EVALUATION OF PROCESS MINING FOR A TYPICAL PLANT PLANNING TASK

Pierre Grzona [1], Luise Weißflog [1]

[1] Univiersity of Technology Chemnitz - Professorship of Factory Planning and Intralogistics, Erfenschlager Straße 73, 09125 Chemnitz
pierre.grzona@mb.tu-chemnitz.de
luise.weissflog@mb.tu-chemnitz.de

Annotation: The paper gives an introduction into the field of factory planning with the specific task of the evaluation of the logistics costs for an existing system with the help of VDI-4405 and the usage of Process Mining. The actual challenges in intralogistics are shown and is explained how in a traditional manner a material flow analysis can be performed. Based on these foundations a new concept for the incorporation of Process Mining into an existing material flow analysis workflow has been elaborated with the help of an industry 4.0 data driven monitoring concept. In a case study the planning efforts for an interview and type representative material flow modelling is then compared with the efforts needed to fulfil the same task with the help of Process Mining.

1 Introduction

Since the onset of the first industrial revolution in the 18th century, the world has continually grappled with balancing the need to increase production to match growing demand with dwindling natural resources [1]. Manufacturing is a longstanding field that has been integral to daily human life. It has assisted humanity not only through the provision of daily utilities but also through significant indirect contributions to social and economic spheres. The effects transcend the mere simplification of daily life. It holds substantial sway over the nation's GDP and impacts variables pertaining to living standards, science, technology, the environment, and national security, to name a few [2]. The recent social changes have brought about various transformations in the field, resulting in an increase in customer demand diversity as well as an intensely competitive market environment [3].

The integration of digital technology in numerous daily activities has resulted in significant transformations in various aspects of life [4]. Covid-19 has accelerated the implementation of digital technologies in daily life [5]. The utilization of digitalization in manufacturing has brought about substantial alterations to the industry, creating competitive advantages for businesses [6], but has also resulted in changes of the needed skill sets and competencies of industrial engineers [7].

https://doi.org/10.24132/PI.2023.11986.134-146

The fourth industrial revolution, known as Industry 4.0, is having a transformative impact on the world. Its introduction in 2011 sparked significant interest from the public, industries, and governments globally [6]. This palpable phenomenon is not a mere buzzword, as supported by its increasing influence [8–10]. Industry 4.0 and its associated digital transformation are advancing rapidly [11]. This technological revolution has considerably changed the way individuals and organizations work and think. It has generated optimism by opening up possibilities for future opportunities. Big data analytics, supply chain management, the internet of things, artificial intelligence, cybersecurity measures, augmented reality, cloud computing, new business models, horizontal and vertical integration, additive manufacturing, and autonomous robotics are commonly cited as the essential components of Industry 4.0 [12, 13].

For complete optimization and automation of production, it is essential to integrate these pillars into the production process. This represents a notable progress towards enhanced production efficiency and addressing the varying requirements of customers. The combination of accepted industrial engineering methods with state of the art technologies can help to overcome the challanges in Small- and Medium Sized Enterprises (SME) [14].

2 Challenges in Intralogistics planning

2.1 Challenges in intralogistics

The manufacturing process involves several activities, from the sourcing of raw materials to the actual production of goods and their final delivery to consumers. These activities need to be optimised through manufacturing and logistics operations [15]. Logistics is all the transport activities associated with the organisation. Intralogistics mainly deals with the activities related to internal transport within the physical boundaries of a company. An optimised intralogistics system in terms of higher operational performance, sustainability and uptime is required to realise and maintain higher business value [16]. Every industry is witnessing digitisation and adopting it as a standard in this new era of Industry 4.0. The main reason for this is the need to remain competitive, meet customer demands and be efficient [17]. Technology is the main means of delivering Industry 4.0 solutions. This trend can also be seen in the intralogistics sector, as intralogistics is a very important area for the application of Industry 4.0 technologies [18]. Intralogistics has always been the data-driven part of the business and has become more important than ever. Industry 4.0 has provided the prospect of live tracking of product and material flows along with risk management and improved transport management [19]. It might be argued that the vision of Industry 4.0 can only become a reality when intralogistics department of a system can become capable enough to provide the inputs that are required not only to the physical layer of the material flow [18].

The competitive advantage of any industry depends on its intralogistics system's capacity to quickly adapt to changing environments. However, achieving this is challenging due to the high instability and dynamics present in most intralogistics systems. Companies can enhance their flexibility by correctly identifying the actual circumstances. The intralogistics process and its deviation from planned processes in a fluctuating environment may result in a significant number of discrepancies between real and intended processes. Knowledge and precision are essential in addressing this issue. Nevertheless, modern technologies potentially solve the challenge of keeping data up to date.

