Process Mining as Alternative to Traditional Methods to Describe Process Performance in End-to-End Order Processing of Manufacturing Companies

Günther Schuh[1], Andreas Gützlaff [1], Seth Schmitz [1], Calvin Kuhn[1], and Noah Klapper[1,2]
[1] Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen University, Aachen, Germany
[2] Digital Capability Center (DCC) Aachen / ITA Academy GMBH, Aachen, Germany
Email: g.schuh@wzl.rwth-aachen.de,{a.guetzlaff, s.schmitz, c.kuhn, n.klapper}@wzl.rwth-aachen.de

Abstract—**The description of process efficiency remains a key factor for manufacturing companies competing in volatile markets. Since describing the process performance requires the consideration of all order-fulfilling activities, focusing on the end-to-end order processing process is crucial. Classical techniques for process description are time- and cost-intensive while relying on situational impressions. Consequently, improvement approaches are based on gut feelings and cannot consider dynamic process behaviour. Process Mining can be used for fact-based and objective process descriptions. However, today's process mining applications are mainly conducted in partial processes with similar order types. In the end-to-end order processing, multiple orders with one-to-many and many-to-many relationships exist that need an object-centric process mining approach. This paper presents a methodology for the application of process mining in end-to-end order processing with multiple order types. Based on data from software infrastructure, the integration of the methodology provides manufacturing companies with process models and process performance indicators to describe their PP in end-to-end order processing processes.**

Index Terms—**object-centric process mining, order processing, manufacturing companies**

I. Introduction

To compete in fast-paced environments, manufacturing companies describe their process performance (PP) in order to assess their competitiveness. PP measures processes' progress towards their objectives [1] and a process consists of numerous sub-processes and activities. Describing PP includes mapping as-is processes and measuring Process Performance Indicators (PPI). Thus, PPI need to be unequivocally determinable [2].

Assessing the competitiveness of manufacturing companies raises the need to describe the PP of the entire end-to-end order processing process (ETEOPP) [3]. The ETEOPP includes all technical-operative core processes, reaching from sales processes and manufacturing processes to shipping processes, and describes the sequence of operational processes transforming customer inquiries into saleable products [4]. Notwithstanding, as 96% of process optimisation projects are realized in manufacturing processes, most ETEOPP sub-processes are disregarded in PP descriptions. However, disregarded sub-processes make up 70% of the end-to-end process time. As a result, not taking the ETEOPP into consideration results in crucial non-transparencies for PP improvements [5]. Limitations are biased participants, large time consumption and limited abilities to capture process dynamics in paper-based techniques [6]. Further industry insights show that describing the ETEOPP is a significant problem. 62% of companies have only documented less than 25% of their processes and only 2% of companies have an overview of their complete process landscape [7].

Process Mining (PM) can be applied to tackle deficits in process descriptions with a fact-based, objective and precise method. PM aims to discover, monitor and improve business processes using event data stored in event logs. However, PM has only been applied to single departments and partial processes with similar order types, respective order-IDs [8]. A three-phase framework is introduced in previous work to address PM in ETEOPP and shows the impact of data-based approaches on process analysis [3]. This paper provides an approach to merge multiple order types and calculates PPI as well as process models to expand the second phase of the framework. The remainder of this paper is structured as follows: Chapter two outlines the importance of PM for order processing. Chapter three presents the methodology for merging multiple event logs to apply PM across the ETEOPP. Chapter four validates the methodology using a dataset. In the fifth chapter, the results of the paper are summarised and an outlook on further research is given.

Manuscript received December 11, 2021; revised January 10, 2022.

doi: 10.18178/joams.10.1.1-8

II. Importance of Process Mining for Order Processing

Due to diverse order types, parallel or sequential activities in ETEOPP, process variances are often higher than assumed in manufacturing companies [9]. In the following, prerequisites for the application of PM in ETEOPP are outlined.

Process discovery as one type of PM algorithmically converts event log data into a process model [10] and quantifies indicators such as frequency, duration or throughput times. Regarding the ETEOPP, process models must display event data emerging from different departments of a company. However, event data of the ETEOPP are scattered across multiple information systems such as Customer Relationship Management systems, Enterprise Resource Planning systems and Manufacturing Execution Systems [11]. Thus, data from multiple information systems must be defined in a data model and merged in an event log before PM techniques can be applied.

In the ETEOPP, order-IDs appearing in events can be categorised by different object types (OT). Each OT characterises orders that are processed in partial business processes. For instance, customer orders (order-IDs of sales processes as one OT) could contain several articles represented by multiple manufacturing orders (order-IDs of manufacturing processes as a second OT). A customer order can be split and joined in various OT throughout the ETEOPP. Resulting multiple order-IDs must be considered as process instances for evaluable results of PP descriptions across the ETEOPP [12].

The eXtensible Event Stream (XES) is the common format for event logs and PM applications but only represents one single OT [13]. A different format is required to represent multiple OT for an ETEOPP. An Object-Centric Event Log (OCL) combines multiple OT within a single data table [14]. In this paper, an OCL is a two-dimensional, column-structured table with multiple OT (respective order-IDs), related activities and timestamps as data attributes [14]. This enables the tracing of orders with multiple order-IDs across processes. However, describing the PP requires transforming the OCL into an XES-structured data table to apply traditional PM algorithms.

