A Framework for Multidimensional Business Process Forecasting

Aurélie Leribaux[1]

[1]Research Centre for Information Systems Engineering, KU Leuven, Belgium

Abstract

Accurate forecasting of business process models is essential for effective process management and decision-making. However, existing techniques often focus narrowly on predicting control flow, overlooking the rich, multidimensional nature of real-world processes. This research addresses this gap by proposing a comprehensive framework that forecasts interactions between control flow and critical process dimensions, including resources, data objects, decision logic, and performance metrics. Leveraging recent advances in process knowledge graphs and graph neural networks, the proposed approach represents event data as multiple time series, each capturing a distinct behavioral aspect of the process. This multidimensional perspective enables the development of neural network-based models that predict the evolution of complex process structures rather than isolated outcomes. The research agenda is organized into four key pillars: enhancing process representations and predictive models, integrating simulation for process optimization, ensuring model explainability, and validating findings through real-life case studies. By uniting these elements, this work aims to advance the field of predictive process monitoring, offering more accurate, scalable, and interpretable forecasting solutions for business process management.

Keywords: *Predictive Process Monitoring, Multidimensional Process Models, Graph Neural Networks, Simulation, Explainable AI*

1. Motivation and Background

Predictive Process Monitoring (PPM) has emerged as a key component of Business Process Management (BPM), enabling organizations to make proactive decisions by predicting the future states of ongoing process instances. Traditional PPM techniques typically focus on predicting specific outcomes, such as the remaining time of a case or the next activity to be executed, often using classification or regression models trained on historical event logs. While effective in predicting instance-level outcomes, most PPM approaches rely on control-flow-centric models, which focus on the sequence of activities but often neglect other important process dimensions, such as resource usage, data interactions, and decision logic. Even so, there are alternative solutions that take into consideration the data-flow of processes, although there is no common agreement on how to address the problem, given the particular nature of each problem domain. Consequently, these approaches fail to capture the complex, multidimensional nature of real-world business processes generically.

Process Model Forecasting (PMF) was introduced to address some of the limitations of PPM [1]. Instead of predicting individual outcomes, PMF aims to forecast the evolution of the entire process model over time, enabling organizations to anticipate structural process changes. Early PMF work primarily focused on forecasting Directly-Follows Graphs (DFGs) as time series, providing insights into control-flow-level dynamics. However, even these approaches often remain limited to univariate time series forecasting, lacking the integration of additional dimensions, such as resources, data, decision points, and performance metrics, that are crucial for understanding complex process behaviors.

Recent research has begun to address this research gap by extending PMF beyond simple control flow predictions. For example, a time-series-based approach for forecasting process models based on sequence data was introduced [2], yet primarily focused on predicting Directly-Follows relations without integrating other process dimensions. Similarly, a recurrent neural network approach for

Proceedings of the Best BPM Dissertation Award, Doctoral Consortium, and Demonstrations & Resources Forum co-located with 23rd International Conference on Business Process Management (BPM 2025), Seville, Spain, August 31st to September 5th, 2025.

predicting process constraints was proposed [2], but without accounting for resources or data perspectives. Additionally, a data model to enable multidimensional process mining was presented, including dimensions such as material flow and information flow, demonstrating the feasibility of a comprehensive representation of process dynamics [3]. Moreover, recent advancements in object-centric process mining emphasize the need to model event data that includes interactions between different objects within processes, enabling a more comprehensive representation of process behavior [4, 5].

Building upon these foundations, this research project aims to advance predictive business process modeling by developing neural network-based forecasting models that capture the multidimensional nature of real-world processes. Unlike traditional methods that focus solely on control-flow, our approach extends to predicting interactions between and among four key dimensions: (1) resources, (2) data objects, (3) decision logic, and (4) performance metrics. This work is grounded in representing event data as multiple time series, each encapsulating a distinct behavioral aspect of the process. This multidimensional approach leverages advances in both process knowledge graphs (PKGs), event knowledge graphs (EKGs), and graph neural networks (GNNs), facilitating the prediction of process dynamics at the process level rather than merely predicting individual outcomes at the instance level. By integrating these dimensions into a cohesive forecasting framework, this research aims to deliver more accurate, actionable, and explainable insights for BPM.

