

Data-Driven Automatic Repair Order Scheduling in the Aerospace Industry

Felix Schumann¹, Michel Kunkler¹, Silvia Burger² and Stefanie Rinderle-Ma¹

¹ Chair of Information Systems and Business Process Management,
TUM School of Computation, Information and Technology,
Technical University of Munich
85748 Garching, Germany

² Liebherr-Aerospace Lindenberg GmbH,
88161 Lindenberg im Allgäu, Germany

Abstract

Adherence to delivery deadlines is critical for organizational success. In execution environments such as repair shops, the prioritization of open orders significantly impacts due-date performance. This research develops an automated task scheduling approach driven by organizational data to prioritize repair orders. We present a pipeline that transforms process event log data into a mathematical optimization problem while accounting for inherent process uncertainties. Adopting a “data-first” methodology, we prioritize data discovery and insight generation before developing the final optimization artifact. Our findings indicate that while a “data-first” approach enhances process transparency, the resulting optimization and prediction quality are highly sensitive to data quality. The case study reveals that process mining can be improved through effective event abstraction, yet challenges remain with sparse data and the verification of optimization models when applied to real-world datasets.

1. Introduction

Liebherr-Aerospace is a leading global supplier of integrated systems and components for the aerospace industry. Liebherr-Aerospace’s portfolio includes the manufacturing and maintenance of a plethora of different aircraft components used in both civil and military aircraft. Their regular maintenance, repair, or overhaul (MRO) plays a critical role in sustaining aircraft safety and is therefore required by strict aviation regulations. Aircraft operators (e.g., commercial airlines) are therefore regularly required to send various components of their aircraft for MRO; hence, Liebherr-Aerospace operates several MRO shops worldwide. This case study focuses on the operations within a single MRO shop at one of Liebherr-Aerospace’s locations, which handles parts for one component domain. For aircraft operators, it is of utmost importance to have their shipped-in components returned in a planned and timely manner, as this relieves them of the need to stockpile large quantities of replacement parts and minimizes unscheduled aircraft-on-ground time. Therefore, aircraft operators typically maintain contracts with Liebherr-Aerospace that stipulate MRO within a specified number of days (turnaround time). The aircraft components to be repaired that arrive at an MRO shop undergo a range of activities, including an initial inspection, an actual maintenance or repair activity, scrapping (per the customer’s decision), and, after maintenance or repair, a functionality test. A process case must be completed within the turnaround time, but the duration of the individual activities depends strongly on the type of the component to repair. Moreover, some components are more prone to failing the functionality test, which consequently leads to a repetition of the repair and another functionality test.

Prior to our study, the planning of activities within the MRO shop was conducted manually, primarily based on domain knowledge. This state-of-the-art planning sometimes led to missed due dates. We suspect that optimization potential can be lifted by leveraging combinatorial optimization approaches.

Proceedings of the Industry Forum (BPM 2026), Toronto, Canada, September 27th to October 2nd, 2026.

*Corresponding author.

✉ felix.schumann@tum.de (F. Schumann); michel.kunkler@tum.de (M. Kunkler); silvia.burger@liebherr.com (S. Burger); stefanie.rinderle-ma@tum.de (S. Rinderle-Ma)

ORCID: 0009-0006-5431-1922 (F. Schumann); 0000-0002-1920-7322 (M. Kunkler); 0000-0001-5656-6108 (S. Rinderle-Ma)



© 2026 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

In our study, we developed a prototypical artifact that i) automates the planning of activities within the MRO shop that is under investigation in our case study, and ii) optimizes the planning towards reduced turnaround time deviations through more efficient scheduling. In this work, we reflect on a successful strategy for applying business process management (BPM) techniques, pinpointing pitfalls and discussing gaps in BPM research that emerged during the development of the artifact. The remainder of this paper is structured as follows: In Sec. 2 we describe the problem statement in detail, in Sec. 3 we describe the methodological framework that we applied for the development of our artifact. In Sec. 4 we outline our findings, including successes as well as pitfalls during the project journey. Sec. 5 iterates on the relevance of this industry study and Sec. 6 describes the scope and limitations of this work. Lastly, in Sec. 7 we discuss our findings and conclude this work.

