2025 29th International Conference Information Visualisation (IV) | 979-8-3315-7741-4/25/$31.00 ©2025 IEEE | DOI: 10.1109/IV68685.2025.00058

Process Mining for Production Optimization in Smart Manufacturing

Dirk Burkhardt
Software AG
Darmstadt, Germany
dirk.burkhardt@softwareag.com

Juliane Harbarth
Software AG
Darmstadt, Germany
juliane.harbart@softwareag.com

Andreas Görmer
Software AG
Darmstadt, Germany
andreas.goermer@softwareag.com

***Abstract*—Smart Manufacturing is currently the main objective when production manufacturers digitalize their plants to face current regulations and requirements to optimize costs, consumed resources, and sustainability. The focus is usually on extracting data from production and making it "analyzable". However, the results often neglect advanced options to optimize the production, may it in regards to the production process itself or in regards to consumed resources of the produced goods in specific. As the main contribution, this paper describes a novel approach to consider process mining appending to plant digitization and IoT analytics. As a result, the entire production process becomes transparent and therewith analyzable, but also the concrete consumed resources per produced good, per group, or as a whole can be analyzed. As the application benefit, the paper also outlines some advanced analysis capabilities to identify production optimizations based on process mining.**

***Index Terms*—Process Mining, Smart Manufacturing, Industry 4.0, IoT Analytics, Visual Analytics**

I. Introduction

The manufacturing industry is currently facing many upheavals worldwide. Companies in Europe, in particular, are facing a unique set of challenges. In addition to global competition, companies are being challenged by issues such as sustainability, more efficient resource usage especially in regards to energy usage, and a shortage of skilled workers.

To counter these difficult challenges, many companies have already started to plan and implement the digitization of their production lines and systems. Through digitization, companies hope to gain significantly more transparency and understanding of where further optimization and therefore saving potentials lies. And indeed, many reports and studies [1] [2] show that the digitization of production creates a much more comprehensive understanding of production, and the resources used.

This digitization of production is often referred to as Smart Manufacturing or Industry 4.0. As a rule, systems are used to monitor and analyze current production based on data. The usual main focus here is thereby on reading either the installed sensors or additional sensors and displaying them to the analyst [3]. These can be instantaneous consumption data, such as electricity values, or current switch and controller positions, as well as readings from consumption meters. This is often followed by aggregated views that provide an overview of consumption over a certain period or a customer's entire order [4]. Today, optimizations are usually made based on these aggregated views.

However, one challenge of these aggregated views are that it is still a perspective on the production plant, on which is tried to draw conclusions about the production goods and make corresponding optimizations. Currently, there is often a lack of approaches for making concrete statements about the resources consumed for a specific piece goods produced. But only with such a view that enables significantly beyond-going optimizations, for example, to optimize power consumption or reduce manual personnel costs, the full potential of smart manufacturing in regards of resource optimization can be achieved.

This paper therefore presents a novel approach for combining IoT analysis with process mining to obtain a much more comprehensive picture of production and outline far-reaching optimizations based on this. In contrast to IoT analysis, process analysis tracks the processing of the workpiece across the entire production line and allocates the efforts, consumptions, and errors incurred in each case. This also applies if workpieces undergo post-treatment, for example, and are therefore subject to extended processing. This specific piece goods analysis makes it possible to carry out comprehensive investigations and also to identify where there is great potential for savings.

II. Visual Analytics in Smart Manufacturing

Typical approaches to data visualizations focus on generating graphical views of data, which are often almost static. But even a wisely chosen visualization in regards of the given data, user, and task [5] cannot fulfill the demands on smart manufacturing analytics. In contrast, Visual Analytics aims for a stronger coupling of (data-)models to the visualization, which can be adjusted or replaced and offers therewith much more interactivity and analysis capabilities on the original data.

A. Visual Analytics Model

Thomas and Cook [6] define visual analytics as the science of analytical reasoning facilitated by interactive visual interfaces. Keim et al. [7] define visual analytics more precisely as: Visual analytics combines automated analysis techniques

2375-0138/25/$31.00 ©2025 IEEE
DOI 10.1109/IV68685.2025.00058

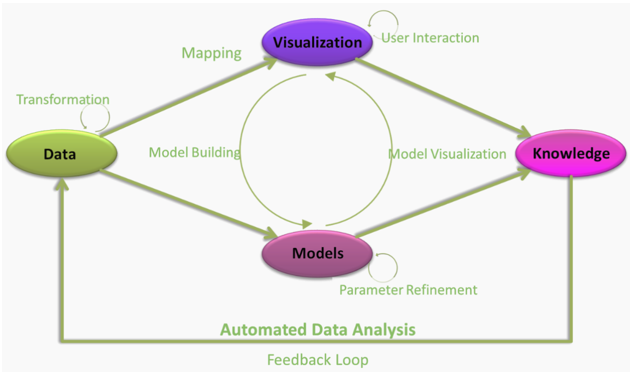

Fig. 1. The figure shows the refined Visual Analytics process model by Nazemi [8], which was originally created by Keim et al. [7]

with interactive visualizations for an effective understanding, reasoning, and decision-making on the basis of very large and complex data sets. The goals of visual analytics are summarized as [7]: (1) synthesize information and derive insight from massive, dynamic, ambiguous, and often conflicting data, (2) detect the expected and discover the unexpected, (3) provide timely, defensible, and understandable assessments, and (4) communicate assessment effectively for action.

The published visual analytics model of Keim et al. [7] sketches the processing of data visualizations and should lead to knowledge (see Fig. 1). The main step is the transformation of data to a (data-)model that can be transformed through mathematical functions to extract certain data aspects and gain new insight. The original model of Keim et al. was refined by Nazemi [8] to be more appropriate for applying it in modern visual applications.

