

Concepts for improving the quality of production plans using machine learning

Lukas Lingitz¹, Wilfried Sihm^{1,2}

¹ Fraunhofer Austria Research GmbH, Theresianumgasse 27, A-1040, Vienna, Austria

² TU Wien, Institute of Management Science, Theresianumgasse 7, A-1040, Vienna, Austria

ABSTRACT

There are always deviations between production planning and subsequent execution. Furthermore, it has been found that the reliability of production plans and thus Planning Quality (PQ) can drop down to 25 % in the first three days after plan creation [1]. These deviations are caused by uncertainties, such as inaccurate or insufficient planning data (including data quality and availability); inappropriate planning and control systems; and unforeseeable events. Production planners therefore use buffers in the form of inventories or extended transitional periods to create possibilities for implementing corrective measures in production control. Buffers, however, lead to increased coordination and control efforts as well as to negative effects, particularly on the inventory, throughput time, and capacity utilisation. The potential for more accurate planning remains largely unexploited. The objective of this paper is to investigate the possibilities of increasing planning quality. Within a case study, the authors demonstrate how machine learning can be used to predict cycle times. Furthermore, the increased accuracy compared to the current method is shown. Based thereon, two approaches are presented, focusing on the reduction of gaps between the master data and predicted data used during the production planning process. Moreover, further research needs are identified.

Section: RESEARCH PAPER

Keywords: production planning; planning quality; master data; prediction; machine learning

Citation: Lukas Lingitz, Wilfried Sihm, Concepts to Improve the Quality of Production Plans using Machine Learning, Acta IMEKO, vol. 9, no. 1, article 6, March 2020, identifier: IMEKO-ACTA-09 (2020)-01-06

Editor: Lorenzo Ciani, University of Florence, Italy

Received November 6, 2019; **In final form** February 10, 2020; **Published** March 2020

Copyright: This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by the European Union under the programme Electronic Component Systems for European Leadership (ECSEL) EU, Project: Power Semiconductor and Electronics Manufacturing 4.0 (Semi40) (Grant Agreement No. 692466).

Corresponding author: Lukas Lingitz, e-mail: lukas.lingitz@fraunhofer.at Fehler! Linkreferenz ungültig.

1. INTRODUCTION

Industry 4.0, Cyber-Physical Production Systems (CPPSs) [2], [3], and the Industrial Internet of Things have a significant influence on Production Planning and Control (PPC) today and will do so in the future. Deploying CPPSs raises several challenges for industries, as addressed in [4]. These challenges include the extraction of knowledge from heterogeneous data sources; interoperations with production information systems; and new possibilities in terms of the changeability, adaptability, and reconfigurability of production systems. Compared to traditional production planning, which is based on a generally static knowledge base, smart factories enable the collection and exchange of real-time information between products, machines, processes, operations [5], and systems. The application and exchange of data by the elements of a smart factory lead to an automated and decentralised production, which is an essential

characteristic of Industry 4.0 [6], [7]. However, there is a need to study how the different solutions – enabled by digitalisation – can support PPC and contribute to an increased corporate competitiveness [8].

Planning Quality (PQ) is a commonly used term in industrial practice when discussing PPC. However, the term is not clearly defined in the scientific context. Therefore, the authors give a proposal for a new definition of the term, trying to set a widely accepted standard. After defining PQ, a case study is given, wherein the suggested PQ has been increased through the application of Machine Learning (ML) for lead time prediction. Lastly, the authors give an outlook of two novel approaches that will be the main subject of an ongoing research project.

The paper is structured as follows. In section 2, the authors give a literature review concerning PQ and the application of ML in the field of PPC. In section 3, a definition of the term PQ is given. A case study from the semiconductor industry in section

4 demonstrates the accuracy of ML for cycle time prediction compared to the currently used static approach. Section 5 relays the results of the case study and presents different approaches to how to improve PQ by applying ML. In section 6, the authors discuss these approaches and identify further research needs. Finally, a conclusion and outlook are given.

2. RELATED LITERATURE

Literature on PPC approaches the phrase PQ from different angles, considering several influencing factors. The first aspect focuses on classical logistical targets of PPC, such as flow times, due dates, setup costs, and product features. Since accurate and high-quality planning data is one of the most important parts of a good production plan, master data management is an important aspect in our discussion as well [9]. These first two approaches do not deal with (planned or unforeseeable) changes or uncertainties of the production environment, which are typical elements of the paradigm shift towards Industry 4.0 [10]–[12]. Therefore, robustness and resilience consider such disturbances. A production plan is robust if its performance is guaranteed – even when facing events not known at the time of planning [13], [14]. Resilience is, in contrast, the ability of a system to cope with changes of all kinds [15]. Two efficient ways of dealing with uncertainty are the application of stochastic fuzzy models and the use of adaptive and cooperative approaches [16]. The overall objective of PPC is the creation of reliable production plans, so their realisation on the shop floor should be close to – or ideally the same as – the production plan as it was originally planned. The deviation between planning and reality on the shop floor increases up to 75 % after just three days in medium-sized mechanical engineering enterprises [1]. As shown in the literature review, it is desirable to have more reliable production plans. Measuring the quality of the prediction, as an alternative, may reveal a potential for bridging the gap between planned and actual figures. Besides, the success of a good production plan depends on the decision-making process itself. In the era of Industry 4.0, the automation of decision-making processes and the level and way of human engagement are also essential topics [17], [18]. It can be concluded that there is no exact definition of PQ. In this paper, a novel industry-oriented concept for measuring, evaluating, and improving PQ will be developed.