2. Problem Description

Currently, the MRO shop we focus on at Liebherr-Aerospace schedules activities through a mix of manual assignment and a pull principle, where workers decide for themselves which task to pick up next. The current practice has three main disadvantages: i) the manual scheduling requires a high amount of manual effort; ii) the manual scheduling can only be performed by a few workers with years of hands-on experience, which creates a bottleneck for the scheduling process; and iii) the scheduling might not exploit combinatorial efficiencies, which may lead to a higher number of turnaround time deviations.

Implementing an automated scheduling solution for the MRO shop is not trivial because the scheduling problem is based on a complex process with a high degree of uncertainty. Moreover, the scheduling solution has to be embedded into an existing IT infrastructure (brownfield development) with different data-sources with unknown relevance of the individual data-sources and unknown data-quality. In the following, we will describe the MRO process for the shops domain in detail.

The MRO process for the repair shop under investigation: Once a customer turns in a component, Liebherr-Aerospace creates a *repair dossier*, which resembles a process case and for which the contracted turnaround time must be adhered to. The component is received at a goods-in and routed to the MRO shop. In the MRO shop, the component is first tested, then disassembled, and based on the findings from the initial test, a *repair offer* is created and sent to the customer. At the time the repair offer is sent, the turnaround time counter is paused, and it only restarts once the customer accepts the repair offer ("Response to Repair Offer" in Fig. 1). Once the customer accepts the *repair offer*, the MRO shop conducts the repair, assembles the component, and subsequently runs a functional test. The functional test can fail, leading to a repeat of the repair, assembly, and functional test. Lastly, once the part passes the functional test, it is sent back to the customer. A case-centric view of the process is shown in Fig. 1. All activities in the MRO shop are manual and executed by employees, who are the main resources for the scheduling problem. Due to holidays or sickness, the number of available employees varies significantly every day, and within each day, the MRO shop's employees work flexi-time. Moreover, not all employees in the MRO shop are qualified to execute all activities. Two main types of activities exist, testing activities and repair activities, and the employees are typically qualified for one of these two types of activities, but not for both. The MRO process is subject to a high degree of control-flow and activity duration uncertainty. The control-flow uncertainty is caused by the redo loops of the repair-functional test sequence. The repair of a part can require rework if it fails the functional test. The activity durations depend strongly on the type of the component. For each activity, a standard duration for each component type is known, but the actual duration can differ from the standard duration.

3. Methodology

This study follows the design science research methodology (DSRM) proposed in [1] in which the design and development of an artifact is the core of the research. The DSRM process model consists of six phases

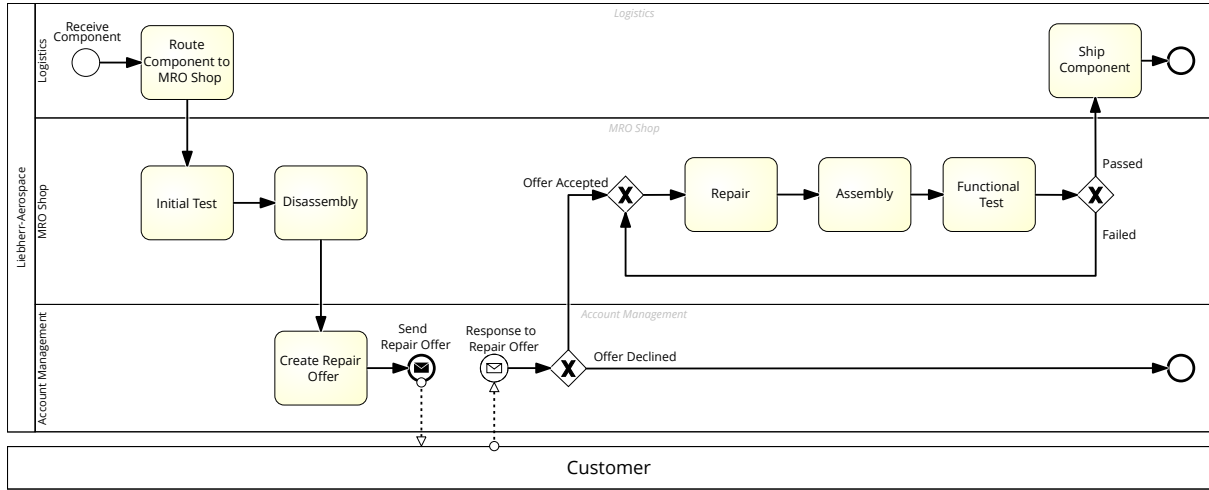


Figure 1: MRO process for the shop domain at Liebherr-Aerospace.