In industry, widespread uncertainty exists regarding the suitability of available data for data-based analysis [15]. Thus, data requirements for data-based PP descriptions must be clearly defined. For PM applications in ETEOPP it is assumed that, according to the first guiding principle [16], partial event logs are available in sufficient quality (i.e. without noise). Exception is syntactic data inhomogeneity, which results from merging multiple event logs of different information systems. Therefore, an application of PM must consider appropriate data preparation to improve the quality of resulting process models. Lastly, PPI that describe process efficiencies must be calculated for processes, traces and activities.

III. Traditional Methods for Process Descriptions

In this paper, interview-, workshop- or evidence-based methods for process description are summarised as traditional methods. Interview-based methods capture the process by interviews or questionnaires. In workshop-based methods, the process is described through participatory workshops with employees. Evidence-based process descriptions rely on existing evidence such as documentation or process observations [2]. In the following, constitutional characteristics of the methods Architecture of Integrated Information Systems (ARIS) [17], Value Stream Design (VSD) [18], Business Process Model and Notation (BMPN) [19] and aixperanto [20] are derived. These methods represent a sufficient sample of traditional methods due to their widespread notoriety and focus on manufacturing processes. Additionally, each method is characterised by its languages for process description.

A. *Subjective and Unsubstantiated Process Descriptions*

The four methods use interview- and workshop-based principles of process discovery. Additionally, VSD and ARIS apply evidence-based observations to describe processes. Thus, the process descriptions are influenced by natural subjectivity due to human participants' involvement [21]. The subjectivity leads to biased interpretations regarding deviations from real process behaviour and relevant information might be opportunistically hidden by participants [2]. In addition, the process descriptions always just represent a random sample of all existing process behaviours [8]. As a result, the drawbacks are unsubstantiated and subjective process descriptions.

B. *Description of Process Model and PPI*

Process descriptions based on ARIS can be complemented by information objects in order to create transparency and to identify improvement potentials [17]. VSD describes process performance by process descriptions and additional PPI such as throughput time, processing time and waiting time [22]. Aixperanto also creates process descriptions and corresponding PPI during the current state analysis to identify weaknesses. Therefore, a process model and PPI, which together describe the PP, serve for process optimization.

C. *Time-consuming Procedures for Single Process Variants*

ARIS, VSD, aixperanto and BPMN are time-consuming methods, as they involve several participants in interviews and workshops for a longer time. Additionally, evidence-based methods are time-consuming due to process observations and persons may have perceptions of how the process operates, which may be partially incorrect [2]. Lastly, the four methods result only in process descriptions as a snapshot of existing processes. Thus, documented process descriptions need to be updated continuously. In VSD the most expensive step is the description of as-is process performance [23]. As the effort of process descriptions scales with the number of process variants, traditional methods mostly describe non-representative processes and unobserved process behaviour leads to 80% of the problems [10]. Therefore, traditional methods are economically not efficient for a representative process description.

D. Consistent Visualization for the Description of Order Processing

ARIS uses event-driven process chains (EPC) for process visualization that consists of events, functions, rules and resources [24]. The use of EPC for process description by the German company SAP AG has established the method in practice. In VSD, standardised and simple visualizations are used for material flows, while process performance is described in data boxes [18]. For aixperanto, existing methods were adapted to ensure easy-to-use process descriptions. Elements of every process description are process cards that contain the process name and further performance indicators [25]. The company IBM developed and established the BPMN method, whose visualisation includes activities, events, decisions, control flows, connections and data objects. Like ARIS and aixperanto, sub-processes can be visualized in swimlanes that enable a structured description of order processing. As a result, consistent visualizations and modelling languages make process descriptions comparable and evaluable.

E. Integration of Context Knowledge

The methods ARIS, VSD, aixperanto and BPMN are based on direct conversation and interaction with participants in workshops and interviews. Thus, an integration of process-specific context knowledge is possible [2]. Ref. [26] motivates the integration of multiple process participants (e.g. workshops in VSD, aixperanto or BPMN), as no individual can have a holistic overview of the as-is processes. By doing so, feedback about consistent or inconsistent process behaviour is possible due to the integration of context knowledge. This creates a common understanding and commitment for later process improvement. As a result, the integration of context knowledge ensures the right judging of process performance descriptions.

Fig. 1 shows an overview of the constitutional characteristics of the four traditional methods that were considered. On the one hand, the four traditional methods repre-sent the drawbacks of subjective and unsubstantiated process description as well as time-consuming procedures for single process variants. On the other hand, the methods show the characteristics to describe PP based on process model and PPI, in a consistent visualization as well as in a valuation-neutral way to integrate context knowledge. Those drawbacks and characteristics as well as the requirements elaborated in the second chapter need to be addressed with a suitable approach that is described in the following chapter.