A general truth is that data does not bring any added value on its own, but in practice, domain-specific knowledge and algorithms are needed to extract useful information from heterogeneous and scalable data sources [19]. Simple statistical analysis is often not sufficient, as it is time-consuming and often does not lead to the desired results. Hence, automated data extraction and analytical methods are needed. Together with the rise of data science as one of the most popular emerging research and application fields today, ML has gained increasingly high attention in the recent past.

At the very beginning of the development of ML, the vast majority of papers were published in journals related to the topic of computer science. However, with increasing demands of computational capabilities and big data analytics, the area is growing, with far-reaching applications in diverse disciplines. Nowadays, many different disciplines use ML algorithms, as experienced in 2012 [20]. The first ML applications in management science can be found in finance and marketing [21]. In 2009, Choudhary et al. identified that the emerging application of ML in PPC has not been systematically explored [22]. However, in the following years, several research papers were

published in the context of production management, focusing on applying ML to advanced planning and scheduling [23]; quality improvement; process monitoring; and defect analysis [24]. Yet, researchers have not intensively focused on (sub-)topics relevant to PPC – such as flow time prediction, lot cycle time prediction, and lead time prediction; thus, the improvement potential has not been completely identified and maximised.

The results of our literature review determine that the current trend in PPC is to employ ML-based simulation and optimisation algorithms. Furthermore, it can be recognised that the focus of most of the analysed publications is either given to production scheduling (47 %) or other applications (33 %), while the prediction of planning relevant times is rarely considered (20 %), as shown in [19].

3. PLANNING QUALITY – DEFINITION

Until now, there has been no uniform mathematical definition and understanding of the term PQ in the scientific literature within the framework of PPC. Therefore, we briefly explain and define what is meant by the term PQ as we see it.

PQ is foreseen as a key indicator for the planner to assess the reliability of the production plan in the planning phase (time t) and to continuously improve the operational reliability of production plans in forthcoming phases (time $t+n$). In principle:

- The PQ is high if there are ideally no deviations or at least deviations within an acceptable range (which can vary depending on different industrial contexts) between the production plan created in advance and the actual execution.

Based upon this general statement, we need to derive some measurable values to determine exactly if the production plan and the execution show deviations or not. On a macroscopic level, we could say that PQ is high if at least the calculated end date of the last production step of a production order from the production plan meets the actual end date. We can also break this assumption down for every operation. This would mean that PQ is high if every operation/production step of the production order starts and ends on the designated day/shift/hour calculated during the plan creation. This leads to the conclusion that the times that are used for the plan creation need to be very accurate so that the start and end times of the setup, operation, transition, and waiting not only meet on average – if we look to a longer period of time – but also need to match the reality for every order in every situation (e.g. over- and underload). This brings us to the next definition:

- The PQ is high if there are ideally no deviations or at least deviations within an acceptable range (which can vary depending on different industrial contexts) between the planning times created in advance and the actual times that result during the execution phase.

There are many factors that could influence the actual times and cause deviations between the planning times and the actual times. One of the main problems is that we normally use static average values for the planning times. If we would use ML models with a high prediction accuracy and trained with confirmation data from the actual production system to predict the actual times, we could reformulate the statement from above:

- The PQ is high if there are ideally no deviations or at least deviations within an acceptable range (which can vary depending on different industrial contexts) between predicted times and actual times (e.g. lead time, setup time, and operation time). Accurate dynamic prediction models

are required to continuously reduce the deviation between reality and prediction AND

- There are ideally no deviations between the static planning times used for the production plan creation and the predicted times OR
- Instead of static planning times, the prediction model itself is used to create production plans that have a high PQ.

Based on the general statement that PQ measures the deviation between a production plan and the later execution, we now have two possibilities for creating production plans with higher PQ. We achieve this by having accurate prediction models of the planning times on the one hand and by using these models during the planning phase on the other hand. The definition shows two methods. The first option is to adapt the planning times to match the predicted times. In this case, we assume that the planning times are still one static value, so we might adapt the times in several iterations until we reach a convergence point. Alternatively, we can use the dynamic time models, which can be considered as a function with several variables as the input for the plan creation.

The above definitions and statements give rise to the conclusion that there are two possible ways of creating production plans with high PQ. These two methods are shown in two approaches that are described and discussed in sections 5 and 6 and are subject to further research within a funded research project. Within the upcoming case study, the possibility of predicting planning times based on confirmation data is demonstrated.

4. CASE STUDY ON TIME PREDICTION

In this section, the authors give an example of the application of ML for time prediction. After an introduction to the related industrial sector of the use case partner and a general description of the production system, the approach and the results are given. Based on the results, ideas of how to apply ML within the planning process are stated.