(problem identification and motivation, definition of objectives for a solution, design and development, demonstration, evaluation, and communication) that are iteratively executed and four research entry points (problem-centered initiation, objective-centered initiation, design- and development-centered initiation, and a client-/context-centered initiation). As the problem and the objectives of a solution were defined by the industry partner, this study adopts a design- and development-centered initiation, driven by the industry need to develop a scheduling artifact that addresses the turnaround time deviations observed in Liebherr-Aerospace's MRO repair process. In the following, we address the design and development phase of our artifact, which is the core of our study.

3.1. Design and Development

Core to the design and development of the artifact is the question of how the uncertainties in the repair process can be addressed. On the one hand, this means how the artifact can react to uncertain events, e.g., unplanned loops, absence of workers, or changing execution times, and on the other hand, how the artifact can predict such events to proactively take them into account in the scheduling process. Our approach follows a data-first perspective, which means that we use as much data as possible to predict uncertainties and to react to them in the scheduling process.

Project Initialization: In the initialization phase, the underlying information systems, their data, and their semantic representation are analyzed. The relevant sources are identified in workshops and by interviewing domain experts and validated using traditional data analysis and process mining techniques, e.g. process discovery and conformance checking. The outcome is that the repair process is tracked in two different IT systems, one for the activities within the MRO shop itself and one for the whole administrative processes, e.g., repair offer creation. The data is stored in a data warehouse (DWH), which is updated regularly. For the administrative tasks, only the end timestamps are available. For the repair process itself, start and end timestamps are available, which allows us to train a model for predicting the durations of activities in the MRO shop.

Event Log Creation: In order to exploit existing process mining techniques, we merge the data from the two relevant IT systems into a process event log reflecting the MRO process. The event log serves for two purposes: first, it allows us to train machine learning models for predicting the future course of a case and quantify uncertainties, and second, it allows us to get the current state of a case, which serves as the input for the scheduling process.

Uncertainties: The problem encompasses different kinds of uncertainties. Of special importance are the uncertainties of the resource availability, e.g., due to unplanned absences of workers, the uncertainties of when new cases arrive, the uncertainties of the activity durations, and the uncertainties of the activity sequences, e.g., due to unplanned loops. The resource availability is addressed by a daily

manual input of the responsible person of the repair shop, who provides a calendar specifying the resource availabilities for each day in minutes. The arrival of new cases is addressed by the design decision to re-running the scheduling formulation daily before the shift starts in the morning, thereby taking the new cases into account. The uncertainties of the activity durations and sequences are addressed by a data-driven approach, which means that we use the historical data to train a model to predict the future course of a case and quantify the uncertainties. To this end, we apply the probabilistic suffix prediction approach presented in [2], which allows us to predict the future course of a case and quantify the uncertainties of the activity sequences and durations. The probabilistic suffix prediction approach allows to choose between a conservative or a less conservative scheduling approach, by applying a different aggregation schema, e.g. taking the most common suffix, or the longest suffix.

Modeling of the optimization problem We create a scheduling formulation that takes a predicted suffix scenario, and for each activity, its execution durations as input, along with the worker availabilities (for the testers and for the repair workers) for a given day, to create a prioritized list of tasks to be handled that day. Moreover, we also take the due dates of the cases into account. The scheduling formulation schedules a single activity per case and day, which allows us to create a prioritized list of tasks that need to be executed sometime on a given day. Thereby, we address the uncertainty due to the flexi-time availability of the workers in the formulation itself. The scheduling formulation models the described problem as a multi-knapsack problem, where each day is represented as two knapsacks, one for the repair workers and one for the analysis workers. The scheduling and prioritization formulation uses Google OR-Tools' CP-Sat solver¹.