IV. METHODOLOGY

The proposed methodology considers multiple event logs and varying OT to describe the ETEOPP by mapping a process model and calculating PPI. The development is based on existing research to be rigor. First, [33] split up an OCL into an event log for every OT through flattening to apply established PM techniques. Second, describing PP requires mapping of as-is processes and calculating PPI. Ref. [34] discovers a process model and enhanced it using separately calculated PPI before displaying results to the user. In [35], time-based PPI are calculated for the categories process, case and activity.

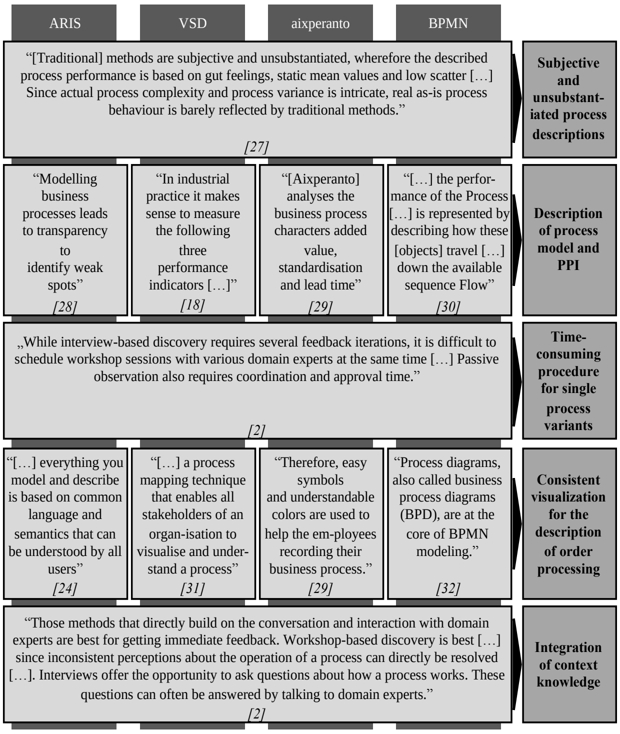

Figure 1. Comparison of the constitutional characteristics of traditional methods [28-32].

Fig. 2 gives an overview of the six-step methodology. Chapter IV-A describes the data tables (DT) in detail as inputs of the methodology. First, the DT are combined into an OCL. Second, the OCL with multiple OT is split into event logs for each OT. Third, event log traces are identified. In the remaining steps, the PPI and process model of the ETEOPP are calculated separately. Thus, fourth, PPI for the activity perspective are calculated. Fifth, PPI for the trace and process perspectives are calculated. Sixth, a process model for the smallest sub-instance OT of the ETEOPP is calculated. The outputs of the methodology are PPI for the perspectives activity, trace and process as well as a process model of the ETEOPP to describe the PP of manufacturing companies.

A. Detailed Description of the DT as Inputs for the Methodology

Each DT is an extracted event log of a partial, department-specific process within a company’s ETEOPP (e.g. sales, manufacturing etc.). A DT is a two-dimensional, column-structured table with order-IDs as process instances as well as their related activities and timestamps as data

attributes. The DT are comparable to the XES-Standard. The timestamps must record the start, the end and the planned end of the activity as well as the time when the order was received. These timestamps are necessary for calculating the PPI for the ETEOPP, which are elaborated in chapter IV-B. The extraction and filtration of the DT from information systems are out of scope of the methodology.

B. Detailed Description of Step One to Three of The Methodology

In step one, the DT are merged to an OCL and extended to trace the ETEOPP from the viewpoint of every OT. To map the ETEOPP, related objects across all OT need to be identified. Two objects across different OT are related to each other, if they occur in the same event within the OCL. The OCL is extended so that every time an object-related order-ID is treated within one event, related objects are complemented to the event. In this paper, the enriched OCL is called End to-End OCL (E2EL). An example of the extension from OCL to E2EL is shown in Fig. 3. In the E2EL, the order numbers 990001 and 990002 can be traced when their related shipping order was packaged in the third event. As a result, the ETEOPP of e.g. an order number can be mapped correctly so that it also includes the packaging activity besides the initiation of an order.

In step two, the E2EL is split into a DT for each OT. Each OT is selected as a case notion and the E2EL is flattened towards the selected OT. Flattening leads to the three problems divergence, convergence, and deficiency [33]. Divergence is the loss of ordering information leading to loops in the process model that do not exist in the real process. Chapter IV-D addresses the divergence problem. Convergence is the replication of an event that is executed for multiple objects, falsifying the real number of events. Chapter IV-C further deals with the convergence problem. Deficiency describes the disappearance of events, which do not include objects of the selected OT. The E2EL diminishes deficiency, as the number of OT included in every event is increased. The outputs of the second step are DT for every OT of the entire ETEOPP.

The input of step three are the resulting DT from step two. The DT events are separated according to their objects to create corresponding traces. By doing so, all event attributes are kept such that no information is lost during this step. The output of step three are traces for every object of the event log. The existence and placement of step three are justified due to the following reasons: First, an event log for every OT is required as input for step six, such that the step cannot be merged with step two. Second, this step prepares the data while step five calculates PPI. Thus, separating both steps allows a better understanding and distinction of the steps of the methodology.