4.1. Introduction

As semiconductor industries offer a high availability of data for the production process, these industries take the lead when it comes to innovative applications in the area of industrial data science. Furthermore, the market is highly competitive, and high productivity as well as short lead times (among others) are the key factors for success.

Most products in the semiconductor industry are built in layers, where the same or similar production steps are repeated in cycles to build integrated circuits on the layers. Nevertheless, most production systems do not consist of rigidly linked machines. Rather, they are made up of highly automated machines that perform specific operations organised in a job shop-type production system. The often broad product spectrum is driven by rapid product innovations and results in a high material flow complexity. Therefore, constant sequencing is needed.

Within the research project, the goal has been to increase the PQ of a semiconductor manufacturer. Due to good data availability, the authors decided to apply ML techniques to build prediction models for lead time prediction that can subsequently be used to analyse the accuracy of the production plans. As the company had no simulation model of the plant and the authors did not build a model, the prediction models were solely built with readily available data from the shop-floor IT systems

(Enterprise Resource Planning [ERP] system and Manufacturing Execution System [MES]).

4.2. Approach and results

After collecting the process confirmation data from the MES, historical data about machines/equipment/work centres and customer-related information, features were generated based on the domain knowledge and experience of the process experts. Furthermore, some features that were relevant from a production logistics point of view were added. Many features that were created can be parametrised in a way such as ‘average lead time of the last x lots’, where x can be any positive integer number. The feature catalogue with the most important features can be found in [19].

According to the domain experts, the overall lead time (hereafter defined as the time span from the beginning of the first operation to the finish of the last operation of an order) is influenced by three main process steps. Therefore, we started the analysis with these steps. After calculating the features, several regression methods were tested to predict the cycle time. For the evaluation of the accuracy, the Normalised Root Mean Squared Error (NRMSE) was chosen as the measurement index. The NRMSE is the square root of the mean of the squares of the deviations between a predicted value and its actual regression dependent value divided by the range of the biggest and smallest regression dependent value for normalisation purposes. The normalisation makes different datasets and models comparable.

The conclusion is that ensemble tree-based methods (bagged regression trees, random forests, and boosted regression trees) outperformed all tested regression algorithms. Furthermore, these methods give quantitative feedback about the feature’s importance i.e. the importance of a variable on the prediction.

After these experiments, the scope of the study was broadened (time frame, products, and process steps). As a result, the lead time prediction for 80 process steps was done.

Also for the broader scope random forest outperformed linear regression, based on different aspects (the need for data cleaning, the training time, and the accuracy of the model). To evaluate the advantages of the proposed approach compared to the current approach – creating production schedules with static target cycle times for every layer of every different product, updated in each quartal – the dynamic cycle time prediction was

Table 1. NRMSE values for the different cycle time calculations.

ProductID	Average	ML	Number of lots
1	17.9	11.7	2228
2	12.5	12.7	1103
3	15.9	13.7	964
4	11.8	18.0	237
5	23.1	18.4	289
6	20.6	18.8	345
7	23.7	20.4	142
8	27.4	25.1	118
9	25.9	29.0	32
10	56.0	41.0	11
11	37.4	43.4	32
12	19.4	66.3	145
13	18.3	-	7
14	117.8	-	9

done for each layer of the 14 most important products with the random forest algorithm. A comparison of the NRMSE values is given in Table 1 for the currently applied static planned cycle times (average) and for the predicted dynamic cycle times (ML). In most cases, random forest gives a more accurate prediction of the cycle time, as the NRMSE values are mostly smaller, especially when the sample is bigger. This is of high relevance in industry, as the lead times, in general, are three to seven weeks long, and even small deviations between predicted and actual end dates, calculated based on the lead time, cause a delay in shipping of several days.

However, this means that the prediction accuracy is sensitive to the amount of data, and under a certain limit, no cycle time prediction with ML could be done. The inclusion of products with a minimum of 200 samples (ProductIDs 1 to 5) leads to the mean NRMSE values for the predicted dynamic cycle times (ML) of 15.5 % and 17 % for the static planned cycle times (average). To highlight the difference between the static and dynamic approach, the real cycle times (black dots) are compared with the currently used static (orange line – some outliers are not printed due to the scaling of the graph) and the predicted dynamic (green dots) cycle times for a given product and layer in Figure 1. It must be mentioned that a rolling forecast was applied to capture trends. This means that the ML algorithm learns from the past 2000 observations and predicts the next 100 observations. Therefore, for the first 2000 observations, no prediction is available.

4.3. Outcome of the project

For the deployment, an interactive web application was developed to support the whole cycle – data cleaning, outlier detection, feature generation, model training, and prediction. The whole process is depicted in Figure 2. For the development of the web application, the authors used the free static programming language ‘R’ with the respective package ‘Shiny’.

The user can upload new data – exported from different IT systems – in a specific predefined format. After uploading, automated data cleaning and an outlier detection process were executed. Several algorithms prepared the data for the next step of feature generation.