2.2 Traditional planning approaches

According to Schenk et. al the „Logistics processes represent a significant cost factor“ in modern factory planning and operations whereas the requirements for logistics are quite complex beneath the classical seven rights of logistics [20, 21]. In general, the aim of logistics planning should be:

- Development of suitable solutions for a (logistical) problem
- Demonstration of their suitability
- Formulation of specifications for several following activities in a factory planning project (procurement, assembly, construction, operations, etc.)

To reach these aims a 6-stepped material flow planning approach is suggested (Figure 1)

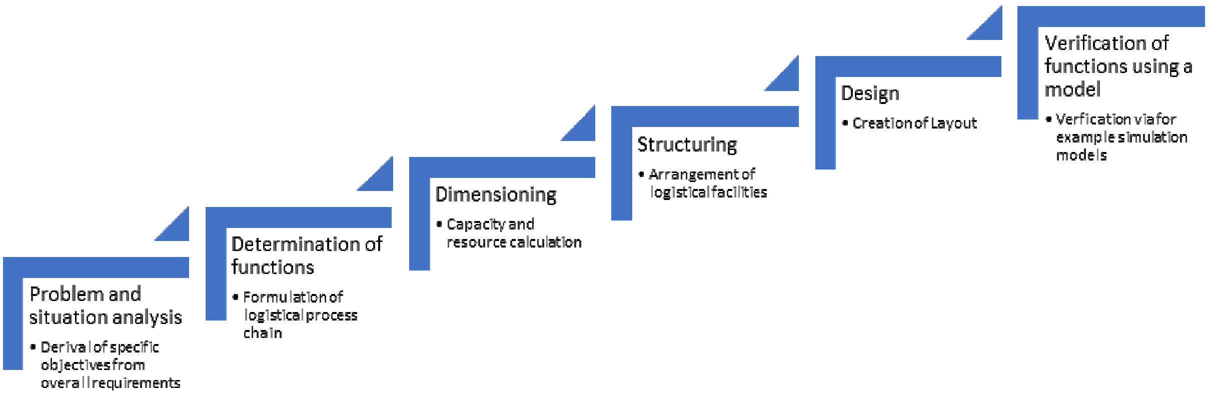

Figure 1: Material flow planning approach [20]

During the 2nd phase the selection of the facilities should follow the rules of efficiency and effectiveness. Additionally, the material flow should be calculated. To represent these, different types of graphical models could be used like graphs, flow or block diagrams. For a mathematical description formulas, connection or transport matrix might be used. To conclude which type of material handling equipment (MHE) fulfils the need of efficiency a cost analysis in form of process cost analysis can be used [20].

A methodology for the calculation of cost is the Activity-based cost analysis in internal logistics (Figure 2) [22]. Typical methods which might be used for the data collection are employee surveys, writing down, assessment and estimation, observations, production data acquisition, time measurement or systems of pre-determined time.