C. Detailed Description of Step Four and Five of the Methodology

In step four, the PPI from the activity perspective are calculated. Input for step four is the E2EL. The E2EL is not modified by flattening, so the activity PPI are not affected by convergence. The five PPI process time, time of response, deadline adherence, mean tardiness and process reliability are calculated based on previous works [27]. In this paper, the calculations for the process time and the deadline adherence are further elaborated. Equation (1) depicts the calculation $PPI_{PT,a}$ for process time PT of activity a. Therefore, the sum over all events E in the E2EL is taken. Each event is filtered for the inquired activity using the expression in equation (2). The process time for each event i is calculated by subtracting the start timestamp TS_i from the end timestamp TE_i. Thus, the process time of an activity is the average duration of all instances of that activity.

$$\mathrm{PPI_{PT,a}}=\frac{\sum_{i}^{E}(x_{i,a}\times(TE_i\text{-}TS_i))}{\sum_{i}^{E}x_{i,a}} \text{ with } \mathrm{PPI_{PT,a}}\in[0,\infty)\ \forall a \quad (1)$$

$$x_{i,a}=\begin{cases}1;\ \text{Event i includes activity a}\\0;\ \text{Event i does not include acitivty a}\end{cases}\quad \forall i,a \quad (2)$$

Equation (3) shows the calculation of $PPI_{DA,a}$ for the deadline adherence DA of an activity based on equation (2). Equation (4) checks if an event i has been completed on time by comparing the end timestamp TE_i with the planned end timestamp TP_i.

$$\mathrm{PPI_{DA,a}}=\frac{\sum_{i}^{E}(x_{i,a}\times y_i)}{\sum_{i}^{E}x_{i,a}} \text{ with } \mathrm{PPI_{DA,a}}\in[0,1]\ \forall a \quad (3)$$

$$y_i=\begin{cases}1;\ TE_i\text{-}TP_i\le 0\\0;\ TE_i\text{-}TP_i>0\end{cases}\quad \forall i \quad (4)$$

In step five, PPI for the trace and process perspectives are calculated. Equation (5) displays the calculation of $PPI_{PT,j}$ for the process time PT for the trace of an object j. Each object j has a trace with several events E_j. $PPI_{PT,j}$ is calculated based on two timestamps that are differentiated by two indices. The first index refers to the object of the trace, the second to the position of the event of this object in the trace. Consequently, $TS_{j,1}$ is the start timestamp of the first event in the trace of object j. TE_{j,E_j} is the end timestamp of the last event in the trace of object j.

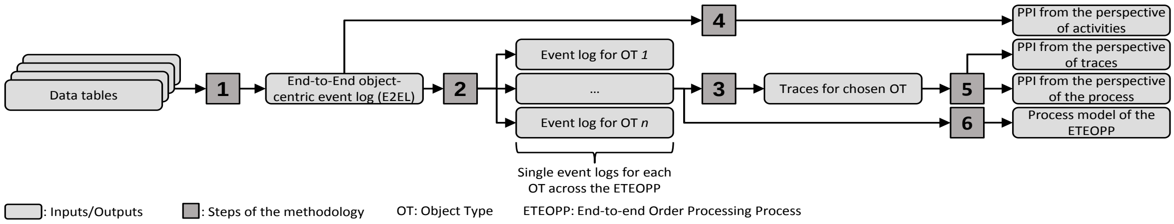

Figure 2. Six-step methodology as well as inputs and outputs.

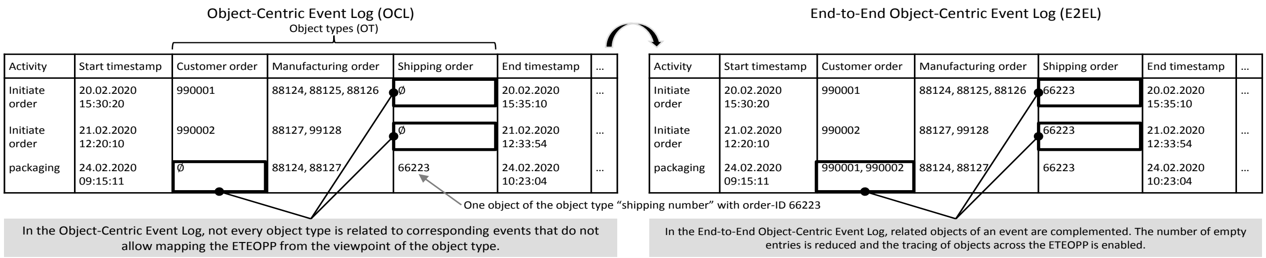

Figure 3. Exemplary visualisation of the extension from OCL to E2EL.

$$PPI_{PT,j}=TE_{j,E_j}-TS_{j,1} \text{ with } PPI_{PT,j} \in [0,\infty)\ \forall j \tag{5}$$

Equation (6) is the calculation of the process time $PPI_{PT,p}$ for the process. The process has several traces T. The process time for the process perspective is the average of all process times of the traces in that process (see equation (5)).