B. Visual Analytics of IoT Data

The most common but also generic form of using visual analytics is using basic forms to analyze IoT data, such as temperatures, pressures, speeds, or flow rates. Most of the available systems, such as Cumulocity[1], already provide basic visualization components to analyze the data. But also cross connections to advanced analytical systems such as Tableau or Microsoft Power BI are common, besides more sophisticated systems with a stronger focus on plant workers, such as TrendMiner [3], [9]. The main idea for all of them is to enable analysis close to the sensor data.

A less common approach for this type of visual analytics is the creation of specific visualizations that are usually not part of common platforms. One of these examples is LiveGantt [10], which is a visual analysis approach used in analyzing large-scale manufacturing schedules consisting of numerous production tasks and resources. As the main features, it utilizes task aggregation and resource reordering algorithms to deal with schedule data at first and then uses a new Gantt chart to visualize the results of these algorithms. Another kind of approach is the visual analysis solution allowing test engineers to interactively steer ensembles generated in the performance test of automobile power system published by Matkovic et al. [11].

C. Visual Analytics of Production and Plants

Besides the generic visual analytics approaches, there are some production or plant-oriented analytical systems available. The intention is to offer an abstract visual view of the plant or production and apply analysis toward the root causes of problems or optimizations. Some recent visual analytics approaches, as presented by Xu et al. [12], enable real-time monitoring of entire production lines with the help of a sophisticated interface. AI methods can already be used to analyze large amounts of data today. A visual analytics system can actively help to set the right parameters for an AI application and evaluate the results [13]. These two publications are representatives of two research directions within the smart manufacturing research field. On the one hand, starting from the machine to visualize the complete manufacturing process [12], and on the other hand, to make the methods used more transparent [13].

Both methods therefore rely on anomaly detection or novelty detection. In the context of smart manufacturing, large volumes of multivariate data are usually used with unsupervised anomaly detection algorithms. Machine learning methods can also be used here. A data set within machine learning always consists of input values (initial situation) and the appropriate description (label). With large amounts of data (usually several terabytes), the manual annotation of each data set is very time-consuming and labor-intensive, and therefore ultimately cost-intensive and error-prone. This is why unsupervised techniques are used here. Here, the label is missing, and the anomaly is determined using mathematical operations based on the initial situation. Xu et al. generate automatic ensembles of predefined anomaly detection algorithms based on the different data types and visualize their effectiveness. In this way, the best ensembles can be selected for specific data to effectively detect anomalies within the large amount of multivariate data. This methodology makes the path from data to anomaly transparent.

Xu et al. [12] present with ViDX (Visual Diagnostics of Assembly Line Performance for Smart Factories) a dashboard that visualizes the manufacturing process. The interface is divided into five sections: the station map, histograms, an extended Marey graph, and a timeline with a calendar above it. The station map visualizes the sequence and connections of the individual stations. The histogram shows the utilization of the station. The extended Marey graph provides an overview of failures (gaps) or delays (converging graph with delayed further processing) by viewing all stations simultaneously. The timeline can be used to compare the production line scrap and anomalies can be found more quickly. The calendar aggregates the timeline monthly. Ke Xu et al [13] monitor the physical layout and output of an entire factory historically, i.e. over time. Today's factories usually emit even more complex data sets, which will require even more specialized visualizations

[1]Website of the Cumulocity IoT platform: https://www.cumulocity.com (last accessed: 12/07/2024)

in the future. Data visualization thus forms an essential basis within smart manufacturing and ensures that very complex issues are presented in a meaningful way.

III. Foundations to Process Mining and Smart Manufacturing

Process mining [14, p. 25-41] is an innovative discipline at the intersection of data science and business process management that aims to analyze and optimize real business processes based on digital traces. These digital traces, which are recorded in so-called event-logs, contain detailed information about the sequence of activities, time stamps, and resources involved in a business process. By using process mining techniques, companies can visualize and analyze the actual flows of their processes, identify deviations from the ideal or intended processes, and uncover potential for improvement.

The benefits of process mining lie particularly in its ability to bring transparency to complex processes that are often distributed across different IT systems and difficult to capture in traditional, manual analysis methods. For example, process mining can be used to identify inefficiencies, bottlenecks, or compliance violations in real time, enabling companies to react faster to process problems and continuously optimize their operations [15].

A. Event-based Processing

Process Mining is working on event data, which is generated during the application of business processes. This event data, often stored in so-called event-logs, contain usually information about (see also Table. I) [14, p. 128-137]:

- *Case ID:* Each case within a performed business process should consist about a unique ID.
- *Timestamp:* The timestamp represents when the specific event was thrown or rather the activity has started.
- *Activity:* The activity is the title of what has been done during a certain processing step, for instance, checking a ticket in a support system.

Further recommended information that should be logged are:

- *Event ID:* A unique event could be helpful to better distinguish events if log entries could be complex or multiple log entries build a single event entry.
- *Resource(s):* A resource may it in the form of a specific person, terminal, or software that is required or involved in the activity.
- *Duration or Cost:* The duration and/or cost represent how much of a resource was being consumed in the activity or the worth of the resource it represents.

But also, further supplementary dimensions could be considered, which might be useful in further analysis and optimization, such as:

- *Energy or Water Consumption:* Often beside the activity duration or resource costs, some other aspects could be important to consider too, e.g. to measure sustainability or other social aspects.
- *Aligned Tools:* Some specific business processes might require very special tools or systems that in some enterprises could be difficult to consider, e.g. a special machine in a craft business.

