven root-cause analysis for distributed system anomalies," 2016, *arXiv:1605.06421*.

[145] P.-Y. Yin, "A novel spatiotemporal analysis framework for air pollution episode association in Puli, Taiwan," *Appl. Sci.*, vol. 13, no. 9, p. 5808, May 2023, doi: 10.3390/app13095808.

[146] C. Liu, M. Zhao, A. Sharma, and S. Sarkar, "Traffic dynamics exploration and incident detection using spatiotemporal graphical modeling," *J. Big Data Anal. Transp.*, vol. 1, no. 1, pp. 37–55, Jun. 2019, doi: 10.1007/s42421-019-00003-x.

[147] C. Rudin, "Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead," *Nature Mach. Intell.*, vol. 1, no. 5, pp. 206–215, May 2019, doi: 10.1038/s42256-019-0048-x.

[148] J. Pearl and D. Mackenzie, *The Book of Why: The New Science of Cause and Effect*. New York, NY, USA: Basic Books, 2018.

[149] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 30, 2017, pp. 5998–6008. [Online]. Available: https://proceedings.neurips.cc/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf

[150] S. Lundberg and S. Lee, "A unified approach to interpreting model predictions," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 4765–4774. [Online]. Available: https://proceedings.neurips.cc/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf

[151] C. Molnar, *Interpretable Machine Learning*. Morrisville, NC, USA: Lulu, 2020. [Online]. Available: https://christophm.github.io/interpretable-ml-book/

[152] S. Liu, X. Wang, C. Collins, W. Dou, F. Ouyang, M. El-Assady, L. Jiang, and D. A. Keim, "Visual causality analysis of spatio-temporal event data," *IEEE Trans. Vis. Comput. Graph.*, vol. 28, no. 1, pp. 1153–1163, Jan. 2021, doi: 10.1109/TVCG.2020.3030400.

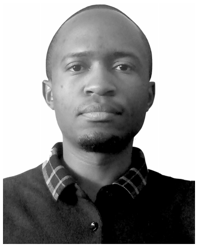
CHRISTIAN M. MULOMBA received the B.S. degree in computer science from the Higher Institute of Computer Science, Programming and Analysis, Kinshasa, Democratic Republic of Congo, and the M.S. and Ph.D. degrees in ICT convergence from Handong Global University, Pohang, South Korea, specializing in the application of machine learning and deep learning. He has extensive experience as an IT and Data Analyst in the banking sector, and the Ministry of Economy of DRC. He is currently an Assistant Professor with Kyungsung University, Busan, South Korea. His research interests include the application of the convergence of causal inference with machine learning, and Bayes' theorem with machine learning to address complex problems in various domains.

VOGEL M. KIKETA received the master's degree in ICT convergence from Handong Global University, in 2020. He is currently pursuing the Ph.D. degree in engineering and smart city global convergence with Seoul National University, South Korea. His expertise spans artificial intelligence, data science, the IoT, smart cities, software engineering, and educational technology. He is a Senior Lecturer with the University of Kinshasa contributing to teaching and research, where he is a Researcher with the ABIL Research Center. Previously, he was an IT specialist in Democratic Republic of Congo's Ministries of Transport and Public Service, focusing on database management and digital payroll systems. He also helped implement campus information systems at prominent universities, such as the Protestant University of Congo (UPC) and the National Pedagogical University (UPN). As a digital entrepreneur, he engages in technology-driven projects. His research interests include leveraging emerging technologies to solve societal challenges and drive advancements in infrastructure, education, and urban development.

DAVID M. KUTANGILA received the master's and Ph.D. degrees in artificial intelligence from the University of Granada, Granada, Spain, in 2003 and 2005, respectively. His major field of study throughout his academic career has been artificial intelligence. He is currently a University Full Professor at the University of Kinshasa and the Protestant University of Congo. Previously, he held the position of the IT Department Manager of Bravo AIR Congo, where he contributed to enhancing IT infrastructure and organizational efficiency. His professional experience demonstrates a blend of academic excellence and industry expertise in artificial intelligence and information technology. His current research interests include advancing artificial intelligence applications in education and resource optimization, while his previous research explored foundational methods in machine learning and intelligent systems. He has authored several publications in the field reflecting his commitment to both academic and practical advancements in artificial intelligence and its applications.