The planner can select a set of different predefined features (e.g. the average lead time of the last x lots, where x can be any positive integer number defined by the user, for more see [19]) and add those features to the dataset (1). Some of the features are parameterisable, as explained above. The user can try different parameters and add the same feature with different

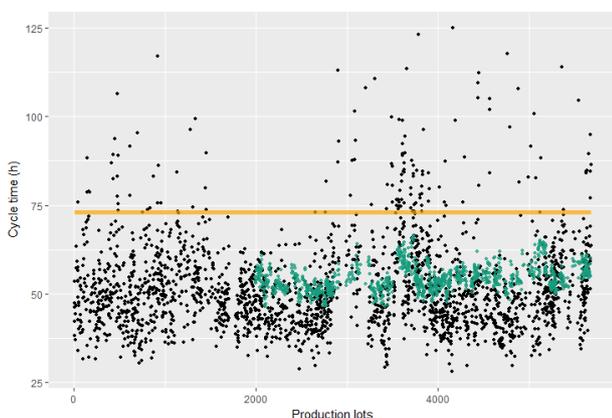


Figure 1. Comparing the real and predicted cycle times of layer 1 with the currently used static cycle time.

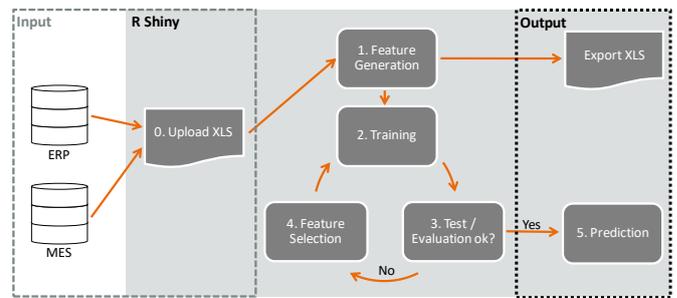


Figure 2. Shiny web application scheme.

selected parameters. In the next step, the user can train (2) different models. He is able to set a number of observations for the training and a number of predictions before the next training. For evaluation, he gets the NRSME (3) and can also see the predicted values and compare them to the currently used target values. Based on the results, he can adapt the parameters of the features and select the best features (4) until he reaches an acceptable accuracy level. Lastly, there is also the possibility of getting a prediction of a certain lot by entering the values of the features (5). An automated implementation of the developed prediction model for the current planning process is in progress, as the features used for training must also be calculated from the production plan itself. For this reason, new approaches need to be developed and researched further in order to be integrated into the application.

What can be clearly seen is that even if the underlying planning data is updated regularly (in the current case, every three months), the cycle time is highly dependent on several influencing factors that cannot be depicted in a single, static planning value. Numerous influencing factors can be exploited with the help of different ML methods. This knowledge – gained from the available dataset – can be used to improve the quality of PPC in CPPSs.

5. APPROACH OF THE INTEGRATION OF MACHINE LEARNING WITH PRODUCTION PLANNING

In this section, the authors pick up the ideas from section 3 and discuss the two possible approaches mentioned above. In the first approach, the planning times are adapted to match the predicted times in several iterations. This approach is called the ‘evolutionary approach’. In the second approach, the dynamic time models are used directly for plan creation. This approach is called the ‘function-based approach’.

5.1. Evolutionary approach

The overall approach consists of six steps, numbered from 0 to 5, while steps 1 to 5 run in ongoing loops and are depicted in Figure 3. It is assumed that in every iteration, PQ increases. This iterative process aims for continuous improvement and is the reason why the authors named the approach ‘evolutionary’.

The planning starts with the generation of an initial production plan, whereby the planning itself is carried out by the conventional planning system, e.g. the ERP or MES of the company, and the planning data is extracted from the master data of the underlying system (steps 0 and 1). The production plan is the input for the second step. In step 2, ML models that have been trained by utilising historical data from the company beforehand are employed to predict different planning times, such as the cycle time, setup time, and operation time. In step 3, the deviations between the planned times and the forecasted

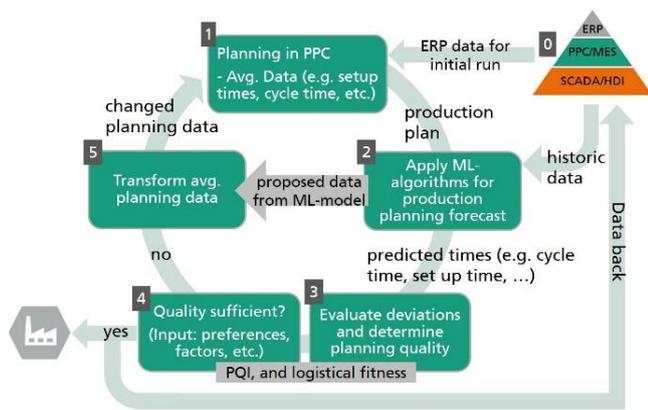


Figure 3. Evolutionary approach.

times and the impact thereof are evaluated. The results are summarised in a dashboard, showing the logistical fitness as proposed by Lödging et al. [25], and the newly developed PQ index. Depending on the impact of the deviation, the quality gives feedback about the reliability of the production plan. In the fourth step, the planner decides whether the production plan meets the preferences and objectives of the company or not. If not, it is possible to change the planning related master data by using the support system, which provides the planner with some alternative proposals to change the data. The overall decision and therefore the responsibility stays with the human (i.e. a human-in-the-loop model) [18]). After the refinement and adjustment of the planning data, the planning can be re-initiated, and the cycle starts again.