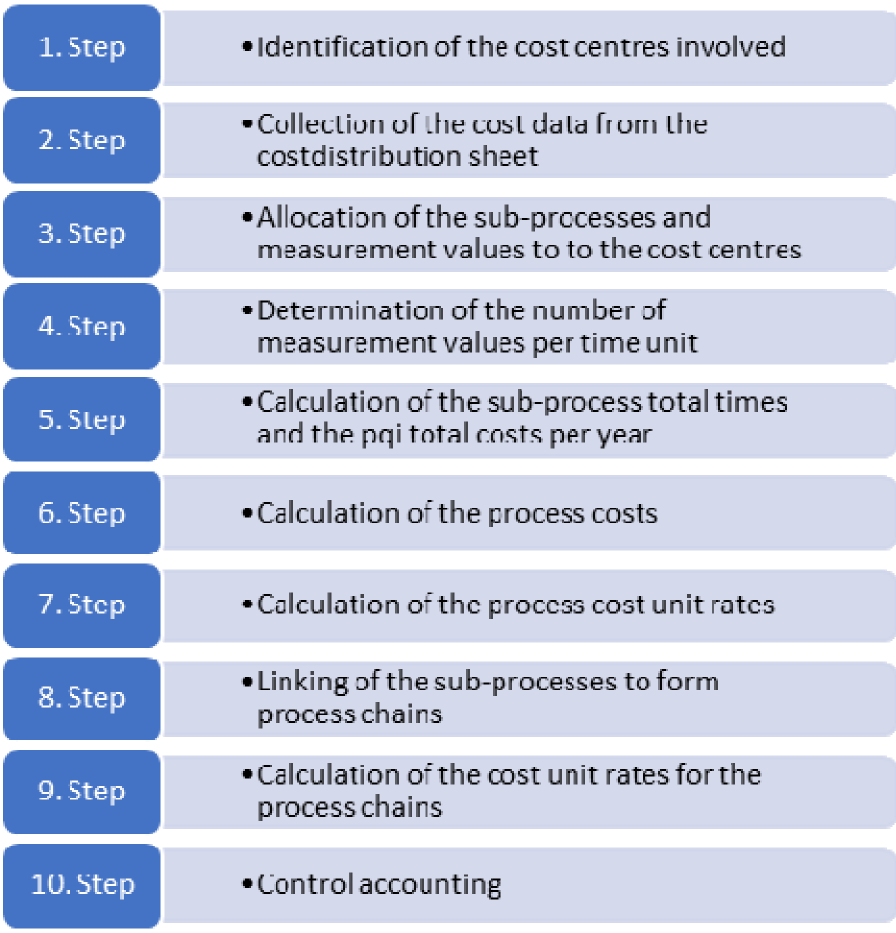

Figure 2: Steps of Activity-based cost analysis [22]

Steps 1 & 2 will define the cost centres involved and cost data so will therefore define the scope of investigation, comparable to the 1st planning step. Most crucial are the Steps 3 to 6 with the definition of the sub-processes and their boundaries, the specific calculation of needed times per investigation period and the aggregation of the sub-process cost to the process costs. To gain a a comparability of these activities the boundaries needs to be chosen similar between the different variants of logistical ressources which needs to be evaluated according to the planning approach. Due to the thinking of logistical process chains this leads to the opportunity of an easy comparison of ideas. The next 3 steps in the cost analysis will lead to the possibility of an overall cost comparison and should be finished by the comparison with other cost data sources. [22]

With the help of the two beforehand mentioned methods confident statements of planned logistical processes and systems can be made.

2.3 Methodological approach and Research questions

In the context of this paper, a methodological approach is employed, drawing upon a case study to elucidate a complex relationship, both providing a tangible explanation and facilitating a practical comparison through a real-world corporate example [23]. The procedure is therefore based on the following process for the selected methodology.

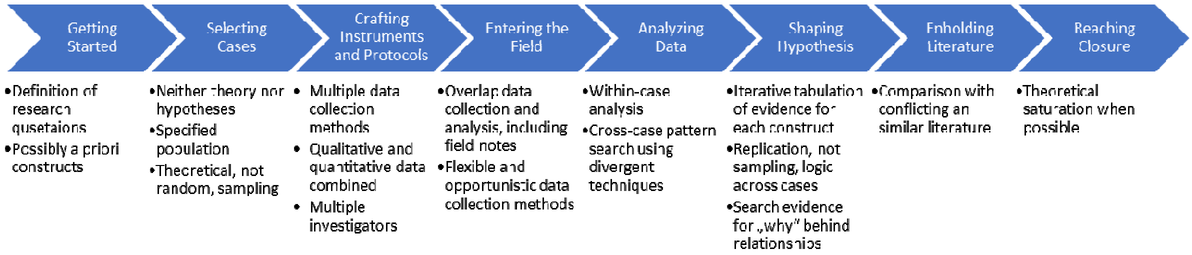

Figure 3: Process for a Case Study Research [24]

This framework enables the derivation of pertinent conclusions to address the research questions outlined below:

- How can data-driven approaches be leveraged in material flow analysis? (RQ 1)
- What additional value do data-driven material flow analyses offer? (RQ 2)

Through the examination of a specific case study, the paper aims to not only illustrate the application of data-driven methodologies in material flow analysis but also to discern the intrinsic benefits that such approaches bring to the understanding and optimization of material flows within operational contexts.

3 Development of a data driven material flow assessment

3.1 Data driven tools and trends in manufacturing

The types and possibilities of data analysis have increased significantly in recent years. The availability of data is growing in most companies and previous boundaries are opening up in various places, so that a precisely defined exchange of data takes place along a supply chain, for example, as in the automotive and aerospace industries. However, antitrust law and data protection often set strict limits that must be adhered to. Based on recording and thus the availability of the correct data, there are a wide range of possible uses. This starts with the generation of information through pattern recognition to forecasting and ultimately a recommendation for action (see Figure 4).