By continuously processing these events in real-time, companies can perform an accurate and timely analysis of their processes, making it possible to identify inefficiencies and exploit optimization potential more quickly. In contrast to traditional approaches, which are often based on periodic evaluations, event-based processing provides dynamic and immediate feedback, which is particularly beneficial in highly volatile business environments [14, p. 25-41] [16].

TABLE I

A fragment of some exemplary event-log: each line corresponds to an event. The data fields colored in red are mandatory, those in green are useful and just optional

IDs		Main Properties		Additional Properties			Supplementary Dimensions		
Case ID	Event ID	Timestamp	Activity	Duration	Ressource	Cost	Energy Consumption	Water Consumtion	Alligned Tools
1	412523	2024-07-21_10:15:30	register request	03:05	Pete	40	30	0	PC
	412524	2024-07-21_10:19:31	examine thoroughly	00:15	Sue	300	5	0	PC
	412525	2024-07-21_10:20:14	check tickt	01:03	Mike	90	15	0	PC
	412526	2024-07-21_10:21:23	decide	02:56	Sara	150	60	0	PC
	412527	2024-07-21_10:24:49	reject request	01:02	Pete	180	11	0	PC
2	415643	2024-07-28_15:24:49	register request	05:57	Pete	40	90	0	PC
	415646	2024-07-28_15:36:47	examine casually	09:42	Mike	350	180	0	PC
	415648	2024-07-28_15:49:09	check ticket	01:14	Ellen	80	15	0	PC
	415649	2024-07-28_15:58:23	decide	03:41	Sara	160	30	0	PC
	415650	2024-07-28_16:04:10	reinitiate request	07:34	Sara	160	120	0	PC
	415651	2024-07-28_16:14:41	examine throughly	02:01	Sean	300	50	0	PC
	415652	2024-07-28_16:19:55	check ticket	01:50	Pete	70	50	0	PC
	415653	2024-07-28_16:28:38	decide	03:27	Sara	160	70	0	PC
	415654	2024-07-28_16:39:14	pay compensation	02:17	Ellen	150	40	0	PC

B. Process Mining for Optimization

Process Mining is a concept that refers to a set of methods and technologies that come under the category of process management [17]. Process Mining's main aim is to examine how processes actually occur, how far they differ from the real model, what are the problems that occurred, what are the best ways to improve the process, and then the start of actual process improvement. This Process Mining technique can be implemented on any process as long as there is proper data stored by the target system.

Process Mining establishes a major connection between business process modeling and analysis, and data mining. In practice, these methods and techniques enable a very strict monitoring of the authenticity and reliability of data from key business processes. Both industry and academia have given Process Mining a great deal of attention, which has resulted in the development of various open-source and commercial tools for Process Mining. However, to demonstrate the complete picture, Process Mining techniques are gathered into four tasks as illustrated in Fig. 2 [18]. Process Mining techniques are [17] [19] [18]:

- Conformance checking
- Process re-engineering
- Operational Support
- Process discovery

Conformance checking identifies and diagnoses the discrepancies and similarities between a process model and an event-log. Conformance checking is also being performed to see whether reality matches the model (as captured in the log) and vice versa. The input process model can be normative

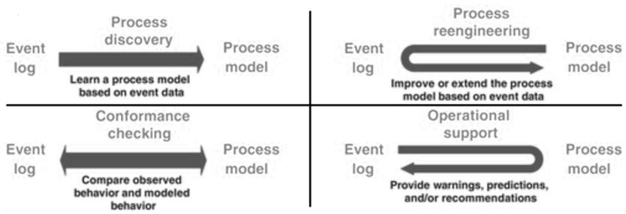

Fig. 2. The four basic tasks of Process Mining (based on [18]).

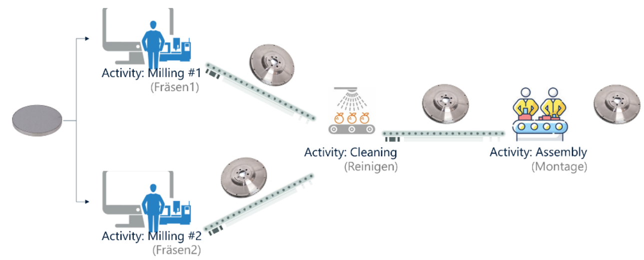

Fig. 3. Simplified illustration of the pilot factory's production line for the manufacture of oscillating disks.

or demonstrative. Furthermore, the process model could have been handmade or discovered through process discovery [18].

Process re-engineering uses event data to improve or extend the model. Both process model and event data are utilized as input, just as they are for conformance checking. However, now the objective is to modify the process model rather than to identify discrepancies. For example, sometimes the model can be ”repaired” to make it more realistic. Further perspectives can be added to an existing process model is also possible. For example, bottlenecks or resource utilization can be demonstrated using replay techniques [14] [18]. Models are updated as a result of process re-engineering. These models can be applied to actual operations to improve them.

Operational support has direct impact on the process, in a way that it issues recommendations, warnings, and predictions directly [14]. As conformance checking happen on the fly, this allows people to stay aware about what exactly the problem is at that particular time and what are the things that deviate. According to the data related to the event and model of the currently running instance of the process, allowing one to anticipate the time left in completion of the current flow, the probability of reaching the legal deadline, as well as the accompanying costs, the likelihood of a case being dismissed, and so forth. It does not improve the process by modifying it but by offering immediate data-driven assistance in the shape of recommendations, warnings, and predictions.

Process discovery provides valuable insights into the actual workflows of business processes and forms the basis for further analyses such as conformance checking and performance analysis [20].

C. Process Mining in Manufacturing

As already mentioned, process mining was originally and mainly intended to analyze and optimize business processes. However, the principal idea and approach works with almost any procedural data, which includes also production data along the manufacturing line. However, nowadays there are no applied approaches known where process mining is used to optimize manufacturing.