2. Research Agenda

To realize the vision of multidimensional process forecasting, this research is structured around four key pillars: (1) Enhancing process representations and predictive models, (2) Integrating simulation for process optimization, (3) Ensuring the explainability of predictive models, and (4) Validating findings through real-life case studies. Figure 1 summarizes the methodology. Each pillar contributes to overcoming the limitations of current predictive approaches, enabling more accurate and comprehensive forecasts. In the remainder, each of the pillars is outlined in more detail, including recent research, research gaps, and proposed contributions and methodology.

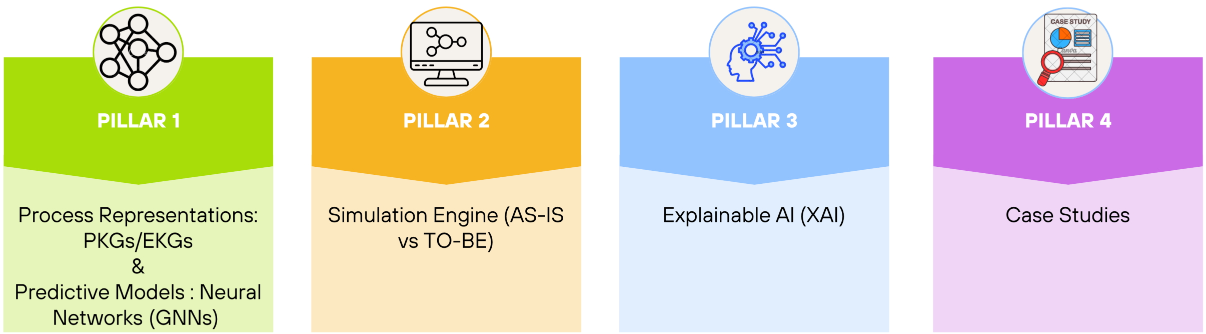

Figure 1: Methodology

2.1. Pillar 1: Enhancing Process Representations and Predictive Models

A fundamental challenge in process forecasting is developing predictive models that accurately capture the multi-dimensional nature of real-world business processes. This challenge motivates the following guiding research question:

- **RQ1:** How can a multi-dimensional process representation and deep learning models be designed to enhance the accuracy, scalability, and interpretability of long-term business process forecasting?

To address this question, this pillar focuses on building a strong foundation for forecasting by emphasizing three key subtopics:

2.1.1. Developing a Multi-Dimensional Process Representation

Existing forecasting models in BPM often focus narrowly on the control flow of processes, overlooking critical aspects such as resource interactions, data dependencies, and decision logic. To address this limitation, this research aims to develop and advance **PKGs** and **EKGs** to build a unified, multidimensional process representation for predictive forecasting.

PKGs and EKGs are both specialized forms of knowledge graphs that capture different facets of process knowledge. PKGs, as applied in several domains such as manufacturing process planning, integrate structured process knowledge from various sources (i.e. process steps, resources, and attributes) and model relationships among them [6]. PKGs can also represent performance metrics and decision logic as attributes or linked entities, enabling comprehensive modeling of process dependencies and requirements. However, these representations are typically static, focusing on structural knowledge rather than modeling dynamic process behavior over time.

EKGs, on the other hand, focus on modeling event-centric knowledge by capturing rich, multi-dimensional relationships between events and entities, such as temporal event relations, event-entity relations, and entity-entity relations [7]. This makes EKGs highly valuable for representing sequential process dynamics and multi-dimensional interactions. However, EKGs are often designed for descriptive analysis, focusing on knowledge extraction and event relations rather than dynamic process forecasting. Moreover, they typically do not explicitly integrate performance metrics or decision logic that are critical in BPM forecasting.

Building on these foundations, this research proposes to enrich and extend KGs into an unified semantic graph. This approach aims to capture both the sequential dynamics of events and the relational dependencies among process elements, such as resources, data objects, and decision logic. By integrating the strengths of both PKGs and EKGs, this research seeks to create a holistic, actionable, and explainable framework for process forecasting that overcomes the limitations of existing models.