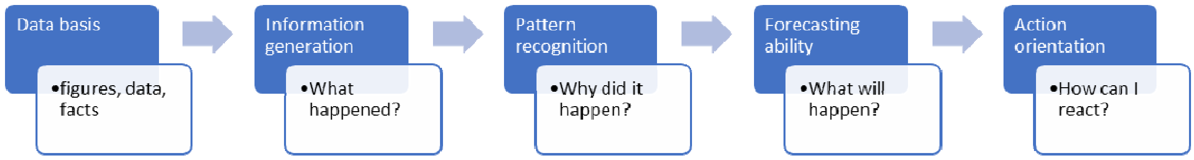

Figure 4: Possible uses of data mining [25]

On the other hand, processes are also becoming increasingly complex, meaning that conventional analysis tools are reaching their limits. As a result, it is gradually becoming essential to use data-based process analytics methods to carry out targeted process analyses. This means that it is necessary to understand what data science and data mining means and how it can be linked to the specific field of process analysis and optimization. [25]

In simple terms, data mining means extracting knowledge from data. Depending on the objective, various methods are used, such as cluster analysis, classification, association analysis or text mining [26]. As a rule, selective results can be expected. In contrast, the aim of process mining is to generate corresponding processes [25].

Prof. van der Aalst, who is considered the inventor of process mining, defines process mining as "[...] the missing link between data science and process science [...]" [27]. This means that a link is created between process models and the associated process knowledge and data-centred analysis methods such as data mining. The aim is to generate automated process models and workflows on this basis and to identify potential for improvement. So this is a technique that offers the possibility of automatically analysing processes. This analysis is based on event logs, which contain a series of events that are recorded during the course of the process. Each event contains a case associated with the execution of an activity. An event log can also store additional information such as the resources used or the costs, etc. Process mining helps various industries to gain data-based insights into different processes and enables further improvement of processes. Process mining analysis can be categorized into three types, namely process discovery, process compliance review and process improvement. [27]

- The first type deals with the topic of discovery. This means that processes are to be created automatically that are not yet known or modelled in this way. Additional information is not included.
- The second type is called conformance testing. This involves comparing known or predefined target processes with the actual processes based on the event logs and identifying any deviations.
- In the third and currently highest expansion stage, so called extension, the aim is to automatically generate improved process models based on the known processes in combination with the event data. [25]

The areas of application for process mining are diverse. An essential prerequisite is the availability of the required data in the right quality and quantity. For this reason, its use is most pronounced in very data-intensive business processes such as purchasing, procurement and invoice settlement (purchase-to-pay or procure-to-pay (P2P)) or in sales (order-to-cash (O2C)). As data collection is also continuously increasing in production and logistics, the proven and potential areas of application are also constantly growing here. However, as process mining is not yet as widespread in intralogistics as in the other examples mentioned, the following use case is intended to focus precisely on this and determine its suitability, particularly in comparison to the classic approach. [28]

3.2 Concept of a Data driven planning approach

Data driven monitoring or analysis platforms in the context of Industry 4.0 might consist of a 4-layer systematics. Typical layers are:

- The Physical layer consisting of the sensors for the data acquisition and actuators
- The Data Aggregation layer which contains the elements for integration, storing and structuring the data
- The Data Analytics incorporating the different elements of the data processing modules for the assessment
- The Presentation layer for presenting the obtained results from the data of the lower layers [29]

Figure 5 shows this concept incorporated into the above-mentioned planning approaches (not including the last steps) according to Schenk et. al and VDI-4405. Whereas the traditional planning approaches relies on interviews and data from limited quality which is then for example integrated with the help of product representatives into the planning premises for a generic data analysis typically done in some kind of standard spreadsheet software. With the new proposed concept based on the usage of process mining it could enhance the data which is used in a Material Flow Analysis Software for the further usage during the logistics analysis. Then the existing planning workflow and capabilities of the Analysis software could be used, which includes the Analysis of Material Flows, Transport Matrix, Cost-Model and Layout representation.