In the literature, there are just very few publications available that propagate the idea of using process mining in manufacturing in a theoretic model. One of these approaches is described by Massaro [21] to provide an innovative approach to model production management in industry, adopting a new “proof of concept” of advanced process mining automatizing decisions, and optimizing machine settings and maintenance interventions. The proposed work provides important parts of engineering management related to the digitization of production process matching with automated control systems setting production parameters, thus enabling the self-adapting of product quality supervision and production efficiency in modern industrial systems. Even though this approach is using process mining, it is very specific and adapted to automatic adjustment of certain parameters but not in analyzing and optimizing the manufacturing line in general. Another approach, developed by Yahya [22], is much stronger orienting on the manufacturing line. The objective of this research is to extend the existing process analysis framework by considering the attribute perspective. The paper gives some fundamental information on how to use it in manufacturing but is in many ways only applicable in a theoretical manner and gives less practical insights in analyzing or optimizing the production.

IV. Process Mining in Smart Manufacturing for Advanced Analytics

For our work, we collaborated with our partner of the PTW team at the TU Darmstadt which allowed us to apply our systems on their pilot factory [23]. One of the biggest advantages of the pilot factory is that we can apply our approach right now and make continuous advancements in data collection and mining without risking real business production. Important to mention is that the pilot factory is strictly orienting on real manufacturing industries, especially in regards of following a strict production line, considering standards and state-of-the-art production machines. This will later enable an application of our solution in real manufacturing enterprises.

Furthermore, in our research project, we also have some engaged small and medium-sized enterprises in the production area, but they are as of today not ready to apply process mining in their current stage of IoT integration. However, it is intended to apply process mining there too.

The pilot factory [23] of the PTW produces among others oscillating disks via the following stations (see Fig. 3):

- *Milling:* The plant offers two different milling machines that differ in some minor features and aspects. The milling is the most important step in creating the oscillating disks.

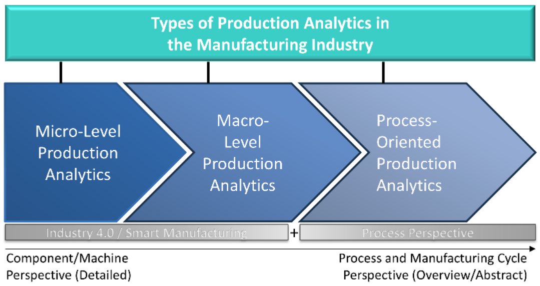

Fig. 4. Schematic representation of the Production Analytics types: Micro-Level and Macro-Level Production Analytics are the current standard in the context of Industry 4.0 and Smart Manufacturing. These approaches need to be expanded to include the Process Perspective in order to enable Process-Oriented Production Analytics.

- *Cleaning:* To clean the disks after the milling, the station enables cleaning via pressured air or water.
- *Assembly:* At the final station the oscillating discs will be assembled, after which they are ready to use.

A. Methodology and Sensorization

As sketched in Fig. 4, we build upon an already sensorized plant that makes use of Industry 4.0 or Smart Manufacturing respectively. This means, that we use the installed Cumulocity IoT platform at the pilot factory to manage the sensors and collect any available data from internal sensors and additional external sensors of the available plant machines via protocols like OPC-UA or Modbus.

When manufacturers digitize their production systems, they usually read the sensors or add sensors to the machines in the plant to extract data on their status. In terms of Industry 4.0 or Smart Manufacturing, this data is generally only suitable for Micro-Level Production Analytics. This means that it is usually only possible to read out and analyze the current statuses, such as which controllers are set and how, what electricity volumes are flowing, how hot certain parts of the system are, or what flow rate certain components have. To carry out Macro-Level Production Analytics, aggregated data is also required. For example, consumption or flow rate per minute or hour. Only statements over longer periods can offer information on the efficiency of the system or the current production speed. However, it is important to mention here that there are mostly statements about the production plant can be made and it is basically hardly possible to make statements about concretely produced products. In fact, Micro-Level and Macro-Level Production Analytics are the current standard when talking about Industry 4.0 or Smart Manufacturing.

In order to be able to make statements about specific products produced or the production process itself, a process perspective must also be added. This means that the products to be produced must be specifically traced during their production process and the exact required effort and consumed resources at each processing station must be recorded and saved in the event-log. This event-log can then be used later with process mining to analyze both, the process and the individual products.

TABLE II
EXCERPT FROM OUR COLLECTED EVENT-LOG FROM THE PILOT FACTORY, ON WHICH BASIS WE CARRY OUT PROCESS MINING:

Bauteil-ID / Case ID	Aktivität / Activity	Startzeitstempel / Start Time	Endzeitstempel / End Time	Energieverbrauch pro Bauteil [Ws] / Energy Consumption per Component [Wh]	Druckluftverbrauch pro Bauteil [l] / Compressed Air Consumption per Component [liter]
230721-05G-002	Fräsen2	21.07.2023 08:50	21.07.2023 09:01	440148	424,9635301
230721-05G-002	Reinigen	21.07.2023 09:02	21.07.2023 09:06	61428	0
230721-05G-002	Montage	21.07.2023 15:18	21.07.2023 15:20	0	2058,168284
230721-05G-003	Fräsen1	21.07.2023 09:03	21.07.2023 09:13	418344	435,6773582
230721-05G-003	Reinigen	21.07.2023 09:14	21.07.2023 09:18	64896	0
230721-05G-003	Montage	21.07.2023 15:21	21.07.2023 15:23	0	2014,221637
230721-05G-004	Fräsen2	21.07.2023 09:15	21.07.2023 09:24	351448	444,2320032
230721-05G-004	Reinigen	21.07.2023 09:24	21.07.2023 09:28	39948	0
230721-05G-004	Montage	21.07.2023 15:24	21.07.2023 15:25	0	1667,531351
230721-05G-005	Fräsen1	21.07.2023 09:25	21.07.2023 09:34	405204	451,5030365
230721-05G-005	Reinigen	21.07.2023 09:34	21.07.2023 09:38	39888	0
230721-05G-005	Montage	21.07.2023 15:26	21.07.2023 15:28	0	1479,537368
230721-05G-006	Fräsen2	21.07.2023 09:35	21.07.2023 09:43	381104	393,425094
230721-05G-006	Reinigen	21.07.2023 09:43	21.07.2023 09:47	38724	0
230721-05G-006	Montage	24.07.2023 15:27	24.07.2023 15:29	0	2041,077885
230721-06G-014	Fräsen2	21.07.2023 11:07	21.07.2023 11:12	264576	265,9553814
230721-06G-015	Fräsen1	21.07.2023 11:14	21.07.2023 11:22	386888	395,1829587
230721-06G-015	Reinigen	21.07.2023 11:23	21.07.2023 11:33	168930	0
230721-06G-015	Montage	24.07.2023 15:34	24.07.2023 15:35	0	1586,962488
230721-06G-016	Fräsen2	21.07.2023 11:23	21.07.2023 11:33	439716	490,7425164
230721-06G-016	Reinigen	21.07.2023 11:33	21.07.2023 11:38	108858	0
230721-06G-016	Montage	24.07.2023 15:37	24.07.2023 15:38	0	698,2635103
230721-06G-017	Fräsen1	21.07.2023 11:34	21.07.2023 11:42	359644	330,7083344
230721-06G-017	Reinigen	21.07.2023 11:43	21.07.2023 11:47	51894	0
230721-06G-017	Montage	24.07.2023 15:39	24.07.2023 15:40	0	239,2651164
230721-06G-018	Fräsen2	21.07.2023 11:43	21.07.2023 11:51	358368	326,2111273
230721-06G-018	Reinigen	21.07.2023 11:51	21.07.2023 11:55	48300	0
230721-06G-018	Montage	24.07.2023 15:40	24.07.2023 15:41	0	644,5509136
230721-06G-019	Fräsen1	21.07.2023 11:52	21.07.2023 12:00	363284	369,1955349
230721-06G-019	Reinigen	21.07.2023 12:00	21.07.2023 12:04	52038	0
230721-06G-019	Montage	24.07.2023 15:42	24.07.2023 15:43	0	737,3271811
230721-06G-020	Fräsen2	21.07.2023 12:01	21.07.2023 12:09	334028	341,1969375
230721-06G-020	Reinigen	21.07.2023 12:10	21.07.2023 12:13	33666	0
230721-06G-020	Montage	24.07.2023 15:43	24.07.2023 15:44	0	954,6189558

B. IoT and Process Data Fusion

To take now the processing aspects into account, the documentation/logging of the processing steps of the piece goods to be produced had to be incorporated. For this purpose, each input material was assigned a unique bar-code identifier. Whenever this material arrives at a station in the production line, it is scanned, as well as when it is finished at the station. The available IoT data is then used to record energy consumption and compressed air consumption for the specific workpiece in addition to the time data and the activity currently being carried out on the workpiece.

Final event data logging looks as shown in Table II. The format is a simple CSV file that can be extended by further data dimensions such as handwork time, consumed water, or consumed auxiliary materials, but maybe later, also data toward occurred processing errors, additional repair costs, or additional human resources and time.

C. Process Mining and Analytics Capabilities

A process mining algorithm, such as the Alpha algorithm, e.g. [24], [25], is used to generate the process model. This uses the event log to extract the individual activities and the corresponding transitions between them, as well as all resource consumption allocated to the activities.

The result can be imagined as a fuzzy cognitive map, where the different transitions between activities have a certain probability. In a further dimension, however, all resources are also assigned to the respective activities. The process model can therefore be used at any time to determine exactly which activity of a specific event log led to this activity in the process model. This complex interlinking of information enables a wide range of analyses to be carried out (see Fig. 5). Essentially, the analyses can be broken down into three levels:

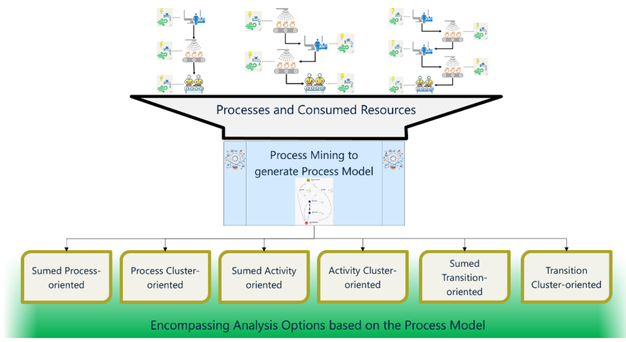

Fig. 5. Processing stages from the collected (process) log data via Process Mining to various analytical purposes.

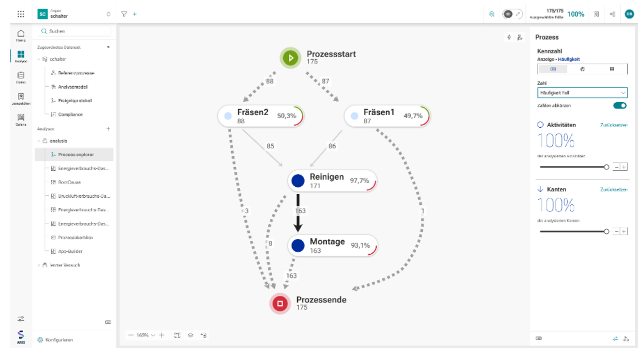

Fig. 6. Automatically generated production process through process mining. The numbers indicate how often a workpiece was processed at a particular station or which process routes were taken.