If the planner is satisfied with the Key Performance Indicators (KPIs) in step 4, he can activate the plan, thus transferring the planning related master data that was used to create the production plan to the planning systems. The innovative characteristics of the proposed evolutionary approach are as follows:

- the provision of an evaluation method for the assessment of the production plans in step 3,
- the creation of a dashboard to comprehensibly provide feedback to the planner through the visualisation of the main KPIs, and
- the establishment of a method for the proposition of more suitable planning data.

5.2. Function-based approach

The second approach differs – the planning cannot be carried out by the standard/conventional planning system. While in the evolutionary approach, the quality of the production plan continuously increases, in the function-based approach, PQ is always set to high, and other logistical fitness factors of the plan are optimised iteratively, as shown in Figure 4.

Step 0 is necessary for generating the prediction models and therefore must be done regularly. Based upon these ML models of planning-related data, a novel planning method generates production plans. The novelty of the planning models lies within the usage of the dynamic set of functions (i.e. an extendable vector of possible functions) for the planning-related data and not a single value or set of values. The general approach seems as follows: orders are subsequently scheduled, always taking the current information of the production plan as information for the prediction models of the planning related data. Some features for the prediction are not available at the time of planning, and

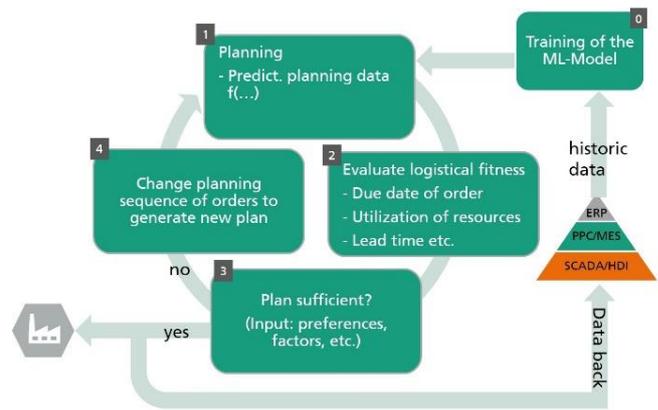


Figure 4. Function-based approach

therefore, these dynamic elements are predicted using different models. As there is no comparable approach in the field of research, this is the highest overall innovation within the approach.

After planning, the evaluation of the logistical fitness follows. The logistical goal weighting of the company (e.g. an average lead time of six weeks and the timeliness of 87 % of all orders) models the benchmark that a production plan must achieve. Since PQ is already high, only these KPIs must be checked. In step 3, the planner again decides whether the plan is acceptable or not. If not, they try another planning sequence. As a support function, the system offers several strategies (based on the order size, due date, customer priorities, etc.). Additionally, the planner can make manual changes. If the sequence is defined, they start the planning again. These steps are repeated until the plan is satisfactory. If so, the plan can be transferred/released to the ERP or MES. In sum, the innovative characters of the proposed function-based approach can be specified as follows:

- Providing a novel planning method for dynamic times.
- Creating a recommendation method to propose changes in the planning sequence.

6. RESULTS AND DISCUSSION

The objective of both approaches is to increase PQ in the sense of smaller deviations between planning and subsequent execution. The key is in the incorporation of uncertainty in the planning phase by carrying out production planning on the basis of dynamic instead of static time values. Only with the use of dynamic time values (e.g. standard times for machining and setup) can the interdependencies of different influencing factors, which occur naturally, be depicted close to reality. These influencing factors can be well known, but most planning system do not offer the possibility of considering them in planning or scheduling activities. In the worst-case scenario, these influencing factors remain unknown or at least undetected. An example of the former case could be fluctuations in the machining and setup times per week and/or shift. This is commonly caused by different skill levels of employees or the actual machine or tools that execute the job. Even though these effects are well known, most companies do not have the resources to levy and document all the effects so that a system would be able to process the information. Still, this information is available in, for example, the confirmation data, and ML is able to quantify the impact to some extent.

In order to be able to represent these dynamics as realistically

as possible by using an ML algorithm, it is essential to ensure high data quality. In both approaches, the prediction provided by the ML algorithm can be as reliable and valid for the actual production system as the underlying historical data source that is used. For further discussion, we focus on different data sources for MES data. We distinguish between Plant Data Acquisition (PDA) and Machine Data Acquisition (MDA) data. MDA data can generally be considered as more reliable than PDA data, as the actual machine status is automatically captured. For example, a status like 'no spindle rotation' and 'no malfunction' indicate that the process has obviously been completed, and a time feedback via MDA is done correctly. In the same case with feedback via PDA, there can either be a time delay in the feedback of the machine operator or feedback is entirely missing.

Therefore, in the case of PDA feedback, higher fluctuations and lower-quality training data are to be expected in comparison to MDA data. However, this does not mean that the approach is only applicable for MDA-generated data or data that is generated with a simulation model of the production system. Longer observation periods and therefore a larger data source help to increase the reliability. In conclusion, it can be stated that the use of data that reflects reality is essential.