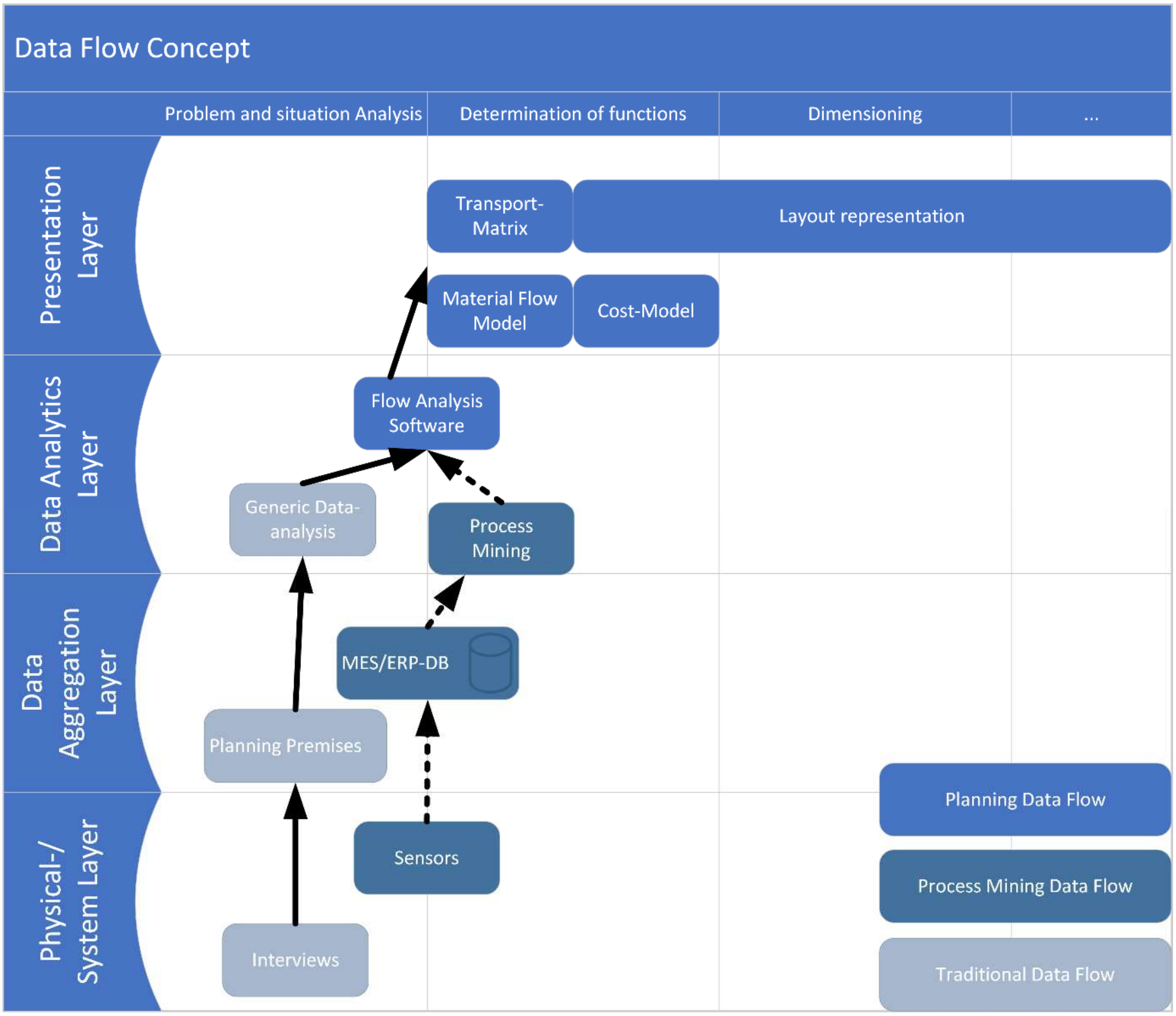

Figure 5: Data Flow Concept

4 Evaluation based on a case study

4.1 Comparison with traditional approaches

The proof of the concept was done by the involvement of a regional manufacturer of speciality nonwovens and composites. The company is a supplier of nonwovens for cosmetics, construction and special cleaning industry. With an assortment of more than 400 products there is a certain degree of transparency needed in the historical grown logistical processes. To increase the transparency and lay the basis for the decision based on the logistical challenges it was necessary to create the Material Flow Model (MFM) to evaluate the usage of other MHE-Technologies.

The typical approach for gathering the data for the MFM is to reduce the number of products to investigate with the help of the type representative method. The challenge here is to determine the right number of representatives with a trade-off between level of detail and their relevance in the production program [20]. In this case it was determined that 3 products represented the overall production best. They consist of either high volume but also logistical high intensity products. Figure 6 shows the example of on product. Due to the limitations of the interview situation and the resource time, the flow was only feasible to document with a range of a percentage distribution. Also, during the interview, it was difficult to stay to the exact product representative because not all of the 5 participants shared the same view of the process because of the flexibility of the manufacturing process different routes were possible. For the logistics the planning premises were defined with the average speed of 3,6 km/h per Forklift and un-/loading times of 60 seconds. The interviews were conducted as one site meeting with 6 persons for 3 hours, an additional online modelling session with 5 persons for 2 hours and a validation session with 5 persons for 1 hour.

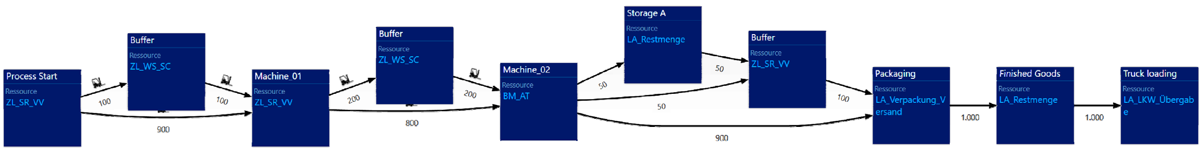

Figure 6: Example of on product

4.2 Usage of Process Mining

For the verification of the modelled material flow model the parallel path of information gathering with the help of PM was used. Figure 7 shows the specific approach for PM according to *Van der Aalst*.