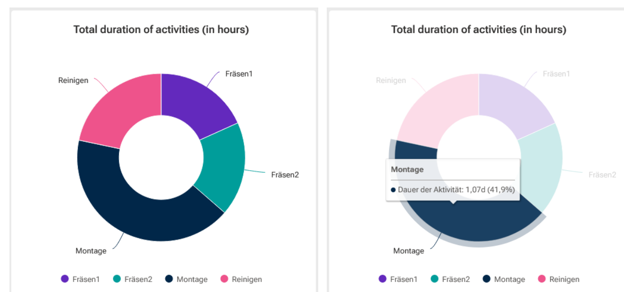

Fig. 7. Overview about the total duration of any activities in hours.

- *Process Level:* Here, the evaluation is always based on the entire process. This could be, for example, evaluations of the shortest or longest process processing or which process has consumed the most electricity.
- *Activity Level:* In contrast to the process level, there is only interest in the processing at a station, such as milling. This could involve variances in production, such as power consumption, or how often errors or abortions occur. If there are several stations and therefore activities, a comparison of efficiency or similar can also be relevant.
- *Transition Level:* Transitions are often rarely considered, but they are not insignificant. In other words, what happens between various activities or work-stations. Regular waiting times could be analyzed here. With reference to the logs, it should be emphasized here that the transitions are not recorded directly and therefore have no resource consumption alignment or something similar. Transitions are therefore recorded indirectly, via the interruption between two activities or work-stations.

Each of the three perspectives can essentially be further broken down into these two considerations:

- *Summarizing Evaluation:* This involves total statements that apply to the majority of all processes, activities or transitions. These can be statements on min, max or average consumptions. Classically, this information functions as performance metric as usually used in dashboards (e.g. [26]).
- *Cluster Evaluation:* Before an evaluation can take place here, the specific clusters must first be selected for cluster evaluations. These can be processes, activities or transitions that take a particularly long time and are therefore above a certain threshold. The focus here is on the evaluation of a subset of existing processes, activities or transitions.

D. Visual Process Analytics in Smart Manufacturing

For Process Mining, we use the market software ARIS Process Mining[2] [27], which we configured to consume the CSV file as defined above. Basically, any kind of process mining solution would be suitable, but ARIS Process Mining enables direct analysis and show the results in diagrams, which simplifies the follow-up work.

As a result, the system generates a process model (see Fig. 6) just based on the event-logs in the CSV file. Furthermore, it shows all possible transitions through the process activities but highlights those that are most often being used. An analysis of the process paths could help to identify possible problems in the manufacturing process.

The higher analytical value is the detailed analysis which can be done by analyzing the additional data dimensions. In contrast to analyzing IoT data, these process data allow a huge analysis variety in perspective of time and resource consumption.

An example is the analysis of consumed time at the different stations which is essential if a large production is planned, and each station should take approximately the same time. As could be seen in Fig. 7 the different stations need significantly different times. In particular, the assembling takes most of the time.

In Fig. 8 we could check the proportion of consumed resources. What we could see is that the total consumed resources in contrast to the median consumed resources per case/piece seems to be equivalent.

[2]Website of ARIS Process Mining: https://aris.com/process-mining/ (last accessed: 13/04/2025)

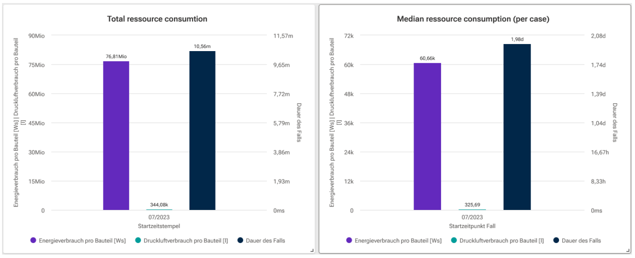

Fig. 8. Comparison of the total resource consumption to the median resource consumption of a case/product.

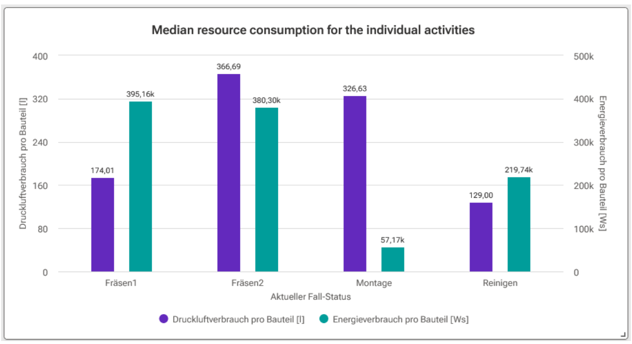

Fig. 9. Comparative view about the median resource consumption for the individual activities/stations on the plant.

Another interesting insight into the production is shown in Fig. 9. In particular, the differences between the two different milling machines are obvious. While milling machine #1 consumes significantly less air pressure in contrast to milling machine #2, it consumes slightly more energy than milling machine #2.

The last insight in Fig. 10 shows the spread between min and max resource consumptions for individual activities. The huge span indicates that the production quality seems to be a problem since some piece goods seems not to be finished the production line. If this indicated anomaly is critical would require further investigation.