2.1.2. Designing GNNs for Process Forecasting

Accurately forecasting business processes requires models that can effectively learn both temporal dependencies and complex structural relationships given their graph-based structures. Hence, this research will design and implement **GNNs** that leverage PKGs to capture interactions between activities, resources, and data elements while maintaining temporal consistency. These GNNs will be specifically tailored to the characteristics of PKGs, ensuring that the models learn from both the sequential and relational aspects of process data.

2.1.3. Adapting GNNs from Other Domains

In addition to developing GNNs specifically designed for PKGs in process mining, this research will also draw inspiration from recent advances. For example, spatio-temporal graph neural networks (ST-GNNs) [8, 9] have shown potential. These models, successfully applied in domains such as traffic forecasting and financial analysis, effectively integrate temporal dynamics with relational structures. This research will investigate how to adapt these techniques for event-driven business processes, ensuring compatibility with the discrete and evolving nature of process event data.

2.2. Pillar 2: Integrating Simulation for Process Optimization

Building upon forecasting in Pillar 1, an essential aspect of this research is simulating and optimizing future process states. While forecasting provides valuable insights into potential future process behaviors, simulation enables stakeholders to evaluate and optimize these behaviors under various scenarios.

Existing research has explored the integration of process mining and simulation, particularly through discrete-event simulation (DES) engines that utilize static process models extracted from event logs [10, 11, 12]. However, these simulation engines often operate on static models, which limits their ability to incorporate dynamic process changes or to simulate complex multi-dimensional process dynamics. They typically do not integrate predictive models that capture how processes evolve over time, especially across multiple process dimensions.

Notable recent developments focus on user perspectives to perform 'what-if' analyses. For example, SIMPT [13] leverages time-aware process trees to enable interactive simulations, allowing process owners to explore process improvement scenarios dynamically. However, even these developments remain limited in their capacity to integrate predictive, multi-dimensional process models

This research pillar aims to bridge these gaps by integrating the predictive models established in Pillar 1 into a comprehensive simulation engine. By transforming PKGs into data-aware, executable process models, this simulation engine will support both AS-IS and TO-BE analyses, empowering stakeholders to assess potential improvements, anticipate dynamic process changes, and mitigate bottlenecks more effectively. The approach aligns with recent calls for more flexible, knowledge-driven simulation frameworks that can dynamically incorporate forecasted process states [3]. This leads to the guiding research question for this pillar:

- **RQ2:** How can simulation and predictive models be jointly integrated to effectively inform process improvement and optimization, while addressing the dynamic and multi-dimensional nature of real-world processes?

2.3. Pillar 3: Ensuring Explainability in Predictive Models

As predictive models become increasingly sophisticated, particularly those leveraging advanced deep learning architectures like GNNs, ensuring their transparency and interpretability becomes essential for trust and informed decision-making in BPM. Moreover, representing the multidimensional nature of business processes, like PKGs, further compounds the challenge of model explainability, as it introduces intricate dependencies between process activities, resources, and data elements. Therefore, this pillar extends Pillar 1 by specifically addressing the heightened challenge of model explainability and ensuring that predictive models remain not only powerful but also transparent and trustworthy for end-users.

Recent research in process mining highlights the growing importance of explainable AI (XAI) to support user trust and effective decision-making [14]. General XAI techniques, including feature attribution, local explanations, and model-agnostic methods, have been extensively surveyed [15], providing a solid foundation for integrating interpretability into deep learning workflows.

Given that this research project leverages GNNs to model and predict PKGs, explainability techniques tailored to GNNs are particularly relevant. Recent research provides a comprehensive taxonomy of XAI methods for GNNs, highlighting the challenges of explaining graph-based predictions [16]. By incorporating these insights, this research will develop XAI techniques that make PKG-based forecasting models more interpretable and actionable for process stakeholders. This aligns with the broader goal of empowering stakeholders with predictive tools that are not only accurate and comprehensive but also transparent and trustworthy. The corresponding pillar, therefore, addresses the following guiding research question:

- **RQ3:** How can explainable AI techniques improve the interpretability and trustworthiness of predictive business process models based on graph neural networks?