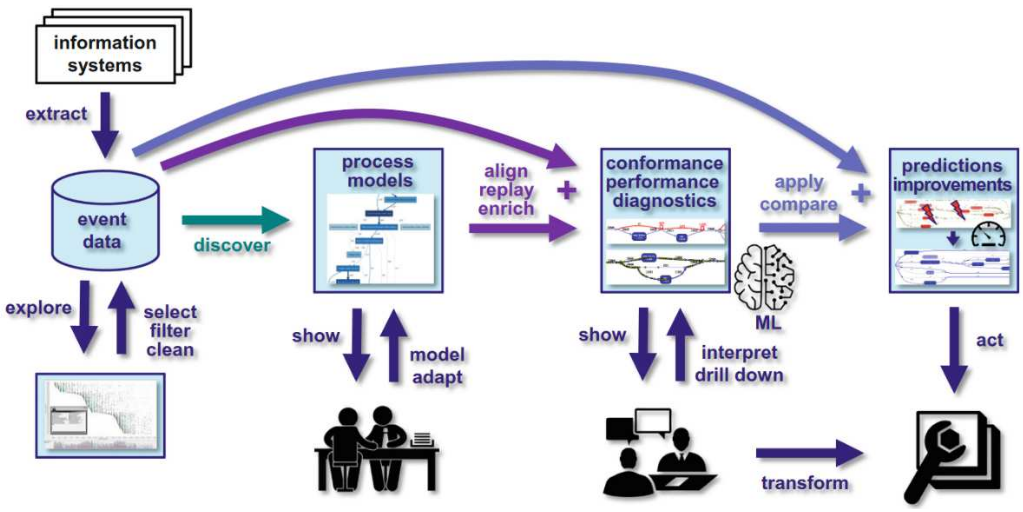

Figure 7: Procedure for the use of process mining [27]

The Data source to create the initial event log were the booking data of the Enterprise Resource Planning System. Due to the actual level of digitalization, it was possible to acquire at each machine specific information on product, time and order number. From this the event log for the analysis with the Celonis Process Mining environment was created and transferred to the MFM in the planning software Vistable (Figure 8). With the help of this model the overall utilization of the MHE could be assessed by the combination of MFM and layout. There the overall transport intensities for each source-sink connection will be visible. This is the basis for calculation of the process costs.

Due to the manual type of transport these are mainly based on labour costs. The process costs of 1 hour of transport can be calculated with 25 €, including 20% for the MHE used. With these information the overall process costs sheet was filled. In the last step the basis of these calculations was verified with the help of a MHE-Tracking database. The results showed that the overall usage of the MHE was 3 times higher than expected from the MFM. Further investigation showed that the MHE have a high number of empty travels and their average drive speed is in reality about 20 % slower than assumed. The efforts needed to create the process model consisted of a 1 hour online meeting with 2 persons and 4 hours modelling and transfer of the process model to the Planning Software.