To ensure that the machine status is mapped in production planning, it is also important to consider them as input features of the ML algorithm. Depending on the input data of the machine status, the underlying causes should rather be questioned instead of a direct link between the change of the data and a change of the production plan. For example, following the assumption that the machining times of the machine increase over time, a variety of causes can be assumed, such as a change in the condition of the machine tool or the production machine (e.g. blunt tool or slow feed). In this case, it seems reasonable to invest resources to adjust the machine condition instead of generally increasing the machining times. It is important to provide the production planner with a decision support system in order to present decision options and their effects on the planning system. It seems conceivable to constantly increase the processing times due to a lack of investment in old machines or to restore the original machine condition. In the second case, the actual machining times should be significantly reduced. A major advantage of the function-oriented approach is the time-consuming master data maintenance that is not required. Instead, the approach generates benefits through the dynamic adjustments of the relevant values. However, it should be noted that the logic of the ML algorithm can only create a production plan on the basis of the current data situation. Missing or defective data usually leads to a reduced PQ. Even if the condition of the machine worsens, immediate repair of the machines and the associated shortening of the processing times is generally not to be expected, nor is it appropriate. The permanent monitoring and visualisation of the machine status can alternatively be used to initiate measures to reduce production time. An example is the maintenance or change of the production parameters. As an ideal solution, the integration of a logic into the ML algorithm can be considered, which enables the examination of data from different data sources, machines, etc. The algorithm can assist in the interpretation and decision-making or, if required, perform appropriate weighting of individual data and states. By weighting, the effects of individual influences on production planning can be adapted, and implicit knowledge of planning can be represented to further increase PQ.

6.1. Evolutionary approach

Since this approach uses the existing production planning system, there are fewer interventions in the current existing system. This means that the existing systems do not become obsolete and can still be used to generate production plans. The evolutionary approach is meant to be an additional Decision Support System (DSS) to the actual planning system. Within this DSS, a comparison is made between the planned times and the predicted times (cycle time, setup time, operating time, etc.). Feedback to the planner is given in the form of various KPIs and the PQ index. Furthermore, the planner gets feedback on why a certain time prediction is different from the planned time. Based on the chosen ML method, 'drivers' for predicted times can be identified, and a note to the planner is given. Since the system is an additional system, the implementation is expected to be easier and quicker. Furthermore, the acceptance by the planners is expected to be higher, as the actual planning is still done by the existing system. The transparency and sovereignty of the planner about decision-making is an important factor for the acceptance of employees and proves a significant advantage of this approach.

6.2. Function-based approach

As the production plan is created directly using the time prediction models and these models get updated frequently, the quality is independent from the stored master data. Furthermore, the 'dynamic' models are always updated automatically. The authors expect this approach to require a smaller number of iteration cycles because the system always uses the planning data that most likely depicts the later execution. The iterations are needed to meet the objectives for the logistical KPIs. Therefore, the number of iterations is expected to be similar to the current number of iterations, which are needed for the creation of a proper plan today. However, there are several open research questions that need to be answered. Within the function-based approach, the planning algorithm chooses one order after the other. This means that the prediction only considers orders that have already been planned. However, subsequent orders have an impact on the features that are used for the prediction; for example, on the work in progress, when an order arrives at the same workstation, and the prior planned order is not finished. It is therefore crucial to derive correlations from planned orders and unplanned orders in the planning process. According to the authors, further research is needed to define a way in which the necessary features for the time prediction can be determined.

The proposed approach is designed to replace the current planning system with a new planning algorithm. The planner has the option of weighting the KPIs (timely delivery, lead time, stock, etc.) and thus manually adapting the priorities for production planning. Since the PQ index is always high, it is important to check the corresponding KPIs and only implement the production plan when the KPIs deliver sufficient results. A further advantage of this approach is that in contrast to the evolutionary approach, no adjustments of the master data have to be carried out. Instead, a collection of the feedback data is necessary for training the ML algorithm.

6.3. Evaluation of the approaches

For the functionality and correct use of both approaches to increase PQ, it is also important to determine when and which approach is applied. Therefore, in addition to further developing the approaches, attention should be given to creating an evaluation method for the two approaches. This should clearly

demonstrate the advantages of each approach as well as the complexity of implementation. It should also be ensured that the right approach is always used depending on the application, the industry, and other influencing factors in order to achieve optimal results.

6.4. Implementation in ERP/MES

In order to ensure that both approaches are fully functional and can be used separately from the implemented ERP or MES solution, they have to be made accessible regardless of the platform that is used. Different variants of ERP and MES systems in the industry do not represent restrictions for the application of new approaches. Therefore, the transfer of historical data for the training of the ML model is a general solution and must guarantee a data transfer at any time, but the training data can consist of individual characteristics depending on the application case. Even if the accuracy of the production plan prediction increases with the amount of training data used, it is important not to select a disproportionately large amount of data. The horizon of the data also has an important influence. The use of long-term data ensures that the results are based on a long history, while the use of short-term data is suitable for the representation of outliers and random events. The storage of data duplicates in MES and ERP has to be prevented as well.