2.4. Pillar 4: Validating Findings Through Real-Life Case Studies

A crucial step in operationalizing data-driven research findings is to evaluate them in real-world business settings. This research will apply the developed forecasting models (Pillar 1), simulation techniques (Pillar 2), and explainability frameworks (Pillar 3) to real-life case studies across diverse industries. By testing these approaches in practical environments, we will assess their effectiveness, refine the models based on empirical insights, and demonstrate their value for decision-making in BPM, guided by the following final research question

- **RQ4:** How can real-life case studies validate and refine predictive process forecasting models for practical business applications?

3. Progress to Date

This section describes the current progress, focusing on Pillar 1: Enhancing Process Representations and Predictive Models.

3.1. Literature Review

To achieve the four key pillars, it was necessary to conduct a comprehensive literature review. This review serves as a foundation for understanding the current role of time series in the field of process mining, including how and when they are applied. Hence, the aim of this paper is to clarify the extent to which time series analysis has been integrated into process mining methodologies and to uncover potential areas for further exploration and improvement.

As part of this literature review paper, I had the opportunity to present an initial version of the insights from my analysis at the International Conference on Process Mining of the Belgian Operational Research Society (ORBEL) in Maastricht in February. This event not only offered exposure to the latest developments in the field but also provided valuable feedback from experts, which I have incorporated into the further development of my research. Building on this input, I am preparing a journal paper.

3.2. Initial Steps in Enhancing Process Representations and Predictive Models

3.2.1. Exploring Actor-Level Performance Dynamics

This research extends the predictive modeling focus introduced in Section 1 by investigating both the **resource** and **performance** perspectives as critical dimensions in process forecasting. Motivated by recent work, which proposes a decomposition method to disaggregate performance measures according to actor behavior [17], this research seeks to address the limitations of traditional process mining approaches that often collapse diverse actor behaviors into single aggregate measures [18]. Traditional methods overlook distinctions between uninterrupted work by the same actor, resumed work after an interruption, and handovers between different actors.

Although the authors introduced a resource-aware perspective, their work remains high-level and does not explicitly model the temporal causality of actor behaviors, an essential aspect for understanding process dynamics. To address this, our research incorporates a temporal dimension into performance analysis, transforming performance metrics into time series that capture temporal variations in actor behavior. The analysis also seeks to uncover causal relationships among these time series using Granger causality [19], aiming to reveal how different actor behaviors influence process outcomes.

Ultimately, this work aims to develop a comprehensive framework for understanding process performance at the actor level, serving as a descriptive foundation for incorporating the resource perspective into predictive models and exploring interactions among different process dimensions. The goal is to develop this research into a workshop paper for submission to BPM 2025.

3.2.2. Incorporating Resource Perspectives into Predictive Models

Drawing on insights from the actor-level analysis, the research now incorporates resource perspectives into time-series-based predictive models (e.g., ARIMAX). These models will be benchmarked against traditional models that do not account for resource interactions, providing evidence of the added value of multi-dimensional forecasting.

4. Conclusion

This research seeks to advance business process forecasting by combining multi-dimensional process representations, advanced predictive models, simulation, and explainability. The initial progress within Pillar 1 sets a strong foundation for subsequent work on the other pillars, paving the way toward a comprehensive framework for data-driven BPM.

Declaration on Generative AI

During the preparation of this work, the author used GPT-4 to check grammar, spelling, and improve the writing structure. Additionally, ChatGPT was consulted to better understand related concepts, particularly those found in Graph Neural Network (GNN) literature. No AI-generated content was directly copied into the publication. The author reviewed and edited all AI-assisted input as needed and takes full responsibility for the content of this publication.

References

[1] R. Poll, A. Polyvyanyy, M. Rosemann, M. Röglinger, L. Rupprecht, Process forecasting: Towards proactive business process management, in: Business Process Management, Springer, 2018, pp. 496–512.

[2] J. D. Smedt, J. D. Weerdt, Predictive process model monitoring using recurrent neural networks, arXiv preprint (2020). URL: https://arxiv.org/abs/2011.02819.

[3] S. Kroeger, A. Rafles, P. Jordan, C. Soellner, M. F. Zaeh, Data model to enable multidimensional process mining for data farming based value stream planning in production networks, Production Engineering (2024). URL: https://api.semanticscholar.org/CorpusID:273500321.