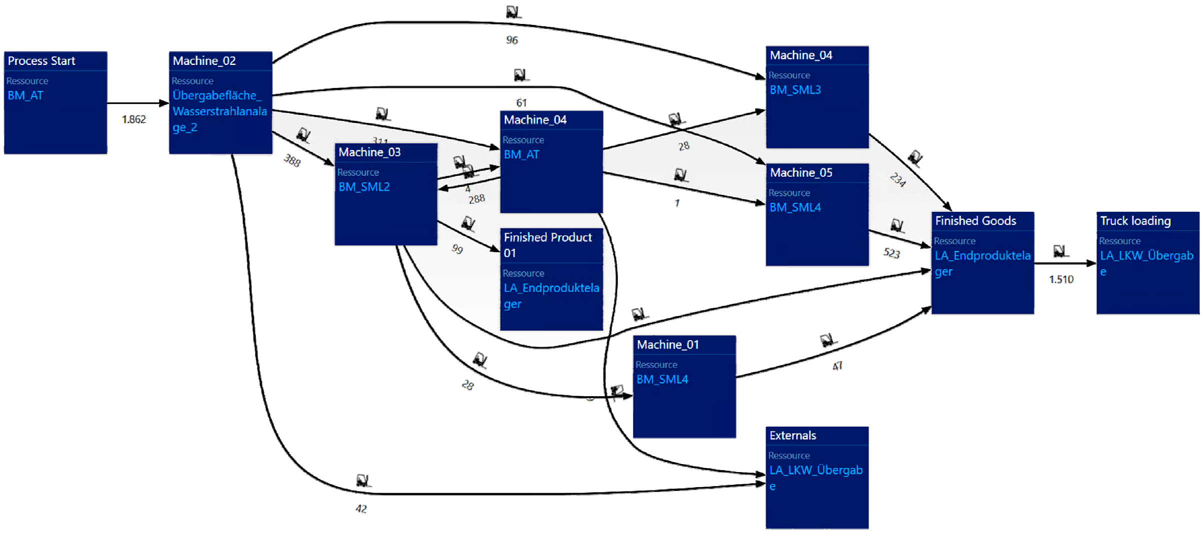

Figure 8: Material Flow Model based on the Process Mining Data

5 Limitations and Conclusions

The comparison between both approaches (interviews and process mining) shows that the efforts for the modelling itself can be highly reduced with the existing data basis. Also, the bias in the modelling of the type representatives can be reduced. A good and valid data base is therefore mandatory, but it was also seen that these data may already be available but not used properly. With the verification of the MFM with the actual MHE utilization data it becomes also clear that especially the assumptions based on Sink-Source connections are too conservative out of the model. In reality it takes more time to fulfil a transport task instead of just the source-sink distance connection. (RQ 1)

The methodology was only tested on one Use-Case with a limited number of resources which were investigated. So, for a more general view more Use-Cases are needed to validate the gains of the process mining int the context of intralogistics analysis.

Nevertheless, new technologies for the data analysis can and should be integrated into existing planning approaches to keep the benefits of specific material flow analysis software which then later can be used for in detail discussions on the layout situation for example. A integration leads to a higher quality of the outcome of further planning steps. (RQ 2)

References

[1] BEIER, Grischa, Silke NIEHOFF, and Bing XUE. More Sustainability in Industry through Industrial Internet of Things? [online]. *Applied Sciences.* 2018, 8(2), 219. Available from: 10.3390/app8020219.

[2] WANG, Ben. The Future of Manufacturing: A New Perspective [online]. *Engineering.* 2018, 4(5), 722-728. Available from: 10.1016/j.eng.2018.07.020.

[3] TAO, F., Y. CHENG, L. ZHANG, and A. Y. C. NEE. Advanced manufacturing systems: socialization characteristics and trends [online]. *Journal of Intelligent Manufacturing.* 2017, 28(5), 1079-1094. Available from: 10.1007/s10845-015-1042-8.

[4] SAREEN, Siddharth. Digitalisation and social inclusion in multi-scalar smart energy transitions [online]. *Energy Research & Social Science.* 2021, 81, 102251. Available from: 10.1016/j.erss.2021.102251.

[5] MIHAI, Florin, Ofelia Ema ALECA, Andrei STANCIU, Mirela GHEORGHE, and Mirela STAN. Digitalization—The Engine of Sustainability in the Energy Industry [online]. *Energies.* 2022, 15(6), 2164. Available from: 10.3390/en15062164.

[6] GILLANI, Fatima, Kamran Ali CHATHA, Muhammad Shakeel SADIQ JAJJA, and Sami FAROOQ. Implementation of digital manufacturing technologies: Antecedents and consequences [online]. *International Journal of Production Economics.* 2020, 229, 107748. Available from: 10.1016/j.ijpe.2020.107748.

[7] BROUM, Tomáš, Petr HOŘEJŠÍ, Miroslav MALAGA, and Pierre GRZONA. Competencies of Industrial Engineers for Implementing Augmented Reality Metadata Systems [online]. *Sustainability.* 2023, 15(1), 130. Available from: 10.3390/su15010130.

[8] ARDITO, Lorenzo, Antonio Messeni PETRUZZELLI, Umberto PANNIELLO, and Achille Claudio GARAVELLI. Towards Industry 4.0 [online]. *Business Process Management Journal.* 2019, 25(2), 323-346. Available from: 10.1108/bpmj-04-2017-0088.

[9] BUER, Sven-Vegard, Jan Ola STRANDHAGEN, and Felix T. S. CHAN. The link between Industry 4.0 and lean manufacturing: mapping current research and establishing a research agenda [online]. *International Journal of Production Research.* 2018, 56(8), 2924-2940. Available from: 10.1080/00207543.2018.1442945.

[10] SCHROEDER, Andreas, Ali ZIAEE BIGDELI, Carlos GALERA ZARCO, and Tim BAINES. Capturing the benefits of industry 4.0: a

business network perspective [online]. *Production Planning & Control.* 2019, 30(16), 1305-1321. Available from: 10.1080/09537287.2019.1612111.

[11] GHOBAKHLOO, Morteza. Industry 4.0, digitization, and opportunities for sustainability [online]. *Journal of Cleaner Production.* 2020, 252, 119869. Available from: 10.1016/j.jclepro.2019.119869.