$$PPI_{PT,p}=\frac{\sum_{j=1}^{T} PPI_{PT,j}}{T} \text{ with } PPI_{PT} \in [0,\infty) \tag{6}$$

Equation (7) shows the calculation for the deadline adherence $PPI_{DA,j}$ for the trace of an object j. The deadline adherence for a trace is the fraction of events in the trace, completed within the planned time frame. Equation (8) compares the end timestamp $TE_{i,j}$ with the planned end timestamp $TP_{i,j}$ of the event i within the trace of object j.

$$PPI_{DA,j}=\frac{\sum_{i=1}^{E_j} y_{j,i}}{E_j} \text{ with } PPI_{DA,j} \in [0,1]\ \forall j \tag{7}$$

$$y_{i,j}=\begin{cases}1; TE_{i,j}-TP_{i,j} \leq 0\\ 0; TE_{i,j}-TP_{i,j} > 0\end{cases} \text{ for } \forall i,j \tag{8}$$

Equation (9) presents the calculation for the deadline adherence $PPI_{DA,p}$ for the process. The deadline adherence for the process is the fraction of traces, of which the last event was completed within the planned timeframe. This is calculated using the expression in equation (8), whereby y_{j,E_j} compares the end timestamp TE_{i,E_j} with the planned end timestamp TP_{i,E_j} of the last event E_j within the trace of an object j.

$$PPI_{DA,p}=\frac{\sum_{j=1}^{T} y_{j,E_j}}{T} \text{ with } PPI_{DA,p} \in [0,1] \tag{9}$$

D. *Detailed Description Of Step Six*

Step six of the methodology uses a discovery algorithm to map the process model of the ETEOPP. The aim is to create transparency of the ETEOPP and to put the calculated PPI into a context. As the popular discovery algorithms cannot deal with multiple OT, a DT of step two is chosen as input. Additionally, independent from the OT viewpoint of which the PP of the ETEOPP should be described and the PPI are calculated, the input for step six must be the DT with the smallest sub-instance OT of the ETEOPP. In a manufacturing company, a product is represented by an article. The OT customer order might contain multiple articles per object, which disqualifies the DT of customer orders as input for step six. Otherwise, if products need to be manufactured one by one, the OT manufacturing order would contain exactly one article per object. The DT of manufacturing orders would then qualify to be selected as an input for step six. In industry, an OT, which contains one article per object, can be defined as the smallest sub-instance OT of the ETEOPP. The selection of the smallest sub-instance OT addresses the divergence problem on process discovery. The convergence problem persists, such that some process instances are duplicated when flattening towards the OT of an article, respective manufacturing order.

As a result, the flattened DT has more events than the original process. This replication of events is acceptable because the PPI are calculated separately and the resulting process model does not display the number of events.

The herein used discovery algorithm is interchangeable as the selection of a suitable discovery algorithm depends on the requirements and data [36].

V. Introducing the Case Study and Validation of the Methodology

The methodology is validated with a dataset. The dataset is based on three order types (i.e. OT) processed through an exemplary ETEOPP depicted in Fig. 4. The process shows various tasks across the departments sales, manufacturing and shipping. The process includes parallel and sequential activities, OR-splits, AND-splits and loops of various lengths to test the robustness of the methodology.

The departments record their activities using different OT. The objects of all OT contain manufacturing orders as the smallest sub-instance as defined in chapter IV-D. A manufacturing order only includes one article, customer orders and shipping orders include one or more articles. Thus, customer orders and manufacturing orders are related one-to-many (1:n), which means that a customer order contains multiple manufacturing orders. Shipping order and manufacturing order are related many-to-one (n:1), which means that multiple manufacturing orders are shipped in the same shipping order. The OT customer order and shipping order are related many-to-many (n:n). In practice, two customer orders are shipped to the same address across three shipping orders.

The OT are processed in the ETEOPP, as shown in the conceptual object-centric Petri net in Fig. 5. An object-centric Petri net extends a regular Petri net by shading transitions and places based on the OT they refer to. Places and arcs of transitions consuming multiple objects are highlighted by double lines [14]. Due to the incomprehensibility of object-centric Petri nets in practice, more intuitive visualizations and established process models (respective miners) are used for the case study.

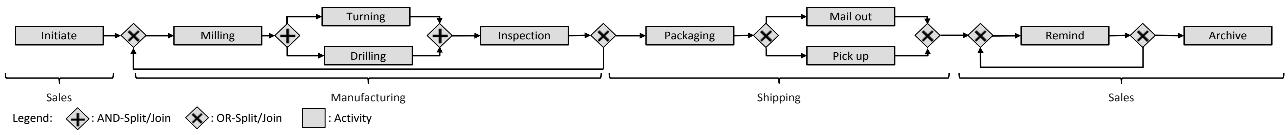

Figure 4. BPMN model of the ETEOPP throughout three departments.