PESCIE H. K. MAMPUYA received the master's degree in business administration with a specialization in computer science and digital marketing and the Ph.D. degree from the Protestant University of Congo, Kinshasa, Democratic Republic of Congo, in 2017 and 2023, respectively. His major field of study is software engineering and digital marketing, reflecting a strong interdisciplinary expertise. He is currently an Associate Professor with the Protestant University of Congo, where he is actively involved in academic teaching and research. Prior to his academic role, he was the Marketing and Service Department Manager of Orange Congo, where he contributed to innovative strategies combining marketing and technology to improve organizational performance. He is the author of the book *Digital Marketing and Information Systems* (Harmattan, 2024), which explores the integration of digital marketing techniques with advanced information systems. His current research interests include software engineering and digital marketing, with a focus on developing innovative approaches to digital transformation and marketing strategies. He continues to contribute significantly to the advancement of digital marketing and computer science, bridging the gap between industry practices, and academic research.

JUNIOR N. MUKENZE is currently pursuing the Ph.D. degree in AI engineering with Handong Global University. He is a scholar and an innovator specializing in artificial intelligence, ICT, and project management. His research interests include machine learning, deep learning, and predictive analytics to address real-world challenges. With over a decade of professional experience, he is the Head of the IT Office, Ministry of Defense, Kinshasa, and is the CEO of MKZ Futuris. He has held roles as a GIS Engineer at JICA, an Assistant Project Manager at FINCA, and a Teacher/IT Specialist at the College Sainte Agathe. He has presented research at international conferences, including waste recovery in DR Congo (Moscow, 2024) and time series forecasting (Dubai, 2024). Recognized for his contributions, he has received appreciation letters and certificates from global institutions. Fluent in multiple languages, with expertise in GIS, programming, and project management, he is committed to leveraging AI and technology for sustainable development and transformative change in emerging economies.



LANDRY M. KASUNZI is currently pursuing the master's degree in security, systems, and networks with a deep passion for emerging technologies and their transformative impact on solving real-world challenges. His primary focus is on the security aspects of these innovations, recognizing that safeguarding technological advancements is essential to unlocking their full potential. Dedicated to lifelong learning, he thrives under pressure and excels at navigating both the technical and social complexities of rapidly evolving technologies. He is a strong team player, believing that collaboration is key to addressing challenges in today's fast-paced and dynamic world. He is also a Research Assistant with the University of Nouveaux Horizons, Lubumbashi, Democratic Republic of Congo. Additionally, he holds the position of the Secretary to the Scientific Research Coordination Office of the University, where he actively contributes to advancing academic research initiatives and fostering innovation.

KYANDOGHERE KYAMAKYA is currently a Full Professor of transportation informatics and the Deputy Director of the Institute of Smart Systems Technologies, University of Klagenfurt, Austria. He is actively conducting research on modeling, simulation, and test-bed evaluations for a series of concepts among others in the context of intelligent transportation systems. In the research addressing transportation systems, a series of fundamental and theoretical tools from the fields of applied mathematics, electronics, and computer science is either extensively exploited or a source of inspiration for innovative solutions and concepts, including nonlinear dynamics, systems science, machine learning/deep learning, nonlinear image processing, and neuro computing. He co-edited more than six books. He has published more than 100 journal articles and more than 100 conference papers.

TASHO TASHEV received the M.Sc. degree in automation and telemechanic and the Ph.D. degree from the Technical University of Sofia (TU-Sofia), in 1985 and 1998, respectively. He is currently a Professor with the Department of Information Measurement Systems, TU-Sofia. His research interests include adaptive control, artificial neural networks, sensorless control, and speed control. He has contributed to the fields of signal processing and control systems through his work on parallel algorithms, real motion, and rotation angle sampling. He has been involved in numerous research projects and authored numerous articles in prestigious international journals and has actively participated in various international conferences and workshops within his areas of expertise.

SELAIN K. KASEREKA received the bachelor's degree in computer engineering from the University of Kinshasa, the M.Sc. degree in intelligent systems and multimedia (SIM) from Vietnam National University, the M.Sc. degree in content digital engineering (ICONE) from the University of La Rochelle, the M.Sc. degree in complex systems: natural and artificial cognition (CNA) from the University of Paris 8, and the Ph.D. degree in applied mathematics from the University of South Africa (UNISA), Florida, South Africa, in 2020. He is currently an Associate Professor with the University of Kinshasa, where he is a Researcher and a Secretary in Charge of projects at the ABIL Research Center. He is also a Mid-Career Research Fellow at the Technical University of Sofia, Bulgaria, until 2026. He has published several journal articles in internationally renowned journals and participated in several international conferences and workshops in his fields of expertise. His research interests include complex system modeling and simulation to machine learning, deep learning, multi-agent systems, applied mathematics, bio-statistics, biomathematics, epidemiology, sentiment analysis, and emotion and stress recognition.