[12] ACETO, Giuseppe, Valerio PERSICO, and Antonio PESCAPÉ. Industry 4.0 and Health: Internet of Things, Big Data, and Cloud Computing for Healthcare 4.0 [online]. *Journal of Industrial Information Integration.* 2020, 18, 100129. Available from: 10.1016/j.jii.2020.100129.

[13] JANMAIJAYA, Manvendra, Amit K. SHUKLA, Pranab K. MUHURI, and Ajith ABRAHAM. Industry 4.0: Latent Dirichlet Allocation and clustering based theme identification of bibliography [online]. *Engineering Applications of Artificial Intelligence.* 2021, 103, 104280. Available from: 10.1016/j.engappai.2021.104280.

[14] GRZONA, Pierre, Franziska BAUMGÄRTEL, and Luise WEIßFLOG. Mitarbeiterzentrierte Arbeitsbereichsgestaltung in KMU [online]. *Zeitschrift für wirtschaftlichen Fabrikbetrieb.* 2020, 115(12), 885-889. Available from: 10.3139/104.112426.

[15] MÖRTH, O., Christos EMMANOUILIDIS, N. HAFNER, and M. SCHADLER. Cyber-physical systems for performance monitoring in production intralogistics [online]. *Computers & Industrial Engineering.* 2020, 142, 106333. Available from: 10.1016/j.cie.2020.106333.

[16] GUDEHUS, Timm, and Herbert KOTZAB, eds. *Comprehensive Logistics.* Berlin, Heidelberg: Springer Berlin Heidelberg, 2012. 978-3-642-24366-0.

[17] GUERREIRO, Guilherme, Paulo FIGUEIRAS, Ruben COSTA, Maria MARQUES, Diogo GRAÇA, Gisela GARCIA, and Ricardo JARDIM-GONÇALVES. A Digital Twin for Intra-Logistics Process Planning for the Automotive Sector Supported by Big Data Analytics. *Volume 2B: Advanced Manufacturing:* American Society of Mechanical Engineers, 2019.

[18] HOFMANN, Erik, and Marco RÜSCH. Industry 4.0 and the current status as well as future prospects on logistics [online]. *Computers in Industry.* 2017, 89, 23-34. Available from: 10.1016/j.compind.2017.04.002.

[19] SCHMIDTKE, Niels, Fabian BEHRENDT, Lisa THATER, and Sascha MEIXNER. Technical potentials and challenges within internal logistics 4.0. *2018 4th International Conference on Logistics Operations Management (GOL):* IEEE, 2018, pp. 1-10.

[20] SCHENK, Michael, Siegfried WIRTH, and Egon MÜLLER, eds. *Factory Planning Manual.* Berlin, Heidelberg: Springer Berlin Heidelberg, 2010. 978-3-642-03676-7.

[21] SEVEN “RIGHTS” OF LOGISTICS. In: P. M. Swamidass, ed. *Encyclopedia of Production and Manufacturing Management.* Boston, MA: Springer US, p. 684.

[22] VDI 4405 PART 1. *Activity-based cost analysis in internal logistics.* Berlin: Beuth Verlag GmbH, 2003.

[23] GÖBEL, Fabian. Case Study Approach. In: Manfred Schwaiger and Anton Meyer, eds. *Theorien und Methoden der Betriebswirtschaft:* Vahlen, 2009, pp. 355-372.

[24] KATHLEEN M. EISENHARDT. Building Theories from Case Study Research. *The Academy of Management Review.* 1989, Vol. 14(4), 532-550.

[25] MOCKENHAUPT, Andreas. *Digitalisierung und Künstliche Intelligenz in der Produktion.* Wiesbaden: Springer Fachmedien Wiesbaden, 2021. 978-3-658-32772-9.

[26] CLEVE, Jürgen and Uwe LÄMMEL. *Data mining.* 2. Auflage. Berlin: De Gruyter Oldenbourg, 2016. De Gruyter Studium. 9783110456776.

[27] VAN DER AALST, Wil M. P. *Process Mining. Data Science in Action.* 2nd ed. 2016. Berlin Heidelberg: Springer Berlin Heidelberg; Imprint: Springer, 2016. 9783662498514.

[28] REINKEMEYER, Lars, ed. *Process mining in action. Principles, use cases and outlook.* Cham: Springer, 2020. 3030401715.

[29] RODRIGUES, Daniel, Paulo CARVALHO, Solange RITO LIMA, Emanuel LIMA, and Nuno Vasco LOPES. An IoT platform for production monitoring in the aerospace manufacturing industry [online]. *Journal of Cleaner Production.* 2022, 368, 133264. Available from: 10.1016/j.jclepro.2022.133264.