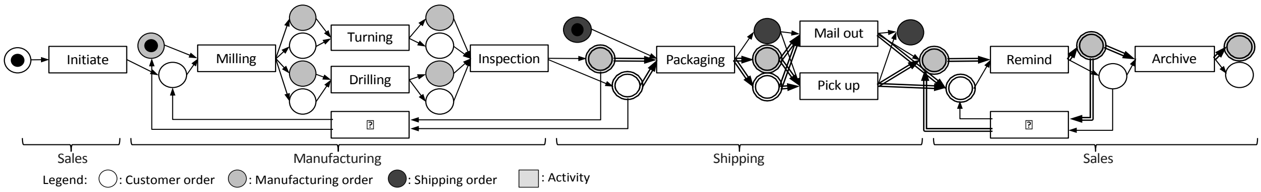

Figure 5. Object-centric Petri net of the ETEOPP, showing the OT treated in each activity.

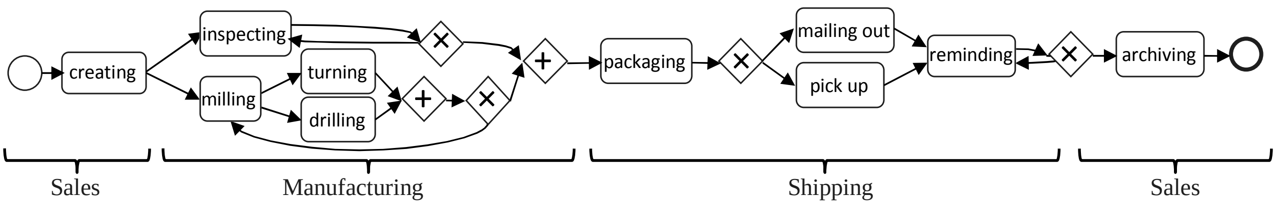

Figure 6. Process model mapped using the inductive min and the OT manufacturing order.

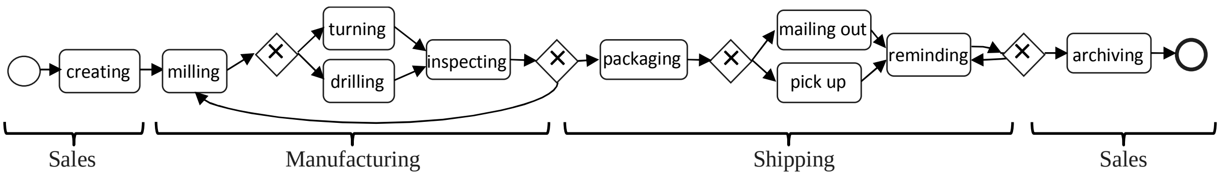

Figure 7. Process model mapped using the heuristic miner and the OT manufacturing order.

The dataset comprises three DT with 41 events involving two customer orders, five manufacturing orders and three shipping orders. Table I shows the first line of the DT from the manufacturing department. Here, the customer order is recorded as data for every activity.

TABLE I. FIRST LINES OF THE DT FROM MANUFACTURING

Manuf. order	Activity	Start	End	Planned End	Order received	Customer order
M378	Milling	11.02. 09:00	11.02. 14:40	11.02. 15:00	05.02. 14:10	13623

Tables I to IV show the resulting PPI for the process time and deadline adherence for selected activities, objects and OT based on the equations 1 to 9. The process time of traces is large compared to the process time of activities, partly because time outside of work shifts were not excluded from the calculations.

TABLE II. PPI FOR THE MILLING ACTIVITY

Activity	Process time [h]	Deadline adherence
Milling	6.73	0.50

TABLE III. PPI FOR THE TRACE OF M28910 FOR THE OT MANUFACTURING ORDER

Object of traces	Process time [h]	Deadline adherence
M28910	773.33	0.54

TABLE IV. PPI OF PROCESSES FOR OT MANUFACTURING AND CUSTOMER ORDER

Process	Process time [h]	Deadline adherence
Customer order	738.67	0.50
Manufact. order	731.73	0.60

Fig. 6 shows the process model, which was mapped using the DT of the OT manufacturing order. This DT was chosen according to chapter IV-D, the resulting process model is valid for evaluating PP independent of the OT chosen to calculate PPI. The event log was extended for process discovery to include 123 events to approximate a bigger dataset. The software ProM 6.9 and the plug-in *Mine* process *tree with Inductive Miner* followed by the plug-in *Convert Process tree to BPMN diagram* were used to describe the ETEOPP process model. The process model is under-fitting. The activity *inspection* is a successor of the activity *initiate*, which is not possible in the real process. One reason is the inductive miner and its trade-off between under-fitting process models and preserving fitness. Here, the heuristic miner was able to produce a better fitting process model (see Fig. 7). For this, the plug-ins *Heuristic net*, *Convert Heuristic net into Petri net* and *Convert Petri net to a BPMN diagram* were applied.

VI. Summary and Research Outlook

This paper demonstrated a methodology for the application of PM in ETEOPP. The six contributing steps merge event logs from companies' information systems into an E2EL and use the results for calculating PPI and discovering the process model. The novelties are the consolidation of multiple event logs of the ETEOPP and the use of an OCL to deal with multiple order types in production companies in the context of PM. Thus, analysis of the ETEOPP can be based on facts and exempt from employees' subjectivity and other external factors. This enables long-term and continuous improvement of PP in projects commencing with the description of as-is PP. An application of the methodology on a dataset results in a visualisation of the ETEOPP process model and calculated PPI.