In each planning run, only the delta from old and new data is to be transferred, which guarantees speed advantages in the data transfer and keeps the transfer duration low. In fact, data transfers are not only possible for order confirmations but also for partial confirmations via MES or ERP. According to the current status, however, only the historical data of a completed order (order confirmation) is used to forecast production plans. Since the time horizon of the data used can influence the result of the prediction, it is also possible to differentiate between two variants of time-based use. In variant 1, the time horizon extends over the entirety of all training data, with all previous training data having an unknown influence on the results. In variant 2, the time horizon can be individually selected according to the principle of ongoing planning, which has advantages, for example, for the more accurate reproduction of short-term events. It is assumed that old data is less valid and can lead to distortions. In the further course of the research project, the advantages and disadvantages of different time horizons on the quality of production planning will be investigated.

An essential basis for the successful introduction and use of these approaches in ERP or MES is an analysis and optimisation of the existing process flows. A large part of the benefits that can be realised by the integration can hardly be evaluated in advance using quantitative criteria. The improvement of the internal production planning is an essential benefit of the presented solution, but its valuable support in the entire order processing is difficult to present. The benefit assessment in particular is considered problematic, since only parts of the achievable benefits can be quantitatively assessed in advance; for instance, adherence to delivery dates or error avoidance. Another element that needs to be considered is the basic hardware requirements at the factory level that the constant use of ML algorithms in planning entails. A quantification of the required computing power is evaluated in the course of a research project.

7. CONCLUSIONS AND OUTLOOK

During the work, a case study for an application of ML for time prediction was given, the term PQ was defined, and two

approaches to incorporate ML models for time prediction in the production planning process were discussed.

The case study shows how ML can be used to predict cycle times based on the readily available data of a semiconductor manufacturer. The results show that the prediction is more accurate than the quarterly adapted static cycle time – the arithmetic average of cycle times of comparable lots within the past quarter (period of three months starting with January). The problem lies in the application of the trained models in the planning phase. Therefore, two approaches of how to utilise these trained ML models during the plan creation are presented. The two approaches are the evolutionary approach and the function-based approach. In the first approach, new plans are created repeatedly, while the used planning times are adapted after each cycle. In addition, PQ is measured by calculating the deviation between planning times and predicted times. If the deviation is within an acceptable level, the planning stops. The latter approach uses functions instead of static values for the planning times. Therefore, there are no deviations between planning times and predicted times. Still, several planning cycles are needed, as other logistical KPIs need to be optimised.

In the first place, an evaluation method for the two approaches should be developed. The function-based approach appears to be smarter due to the usage of a new planning method based on historical data, creating a production plan immediately instead of creating a production plan with master data and then comparing it with predicted data. However, both approaches have to be examined more closely in the further proceedings of the research before an exact assessment can take place.

Therefore, the authors successfully submitted a research proposal called 'MLinPPC' (project number 877446) to the 32nd 'Produktion der Zukunft' call of the Austrian funding agency 'Österreichische Forschungsförderungs-gesellschaft mbH'. With the approval, the two proposed approaches will be implemented and tested together with two industrial companies. Furthermore, the whole process, from data recording to PQ evaluation, will be realised in order to assess both approaches.

In fact, production plans are currently created from ERP or MES data. The acceptance level of employees in the planning department depends strongly on the recognisability of deviations between the predicted and the classic (ERP/MES) production plan. This is a clear advantage of the evolutionary approach. The confidence of the planner in a production plan that was created by an unknown planning logic could initially be low and thus restrict the implementation. Therefore, in addition to the development of the algorithms themselves, further research must be carried out to enable the planner to trace the result of the algorithms. This is especially important during the deployment phase of the new approaches.

Therefore, it becomes clear that no matter which approach is chosen, the transparency of the approach must be consistent and comprehensible for the employee. Only if the two approaches are accepted and trusted can a correct implementation be achieved and the PQ in production planning increased, which will lead to an optimisation of logistics target values.

ACKNOWLEDGEMENT

The authors would like to acknowledge the financial support received within the EU project: Power Semiconductor and Electronics Manufacturing 4.0 (SemI40), which is funded by the programme Electronic Component Systems for European Leadership (ECSEL) (Grant Agreement No. 692466).