Business Process Optimization Pipeline: Combining these steps from event log generation to the finalized schedule creates a pipeline for such a Business Process Optimization (BPO) Problem, which is shown in Fig. 2. Triggered once a day, the pipeline starts with pulling the data from the DWH, and the preprocessing step transforms the transactional data into an event log. From this event log the suffix prediction approach predicts the future course of the open cases and quantifies the uncertainties. Then it feeds this information into the scheduling formulation together with the worker availabilities, which are read from an Excel sheet, and creates a prioritized list of tasks for the workers, which is then also stored as an Excel sheet and sent to the responsible person of the repair shop, who then distributes the tasks to the workers.

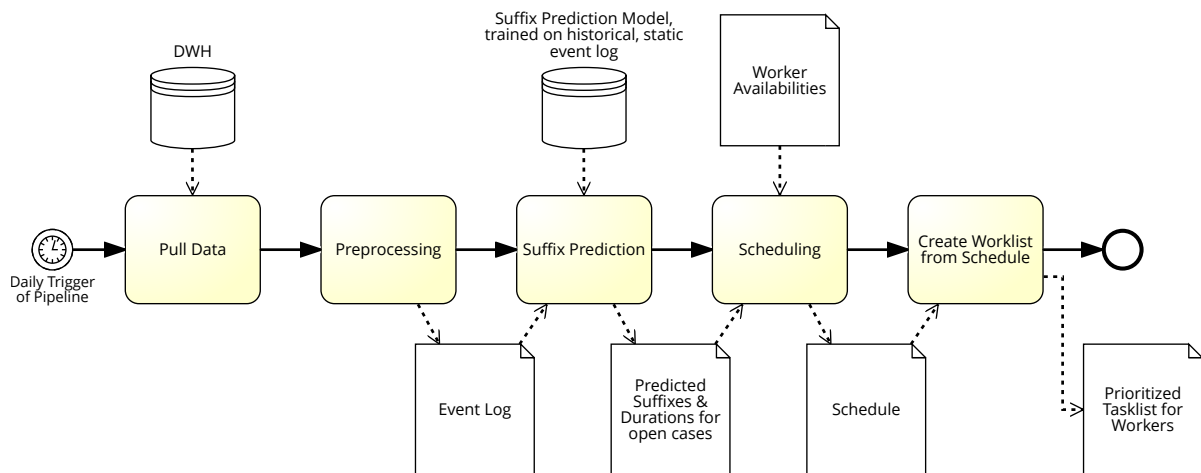


Figure 2: Representation of the BPO-Pipeline: from event log to a prioritized task list for Workers

4. Findings

Throughout this industry study, we applied multiple approaches that have been introduced in the field of BPM and BPO within the last few years. These include the formalization of a mathematical optimization

¹<https://developers.google.com/optimization/cp>

model to solve a decision-making problem, the usage of a suffix prediction approach to predict the process continuation while accounting for process-driven uncertainty, and the application of process mining to analyze and understand the repair process itself and its interaction with the back-office tasks, e.g., writing the repair offers and orders. We approached the problem from a data-first perspective, i.e., using data-driven methods, such as process mining, to generate knowledge about the optimization problem, its necessary inputs, and the constraints that must be considered. These findings are also confirmed with domain experts.

Event Log Data Structure and Process Mining: The event log serves as the overarching data structure used throughout this project. It enabled the use of process mining tools, such as PM4Py [3], and a direct connection to the suffix prediction approach. Further, the event log allows for storing information as part of a single event or for the full case. The usage of start- and end-timestamps is a crucial input for the scheduling formulation, as it represents the task durations.

Although we were able to apply process discovery on the generated event log, the resulting process models were not specifically insightful and required further manual adaptation. One reason is the varying levels of abstraction and granularity in the event log (cf. [4]) due to its generation from two different data sources. In order to address this issue and to generate a valuable process model, a manual adaptation, incorporating feedback from domain experts, becomes necessary.

Gap: An approach that allows a user or domain expert to specify and validate the required abstraction level is needed. When presenting the mined process model to domain experts, they identified superfluous tasks that did not provide any additional knowledge about the optimization problem. With current methods, this adaptation of both the event log and the process model had to be done manually, e.g., by removing tasks from the event log.