The presented methodology expands the second step of a broader approach presented in [3]. As an outlook, the preceding and subsequent steps of the broader approach need to be elaborated before integrating the separate parts into a holistic solution for describing PP in ETEOPP. In particular, an approach for defining the data requirements for the DT from software infrastructure and a user interface to operate the methodology and display the results need to be developed. Furthermore, applications with real company data would have the potential of uncovering development potential.

Next, a data-based approach for process acquisition can always be assisted by classical participative methods, since it helps detecting hidden activities or inefficiencies and further improvement potentials that are not stored in a company's software infrastructure.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Seth Schmitz, Calvin Kuhn and Noah Klapper conducted the research, analysed the data, drafted the manuscript, designed the figures and wrote the paper. Professor Günther Schuh and Andreas Gützlaff provided the expert guidance regarding research and analyzations. The authors discussed the results, commented on the manuscript and approved the final version.

Acknowledgement

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC-2023 Internet of Production – 390621612.

References

[1] A. Del-Río-Ortega, M. Resinas, and A. Ruiz-Cortés, "Defining process performance indicators, an ontological approach," *On the Move to Meaningful Internet Systems (OTM)*, vol. 6426, pp. 555-572, 2010.
[2] M. Dumas, M. La Rosa, J. Mendling, and H. A. Reijers, *Fundamentals of Business Process Management*, Berlin Heidelberg: Springer, 2018.
[3] G. Schuh, A. Gützlaff, S. Schmitz, and W. Van Der Aalst, "Data-based description of process performance in end-to-end order processing," *CIRP Annals - Manufacturing Technology*, vol. 69, pp. 381-384, 2020.
[4] W. Eversheim, S. Krumm, T. Heuser, and S. Müller, "Process-oriented organization of order processing – A new method to meet customers demands," *CIRP Annals - Manufacturing Technology*, vol. 42, pp. 569-571, 1993.
[5] P. Schönsleben, S. Weber, S. Koenigs, and D. Aldo, "Different types of cooperation between the R&D and engineering departments in companies with a design-to-order production environment," *CIRP Annals - Manufacturing Technology*, vol 66, pp. 405–408, 2017.
[6] G. Schuh, A. Gützlaff, S. Cremer, and M. Schopen, "Understanding process mining for data-driven optimization of order processing," *Conf. on Learn Factories*, pp. 417-422, 2020.
[7] P. Harmon and J. Garcia, "The BPTrends Report," *The State of Business Process Management*, 2020.
[8] L. Reinkemeyer, *Process Mining in Action*, Berlin Heidelberg: Springer, 2020.
[9] M. Pospíšil, V. Mates, T. Hruska, and V. Bartik, "Process mining in a manufacturing company for predictions and planning*," Int. J. on Adv. in Software*, vol. 6, no. 3&4, pp. 283-297, 2013.
[10] W. V. D. Aalst, *Process Mining*, Berlin Heidelberg: Springer, 2016.
[11] W. V. D. Aalst, A. K. Alves de Medeiros, and A. J. M. M. Weijters, "Generic process mining," in *Proc. 26th Int. Conf. on Applications and Theory of Petri Nets*, pp. 48-69 2005.
[12] T. Thaler, S. Ternis, P. Fettke, and P. Loos, "A comparative analysis of process instance cluster techniques," *AISeL Wirtschaftsinformatik Proceedings*, pp. 423-437, 2015.
[13] C. W. Günther and E. Verbeek, *XES Standard Definition*, TU Eindhoven, 2014.
[14] W. van der Aalst and A. Berti, "Discovering object-centric petri nets," *Fundamenta Informaticae XXI*, pp. 1001-1042, 2019.
[15] B. Marr, *Big Data: Using SMART Big Data, Analytics and Metrics to Make Better Decisions and Improve Performance*, Hoboken NJ USA: John Wiley & Sons, 2015.
[16] W. van der Aalst W. et al., "Process mining manifesto," *Business Process Management Workshops. BPM 2011. Lecture Notes in Business Information Processing*, pp. 169-194, 2012.
[17] A. W. Scheer, H. Kruppke, W. Jost, and H. Kindermann, *Agility by ARIS Business Process Management: Yearbook Business Process Excellence*, Berlin Heidelberg: Springer, 2006.
[18] K. Erlach, *Value Stream Design: The Way Towards a Lean Factory*, Heidelberg New York: Springer, 2013.
[19] M. V. Rosing, S. White, F. Cummins, and H. De Man, "Business process model and notation – BPMN," *The Complete Business Process Handbook*, pp. 429-453, 2015.
[20] G. Schuh, T. Potente, F. Bachmann, and T. Froitzheim, "An integrated approach: Combining process management, organizational structure and company layout," *Proc. CIRP sponsored Conf. RoMaC*, pp. 481-494, 2012.
[21] W. an der Aalst, B. F. van Dongen, J. Herbst, L. Maruster, and A. J. M. M. Weijters, "Workflow mining: A survey of issues and approaches," *J Data & Knowledge Engineering*, vol. 47, pp. 237-267, 2003.
[22] N. V. K. Jasti, S. Kota, and K. S. Sangwan, "An application of value stream mapping in auto-ancillary industry: A case study," *The TQM Journal*, vol. 32, pp. 162-182, 2019.
[23] D. Knoll, G. Reinhart, and M. Prüglmeier, "Enabling value stream mapping for internal logistics using multidimensional process mining," *Expert Systems with Applications*, vol. 124, pp. 130-142, 2019.
[24] R. Davis, *Business Process Modelling with ARIS: A Practical Guide*, London Berlin: Springer, 2001.
[25] G. Schuh, *Lean Innnovation*, Berlin Heidelberg: Springer, 2013.