[4] W. M. P. van der Aalst, Object-centric process mining: Dealing with divergence and convergence in event data, in: Software Engineering and Formal Methods, Springer, 2019, pp. 3–25.

[5] D. Fahland, Process mining over multiple behavioral dimensions with event knowledge graphs, in: Process Mining Handbook, Springer, 2022, pp. 274–319.

[6] Y. Xiao, S. Zheng, J. Shi, X. Du, J. Hong, Knowledge graph-based manufacturing process planning: A state-of-the-art review, Journal of Manufacturing Systems 70 (2023) 417–435. URL: https://www.sciencedirect.com/science/article/pii/S0278612523001577. doi:https://doi.org/10.1016/j.jmsy.2023.08.006.

[7] S. Guan, X. Cheng, L. Bai, F. Zhang, Z. Li, Y. Zeng, X. Jin, J. Guo, What is event knowledge graph: A survey, 2022. URL: https://arxiv.org/abs/2112.15280. arXiv:2112.15280.

[8] B. Yu, H. Yin, Z. Zhu, Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting, arXiv preprint arXiv:1709.04875 (2017).

[9] D. Cao, Y. Wang, J. Duan, C. Zhang, X. Zhu, C. Huang, Y. Tong, B. Xu, J. Bai, J. Tong, Spectral temporal graph neural network for multivariate time-series forecasting, in: Advances in Neural Information Processing Systems, volume 33, 2020, pp. 17766–17778.

[10] K. M. Rashid, J. Louis, Integrating process mining with discrete-event simulation for dynamic productivity estimation in heavy civil construction operations, Algorithms 15 (2022). URL: https://www.mdpi.com/1999-4893/15/5/173. doi:10.3390/a15050173.

[11] Y. Wang, An integrative process mining approach to mine discrete event simulation model from event data, Theses, Université de Bordeaux, 2018. URL: https://theses.hal.science/tel-01952795.

[12] M. Camargo, M. Dumas, O. González-Rojas, Automated discovery of business process simulation models from event logs, Decision Support Systems 134 (2020) 113284. URL: https://www.sciencedirect.com/science/article/pii/S0167923620300397. doi:https://doi.org/10.1016/j.dss.2020.113284.

[13] M. Pourbafrani, S. Jiao, W. M. P. van der Aalst, Simpt: Process improvement using interactive simulation of time-aware process trees, arXiv preprint arXiv:2108.02052 (2021). URL: https://arxiv.org/abs/2108.02052.

[14] N. Mehdiyev, P. Fettke, Explainable Artificial Intelligence for Process Mining: A General Overview and Application of a Novel Local Explanation Approach for Predictive Process Monitoring, Springer International Publishing, 2021, p. 1–28. URL: http://dx.doi.org/10.1007/978-3-030-64949-4_1. doi:10.1007/978-3-030-64949-4_1.

[15] A. B. Arrieta, N. Díaz-Rodríguez, J. D. Ser, A. Bennetot, S. Tabik, A. Barbado, S. Garcia, S. Gil-Lopez, D. Molina, R. Benjamins, R. Chatila, F. Herrera, Explainable artificial intelligence (xai): Concepts,

taxonomies, opportunities and challenges toward responsible ai, Information Fusion 58 (2020) 82–115.
[16] H. Yuan, H. Yu, S. Gui, S. Ji, Explainability in graph neural networks: A taxonomic survey (2020). doi:10.48550/arXiv.2012.15445.
[17] E. L. Klijn, I. Tentina, D. Fahland, F. Mannhardt, Decomposing process performance based on actor behavior, in: 2024 6th International Conference on Process Mining (ICPM), 2024, pp. 129–136. doi:10.1109/ICPM63005.2024.10680657.
[18] A. Tour, A. Polyvyanyy, A. Kalenkova, Agent system mining: Vision, benefits, and challenges, IEEE Access 9 (2021) 99480–99494. doi:10.1109/ACCESS.2021.3095464.
[19] C. Granger, Investigating causal relations by econometric models and cross-spectral methods, Econometrica 37 (1969) 424–438.