REFERENCES

- [1] G. Schuh (ed.), Ergebnisbericht des BMF-Verbundprojektes PROSense. Hochauflösende Produktionssteuerung auf Basis kybernetischer Unterstützungssysteme und intelligenter Sensorik, 2015, pp. 7-8.
- [2] F. Ansari, M. Khobreh, U. Seidenberg, W. Sihm, A Problem-Solving Ontology for Human-Centered Cyber Physical Production Systems, *CIRP Journal of Manufacturing Science and Technology*, Elsevier, 22C (2018) pp. 91-106.
- [3] L. Monostori, B. Kádár, T. Bauernhansl, S. Kondoh, S. Kumara, G. Reinhart, O. Sauer, G. Schuh, W. Sihm, K. Ueda, Cyber-Physical Systems in Manufacturing, *CIRP Annals – Manufacturing Technology* 65 (2016) pp. 621-41.
- [4] F. Ansari, Cyber-Physical Systems, Bericht für Ergebnispapier, 'Forschung, Entwicklung & Innovation in der Industrie 4.0', Verein Industrie 4.0 Österreich (2018) pp. 26-28.
- [5] T. Bauernhansl, M. Hompel, B. Vogel-Heuser, Industrie 4.0 in Produktion, Automatisierung und Logistik, Anwendung, Technologien, Migration, 2014.
- [6] L. Yongkui, X. Xun, Industry 4.0 and Cloud Manufacturing: A Comparative Analysis, *Journal of Manufacturing Science and Engineering* 139 (2016).
- [7] A. Sanders, C. Elangeswaran, J. Wulfsberg, Industry 4.0 Implies Lean Manufacturing– Research Activities in Industry 4.0 Function as Enablers for Lean Manufacturing, *JIEM* 9(3) (2016) p. 811.
- [8] S. S. Kamble, A. Gunasekaran, S. A. Gawankar, Sustainable Industry 4.0 Framework – A Systematic Literature Review Identifying the Current Trends and Future Perspectives, *Process Safety and Environmental Protection* 117 (2018) pp. 408-425.
- [9] A. Hees, G. Reinhart, Approach for Production Planning in Reconfigurable Manufacturing Systems, *Procedia CIRP* 33 (2015) pp. 70-75.
- [10] G. Schuh, T. Potente, S. Fuchs, C. Hausberg, Methodology for the Assessment of Changeability of Production Systems Based on ERP Data, *Procedia CIRP* 3 (2012) pp. 412-417.
- [11] P. Nyhuis, T. Heinen, C. Rimpau, E. Abele, A. Wörn, Wandlungsfähige Produktionssysteme. Theoretischer Hintergrund zur Wandlungsfähigkeit von Produktionssystemen, *Werkstattstechnik online* 98 (2008) pp. 85-91.
- [12] H.-P. Wiendahl, H. A. ElMaraghy, P. Nyhuis, M. Zäh, H.-H. Wiendahl, N. Duffie, M. Brieke, Changeable Manufacturing – Classification, Design and Operation, *Annals of the CIRP* 56(2) (2007) pp. 783-809.
- [13] T. Tolio, M. Urgo, J. Váncza, Robust Production Control Against Propagation of Disruptions, *CIRP Annals* 60(1) (2011) pp. 489-492.
- [14] D. Gyulai, A. Pfeiffer, L. Monostori, Robust Production Planning and Control for Multi-Stage Systems with Flexible Final Assembly Lines, *International Journal of Production Research* 55(13) (2017) pp. 3657-3673.
- [15] U. Bergmann, M. Heinicke, Resilience of Productions Systems by Adapting Temporal or Spatial Organization, *Procedia CIRP* 57 (2016) pp. 183-188.
- [16] D. Gyulai, B. Kádár, L. Monostori, Robust Production Planning and Capacity Control for Flexible Assembly Lines, *IFAC-PapersOnLine* 48(3) 2015, pp. 2312-2317.
- [17] D. M. D'Addona, F. Bracco, A. Bettoni, N. Nishino, E. Carpanzano, A. A. Bruzzone, Adaptive Automation and Human Factors in Manufacturing – An Experimental Assessment for a Cognitive Approach, *CIRP Annals* 67(1) 2018, pp. 455-458.
- [18] F. Ansari, P. Hold, W. Sihm, Human-Centered Cyber Physical Production System: How Does Industry 4.0 Impact on Decision-Making Tasks?, *Proc. of the IEEE Technology and Engineering Management Society Conference*, 27 June-1 July 2018.
- [19] L. Lingitz, V. Gallina, F. Ansari, D. Gyulai, A. Pfeiffer, W. Sihm, L. Monostori, Lead Time Prediction Using Machine Learning Algorithms – A Case Study by a Semiconductor Manufacturer, *Procedia CIRP* 72 (2018) pp. 1051-1056.
- [20] P. Domingos, A Few Useful Things to Know About Machine Learning, *Commun. ACM* 55(10) (2012) p. 78.
- [21] C. Rainer, Data Mining as Technique to Generate Planning Rules for Manufacturing Control in a Complex Production System, in: *Robust Manufacturing Control*, Springer Berlin Heidelberg (Lecture Notes in Production Engineering). K.Windt (ed.), 2013, pp. 203-214.
- [22] A. K. Choudhary, J. A. Harding, M. K. Tiwari, Data Mining in Manufacturing – A Review Based on the Kind of Knowledge, *J Intell Manuf* 20(5) (2009) pp. 501-521.
- [23] B. Csáji, L. Monostori, Value Function Based Reinforcement Learning in Changing Markovian Environments, *Journal of Machine Learning Research (JMLR)*, MIT Press and Microtome Publishing 9 (2008) pp. 1679-1709.
- [24] Y. Cheng, K. Chen, H. Sun, Y. Zhang, F. Tao, Data and Knowledge Mining with Big Data Towards Smart Production, *Journal of Industrial Information Integration* 9 (2018) pp. 1-13.
- [25] H. Lödding, *Handbook of Manufacturing Control – Fundamentals, Description, Configuration*, Springer, Berlin, Heidelberg, 2013, 978-3-642-24458-2.