[26] F. Milani, *Digital Business Analysis*, Cham: Springer Nature Switzerland, 2019.

[27] S. Schmitz, F. Renneberg, S. Cremer, A. Gützlaff, and G. Schuh, "Definition of process performance indicators for the application of process mining in end-to-end order processing processes," in *Proc. 10th Cong. of the German Academic Association for Production Technology*, pp. 670-679, 2020.

[28] A. W. Scheer, *ARIS - Business Process Frameworks: Third Edition*, Berlin Heidelberg: Springer, 1999.

[29] G. Schuh, W. Boos, and K. Kuhlmann, "Modelling business processes for limited-lot producers with aixperanto," *IEEE Management and Service Science,* 2010.

[30] S. A. White and D. Miers, *BPMN Modelling and Reference Guide*, Lighthouse Point: Future Strategies Inc, 2008.

[31] M. A. Nash and S. R. Poling, *Mapping the Total Value Stream: A Comprehensive Guide for Production and Transactional Processes*, New York: Taylor & Francis Group, 2011.

[32] V. Stiehl, *Process-Driven Applications with BPMN*, Cham Heidelberg: Springer, 2014.

[33] W. van der Aalst, "Object-centric process mining: Dealing with divergence and convergence in event data software engineering and formal methods," *SEFM*, pp. 3-25, 2019.

[34] S. Leemans, E. Poppe, and M. T. Wynn, "Directly follows-based process mining: Exploration & a case study," *International Conference on Process Mining (ICPM)*, pp. 25-32, 2019.

[35] N. Zaki, A. Awad, and E. Ezat, "Extracting accurate performance indicators from execution logs using process models," in *Proc. IEEE/ACS 12th Int. Conf. of Computer Systems and Applications*, vol. 1, pp. 1-8, 2015.

[36] T. Jouck, A. Bolt, B. Depaire, M. de Leoni, and W. van der Aalst, "An integrated framework for process discovery algorithm evaluation," *IEEE Transactions on Knowledge and Data Engineering ArXiv*: 1806.07222v1, 2018.

Günther Schuh was born in 1958 in Germany. He studied Mechanical Engineering and Business Administration at the RWTH Aachen University, Germany. He was awarded a PhD in 1988 after a period as a research associate at the Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen University and was employed as senior engineer until 1990. In 1990, he took up a full-time post as lecturer in Manufacturing and Industrial Management at the University of St Gallen (HSG), Switzerland. Three years later, in 1993, he was appointed Professor of Industrial Production Management and, at the same time, became a member of the Board of Directors at the Institute for Technology Management. He returned to the RWTH Aachen University in 2002 to take up the position of Chair of Production Systems whilst also serving as a Member of the Board of Directors of the Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen University and of the Fraunhofer IPT in Aachen. Since October 2004, he had additionally been a director of the Institute for Industrial Management (FIR). Professor Günther Schuh is founder and Managing Director of several industrial companies.

Andreas Gützlaff was born in 1989 in Germany. He studied Mechanical Engineering and Business Administration at the RWTH Aachen University in Germany and was awarded with a master degree in 2015 after a period as a research assistant at the Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen University. From 2015 to 2019 he worked as a research associate, group leader and PhD at the Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen University. Since 2019, he is chief engineer and head of department production management at the Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen University.

Seth Schmitz was born in 1991 in Germany. He studied Mechanical Engineering and Business Administration at the RWTH Aachen University in Germany as well as Management Science and Engineering at the Tsinghua University in Beijing, China. He was awarded with two master degrees in 2018 and 2019. In between, he worked for 3,5 years as a research assistant at the Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen University as well as in large manufacturing companies. Since 2018, he works as a research associate and PhD at the Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen University and is head of group Process Management.

Calvin Kuhn was born in 1998 in Germany. He studied Mechanical Engineering at the RWTH Aachen University in Germany and National Tsing Hua University in Taiwan. He was awarded with a bachelor degree in mechanical engineering in 2020. In between, he worked as a research assistant at the Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen University. Since 2020, he studies for a master degree in production engineering with a specialization in corporate organization at the RWTH Aachen University.

Noah Klapper was born in 2000 in Germany. He worked as a working student at the Digital Capability Center (DCC) Aachen as well as in several industrial companies. He studies Mechanical Engineering at the RWTH Aachen University in Germany and works as a research assistant at the Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen University since 2017.