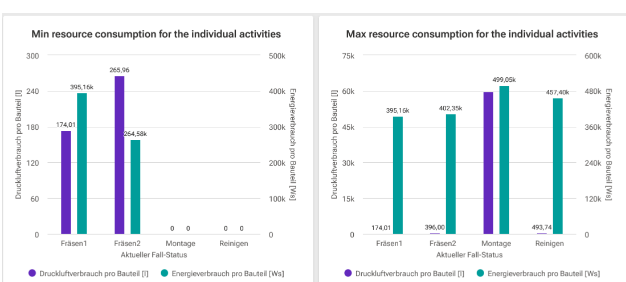

Fig. 10. Comparative view on min and max consumed resources for the individual activities/stations to identify potential failure production (rejected goods).

V. Discussion

We are currently using the approach with our partners in the PTW learning factory at TU Darmstadt, but it is designed in such a way that the approach can also be implemented comparatively easily with our application partners from industry. In particular, the fact that the process model does not have to be defined manually first, but is generated dynamically directly from the logs via process mining, significantly simplifies the application. There is also no need for extensive digitalization of the production facilities beforehand; simple resource measurement instruments are sufficient. The biggest challenge is the traceability of the workpieces to be produced, but the approach can also be applied to tranches or orders. This means that older production plants, in particular, only need to be expanded to include scanning the workpieces at the start and end of processing at each workstation/machine and recording the resources or wear and tear consumed during this time. This makes the application at small and medium-sized producing companies most useful. Compared to digitization for smart manufacturing, this is a considerable simplification that nevertheless leads to very accurate results. In the project, we were also able to easily implement an approach to generate accurate CO2 emissions for a digital product passport (DPP) [28], as it should become standard in the European Union for any sold product.

As the approach does not follow the classic approach of pure digitization of the production line, it is relatively difficult to communicate this as an alternative but often more expedient approach. In the project, however, we had positive experiences overall, especially because the accuracy of the recorded emissions was so precise.

VI. Conclusion

In this paper, a novel approach was presented on how smart manufacturing using IoT can be improved by supplementing it with process mining in order to measure production and consumption much more effectively based on the products produced. This is because process mining can be used to allocate consumption and activities not just at the machine level, but to products along the entire production chain. Special expenses, for example, due to reworking processes, are also recorded and can be analyzed in detail. This means that not only can the production process be comprehensively analyzed and optimized, but also all consumptions for a product can be viewed individually, in clusters, or as a whole. This allows completely new insights into where which expenses actually arise or can even arise in rare cases that finally affect the entire production process.

Acknowledgment

The authors gratefully acknowledge the financial support of the CliCE-DiPP project (grant agreement no. 01MN23023F), funded by the Federal Ministry of Economic Affairs and Climate Action (BMWK). The project supervision of CliCE-DiPP is carried out by the DLR Projektträger. The authors would like to thank all the institutions mentioned for their financial and organizational support.

References

[1] P. Garg, B. Gupta, A. Sar, G. Graham, and A. Shore, "Development and validation of an instrument to measure the perceived benefits of digitalization in manufacturing," *IEEE Transactions on Engineering Management*, vol. 71, pp. 8288–8306, Apr. 2024. [Online]. Available: https://doi.org/10.1109/TEM.2024.3390434

[2] A. Zangiacomi, E. Pessot, R. Fornasiero, M. Bertetti, and M. Sacco, "Moving towards digitalization: a multiple case study in manufacturing," *Production Planning & Control*, vol. 31, no. 2-3, pp. 143–157, 2020. [Online]. Available: https://doi.org/10.1080/09537287.2019.1631468

[3] D. Burkhardt and G. Ristow, "Enabling smart manufacturing with visual analytics for plant workers," in *Datenbanksysteme für Business, Technologie und Web (BTW 2025)*, ser. Lecture Notes in Informatics (LNI) - Proceedings, Volume P-361. Bonn: Gesellschaft für Informatik, 2025, pp. 665–678. [Online]. Available: https://doi.org/10.18420/BTW2025-33

[4] H. Schöning and M. Dorchain, *Big Smart Data – Intelligent Operations, Analysis und Process Alignment*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2024, pp. 611–630. [Online]. Available: https://doi.org/10.1007/978-3-662-58528-3_70

[5] D. Burkhardt, K. Nazemi, P. Sonntagbauer, S. Sonntagbauer, and J. Kohlhammer, "Interactive visualizations in the process of policy modelling," in *Electronic Government and Electronic Participation - Joint Proceedings of Ongoing Research of IFIP EGOV and IFIP ePart 2013*, ser. LNI 221, vol. 221. Bonn, Germany: Gesellschaft für Informatik e.V.., 2013, pp. 104–115, iSBN: 978-3-88579-615-2. [Online]. Available: https://dl.gi.de/handle/20.500.12116/17255

[6] J. J. Thomas and K. A. Cook, *Illuminating the Path: The Research and Development Agenda for Visual Analytics*. USA: IEEE Computer Society, 2005.

[7] D. Keim, G. Andrienko, J.-D. Fekete, C. Görg, J. Kohlhammer, and G. Melançon, *Visual Analytics: Definition, Process, and Challenges*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, pp. 154–175. [Online]. Available: https://doi.org/10.1007/978-3-540-70956-5_7

[8] K. Nazemi, "Adaptive semantics visualization," Ph.D. dissertation, Technische Universität Darmstadt, Darmstadt, 2014. [Online]. Available: http://tuprints.ulb.tu-darmstadt.de/4319/

[9] TrendMiner, "Introducing analytics into industrial environments: How to start using self-service industrial analytics to control business outcomes," Online, Hasselt, Belgium, 2020, https://www.trendminer.com/wp-introducinganalytics-3/ (last accessed: 24/05/2024). [Online]. Available: https://www.trendminer.com/wp-introducinganalytics-3/