Data Quality for Suffix Prediction: Suffix prediction approaches such as [2] require conclusive data for each possible scenario represented in the event log. In our case study, this means a sufficient number of cases for each repaired product type is required to enable good predictions. The tasks in the MRO process result in different execution times for each product type. For components that are repaired only in a few cases, only a small amount of data is available, which hampers the quality of the predictions. Further, we found that the time-tracking system, based on manual inputs, records timestamps at a different granularity level than the tasks that are represented in the process model. In terms of predicted task durations, this led to shorter predicted times than those actually executed, while, at the same time, more loops were predicted. As the suspension of work on a component is tracked in the same manner as a failure, it was unclear whether a task was suspended due to a process failure requiring rework or to environmental factors such as a lunch break.

Gap: Elaborated preprocessing steps are needed to differentiate between required rework loops and other task suspensions. The preprocessing steps were rule-based and manually derived from the data, based on dependencies in the control-flow and resource perspectives in a case. Again, an automated or user-guided approach to identify rules for the preprocessing of an event log is needed.

Evaluation of combinatorial optimization approaches: In theoretical works on BPO, evaluation is either achieved by running simulations on existing event logs [5, 6] or by specifying optimality gaps to the found solutions [7]. The high level of uncertainty in our case study limits the expressiveness of both evaluation approaches in this use case. Through simulation, not all influencing factors on the repair of a part can be covered. External factors, such as the unavailability of workers due to sickness, cannot be well simulated. While the formulated optimization model can be solved to optimality by combinatorial solvers, it is hard to assess how well the prevalent uncertainties are covered in the daily knapsack bucketing approach.

Gap: How to evaluate the potential that can be achieved through the application of combinatorial optimization models, and how to ensure that the application of optimization approaches leads to better deadline adherence, is an open research topic that needs to be addressed, especially since, in industry use cases, using a misaligned optimization model can lead to a negative business outcome.

5. Relevance

In recent years, BPO approaches focusing on resource allocation and scheduling have received close attention in the BPM community [8, 5, 6, 9, 7]. The majority of these papers provide theoretical approaches for BPO. These are evaluated on BPIC Datasets² yet are typically not applied in a real-world setting. This work closes this gap by applying BPO methods in an industry project focused on scheduling and identifies open research gaps in the transfer from theoretical work to industry adoption. Therefore, this work focuses on the design and development, and demonstration steps of DSRM.

As organizations mature in their BPM capabilities, they move towards more proactive BPM methods such as predictions and prescriptions [10]. As seen in this case study and also identified in other industry collaborations, organizations move from creating process transparency to improving the execution of processes. The expectation is to move from pull-based task allocations with a local optimization goal towards the use of optimization algorithms with a global scope to leverage global optimization effects. This work provides insights into the potential and pitfalls that can arise in the development of such global optimization approaches. Moreover, it assesses if a data-first perspective using process mining can help to enable better communication between domain experts and developers

6. Scope and Limitations

The scope of this work focuses on testing novel BPO approaches in a real-world environment. It aims to identify how a data-first approach to building an optimization model can be leveraged successfully to create an optimization model based on process data. To analyze this setting, the scope ranges from identifying the data sources, over to analyzing the data with established methods such as process mining, to using that data to derive predictive models and build a prescriptive optimization model. The goal was to build a prototype that can be tested in the real-world environment. In contrast to other case studies, the goal of this study is not to accompany the integration of an existing software tool in an organization's environment, but rather to design a novel solution approach based on the data sources available. As part of this project, a prototype was developed and used for demonstration. In the DSRM demonstration phase, it has been iteratively improved by incorporating client feedback. Yet, one limitation of this work is that the designed software artifact is not used in the daily operations of Liebherr-Aerospace yet.

Further, as already stated in [9], the evaluation and comparison of optimization approaches, especially in running systems, is particularly hard. This also accounts for this case study. The problem of optimization verification is even more challenging, as many external factors impact the due date adherence, such as the current backlog of items to repair, the availability of spare parts, and the availability of workers. Thus, the findings focus on the data-first approach and the quality of the suffix prediction.

7. Conclusion

This case study in the aerospace industry aimed to build a scheduling prototype in a data-driven manner and to analyze how such an approach can support the development of a mathematical optimization model.

We found that using a data-first approach helps during implementation to quickly identify variants and special execution cases relevant to scheduling that were not named in the initial workshops with domain experts. Also, focusing on identifying a process model helps to generate a common understanding among the heterogeneous team members involved in the project. Yet, looking only at data is insufficient. Manual input is required to simplify and streamline process models due to incoherent abstraction and granularity levels and to remove unnecessary noise.

²<https://www.tf-pm.org/competitions-awards/bpi-challenge>

We can also see that the identified process models, in combination with the suffix predictor, can serve as valuable input to a scheduling algorithm. For the predictive approaches, data quality is very important and requires careful handling in our case study. After deploying the algorithm to the real-world process and testing it, it became apparent that measuring the effectiveness of the applied reordering of tasks is highly complex due to the large amount of external factors influencing operational performance. How to quantify the performance improvement produced through the optimization approach remains an open challenge in real-world settings.

Future work should focus on finding better methods for evaluating the quality of the recommended schedules and further improving the consideration of uncertainties within scheduling models. By enabling users to guide preprocessing steps based on a process model, the data used as input for the optimization step could be filtered more easily.

Declaration on Generative AI

During the preparation of this work, the author(s) used ChatGPT, Claude, Grammarly in order to: Grammar and spelling check, Paraphrase, and reword. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the publication's content.

References

- [1] K. Peffers, T. Tuunanen, M. A. Rothenberger, S. Chatterjee, A design science research methodology for information systems research, *J. Manag. Inf. Syst.* 24 (2008) 45–77. doi:10.2753/MIS0742-1222240302.
- [2] M. Kunkler, H. Moustroph, S. Rinderle-Ma, Probabilistic suffix prediction of business processes, in: *Process Mining, IEEE*, 2025, pp. 1–8. doi:10.1109/ICPM66919.2025.11220650.
- [3] A. Berti, S. van Zelst, D. Schuster, Pm4py: A process mining library for python, *Software Impacts* 17 (2023) 100556. doi:https://doi.org/10.1016/j.simpa.2023.100556.
- [4] I. Beerepoot, C. D. Ciccio, H. A. Reijers, S. Rinderle-Ma, W. Bandara, A. Burattin, D. Calvanese, T. Chen, I. Cohen, B. Depaire, G. D. Federico, M. Dumas, C. G. J. van Dun, T. Fehrer, D. A. Fischer, A. Gal, M. Indulska, V. Isahagian, C. Klinkmüller, W. Kratsch, H. Leopold, A. V. Looy, H. A. López, S. Lukumbuzya, J. Mendling, L. Meyers, L. Moder, M. Montali, V. Muthusamy, M. Reichert, Y. Rizk, M. Rosemann, M. Röglinger, S. Sadiq, R. Seiger, T. Slaats, M. Simkus, I. A. Someh, B. Weber, I. Weber, M. Weske, F. Zerbato, The biggest business process management problems to solve before we die, *Comput. Ind.* 146 (2023) 103837. doi:10.1016/J.COMPIND.2022.103837.
- [5] F. Meneghello, A. Senderovich, M. Ronzani, C. Di Francescomarino, C. Ghidini, Proactive data-driven scheduling of business processes, in: *Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence, IJCAI '25*, 2025. doi:10.24963/ijcai.2025/953.
- [6] J. Middelhuis, R. L. Bianco, E. Sherzer, Z. A. Bukhsh, I. Adan, R. M. Dijkman, Learning policies for resource allocation in business processes, *Inf. Syst.* 128 (2025) 102492. doi:10.1016/J.IS.2024.102492.
- [7] F. Schumann, G. W. van der Heijden, S. Rinderle-Ma, Instance configuration and scheduling based on the resource-augmented process structure tree, in: *Business Process Management Forum*, 2025, pp. 110–127. doi:10.1007/978-3-032-02929-4_7.
- [8] M. Kunkler, F. Schumann, S. Rinderle-Ma, Business process optimization: A systematic review of combining concepts from business process management and operations research, in: *Enterprise Design, Operations, and Computing*, 2025, pp. 61–83. doi:10.1007/978-3-032-15140-7_4.
- [9] L. Pufahl, F. Stiehle, S. Ihde, M. Weske, I. Weber, Resource allocation in business process executions - A systematic literature study, *Inf. Syst.* 132 (2025) 102541. doi:10.1016/J.IS.2025.102541.
- [10] R. Poll, A. Polyvyanyy, M. Rosemann, M. Röglinger, L. Rupperecht, Process forecasting: Towards proactive business process management, in: *International Conference on Business Process Management*, Springer, 2018, pp. 496–512.