[10] J. Jo, J. Huh, J. Park, B. Kim, and J. Seo, "Livegantt: Interactively visualizing a large manufacturing schedule," *IEEE Transactions on Visualization and Computer Graphics*, vol. 20, no. 12, pp. 2329–2338, 2014. [Online]. Available: https://doi.org/10.1109/TVCG.2014.2346454

[11] K. Matković, D. Gračanin, R. Splechtna, M. Jelović, B. Stehno, H. Hauser, and W. Purgathofer, "Visual analytics for complex engineering systems: Hybrid visual steering of simulation ensembles," *IEEE Transactions on Visualization and Computer Graphics*, vol. 20, no. 12, pp. 1803–1812, 2014. [Online]. Available: https://doi.org/10.1109/TVCG.2014.2346744

[12] P. Xu, H. Mei, L. Ren, and W. Chen, "Vidx: Visual diagnostics of assembly line performance in smart factories," *IEEE Transactions on Visualization and Computer Graphics*, vol. 23, no. 1, pp. 291–300, 2017. [Online]. Available: https:/doi.org/10.1109/TVCG.2016.2598664

[13] K. Xu, M. Xia, X. Mu, Y. Wang, and N. Cao, "Ensemblelens: Ensemble-based visual exploration of anomaly detection algorithms with multidimensional data," *IEEE Transactions on Visualization and Computer Graphics*, vol. 25, no. 1, pp. 109–119, 2019. [Online]. Available: https://doi.org/10.1109/TVCG.2018.2864825

[14] W. M. P. van der Aalst, *Process Mining: Data Science in Action*, 2nd ed. Heidelberg: Springer, 2016. [Online]. Available: https://doi.org/10.1007/978-3-662-49851-4

[15] L. Reinkemeyer, *Process Mining in a Nutshell*. Cham: Springer International Publishing, 2020, pp. 3–10. [Online]. Available: https://doi.org/10.1007/978-3-030-40172-6_1

[16] S. J. Zelst, B. F. Dongen, and W. M. Aalst, "Event stream-based process discovery using abstract representations," *Knowl. Inf. Syst.*, vol. 54, no. 2, p. 407–435, feb 2018. [Online]. Available: https://doi.org/10.1007/s10115-017-1060-2

[17] A. Corallo, M. Lazoi, and F. Striani, "Process mining and industrial applications: A systematic literature review," *Knowledge and Process Management*, vol. 27, no. 3, pp. 225 – 233, 8 2020. [Online]. Available: https://doi.org/10.1002/kpm.1630

[18] W. M. van der Aalst, "Process discovery from event data: Relating models and logs through abstractions," *WIREs Data Mining and Knowledge Discovery*, vol. 8, no. 3, p. e1244, 2018. [Online]. Available: https://doi.org/10.1002/widm.1244

[19] A. Berti, S. J. van Zelst, and W. M. P. van der Aalst, "Process Mining for Python (PM4Py) : Bridging the Gap Between Process- and Data Science," in *Proceedings of the ICPM Demo Track 2019*, ser. CEUR workshop proceedings, vol. 2374. Aachen, Germany: RWTH Aachen, Jun 2019, pp. 13–16. [Online]. Available: https://publications.rwth-aachen.de/record/774274

[20] W. van der Aalst, A. Adriansyah, and B. van Dongen, "Replaying history on process models for conformance checking and performance analysis," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 2, no. 2, pp. 182–192, 2012. [Online]. Available: https://doi.org/10.1002/widm.1045

[21] A. Massaro, "Advanced control systems in industry 5.0 enabling process mining," *Sensors*, vol. 22, no. 22, 2022. [Online]. Available: https://doi.org/10.3390/s22228677

[22] B. N. Yahya, "The development of manufacturing process analysis: Lesson learned from process mining," *Jurnal Teknik Industri: Jurnal Keilmuan dan Aplikasi Teknik Industri*, vol. 16, no. 2, pp. 95–106, Dec. 2014. [Online]. Available: https://doi.org/10.9744/jti.16.2.95-106

[23] A. Wank, S. Adolph, O. Anokhin, A. Arndt, R. Anderl, and J. Metternich, "Using a learning factory approach to transfer industrie 4.0 approaches to small- and medium-sized enterprises," *Procedia CIRP: 6th CIRP Conference on Learning Factories*, vol. 54, pp. 89–94, 2016. [Online]. Available: https://doi.org/10.1016/j.procir.2016.05.068

[24] W. M. van der Aalst, *Process Mining in the Large: A Tutorial*. Cham: Springer International Publishing, 2014, pp. 33–76. [Online]. Available: https://doi.org/10.1007/978-3-319-05461-2_2

[25] A. Alves De Medeiros, B. Dongen, van, W. Aalst, van der, and A. Weijters, *Process mining: extending the alpha-algorithm to mine short loops*, ser. BETA publicatie : working papers. Technische Universiteit Eindhoven, 2004. [Online]. Available: https://research.tue.nl/en/publications/a8f8afe3-13cf-40d5-9d11-1b3cb07dd594

[26] K. Nazemi and D. Burkhardt, "Visual analytical dashboards for comparative analytical tasks – a case study on mobility and transportation," *ICTE in Transportation and Logistics 2018 (ICTE 2018)*, vol. 149, pp. 138–150, 2019. [Online]. Available: https://doi.org/10.1016/j.procs.2019.01.117

[27] R. Peters and M. Nauroth, *Process-Mining-Software-Tools*. Wiesbaden: Springer Fachmedien Wiesbaden, 2019, pp. 41–57. [Online]. Available: https://doi.org/10.1007/978-3-658-24170-4_8

[28] D. Burkhardt, M. Bock, and A. Kuijper, "Dppviewer: A visual analytics approach for optimizing production chains on digital product passports," in *2025 29th International Conference Information Visualisation (IV)*. IEEE, 2025.