Available online at www.sciencedirect.com

ScienceDirect

Procedia CIRP 86 (2019) 37–42

www.elsevier.com/locate/procedia

7th CIRP Global Web Conference
"Towards shifted production value stream patterns through inference of data, models, and technology"

Towards a connected factory: Shop-floor data analytics in cyber-physical environments

Dávid Gyulai[a,*], Júlia Bergmann[a], Viola Gallina[b], Alexander Gaal[b]

[a]Centre of Excellence in Production Informatics and Control (EPIC), Institute for Computer Science and Control (SZTAKI), Hungarian Academy of Sciences (MTA), Budapest, Hungary, Kende str. 13-17, H-1111 Budapest, Hungary
[b]Fraunhofer Austria Research GmbH, Theresianumgasse 27, A-1040 Wien, Austria

Abstract

In the era of the industrial digitalization, the availability of shop-floor data is not a question anymore, but rather the exploitation of the underlying information. With advanced sensor technologies, detailed data can be obtained about products, resources and processes in near real time, however, still there are gaps between the collection of the data, and the utilization of it. The greatest current challenge is the use of available data in decision-making processes that brings the real business value for companies, to keep their competitiveness and internal efficiency. In the paper, a reference model of an industrial data analytics platform is presented that supports the integration of various analytics solutions with enterprise level decision support tools, such as planning and scheduling systems. The reference model is composed of various layers, supporting the collection, storage and analysis of data coming from various sources. In addition to its business intelligence related dashboarding and visualization functions, it provides the opportunity of linking the analytics results with other software applications. In order to highlight the capabilities of the proposed model, possible application domains and use-cases are presented, reflecting real industrial needs.

Keywords: Data analytics; Internet of Things; Cyber-Physical Systems; Production management

1. Introduction and motivation

Manufacturing data collection and analysis gain more attention from the industry in the era of industrial digitalization, under the umbrella of *Industry 4.0*. The solutions that rely on the *Internet of Things (IoT)* and *cyber-physical* technologies enable to launch an increasing number of projects that aim at utilizing the process-related data in production management. In a general case, the project stakeholders are interested in solutions that provide more insights in the processes, in order to understand better the underlying causalities and root-causes of certain phenomena, related to the manufacturing processes. The success of these projects is typically measured by process performance metrics, and the return on investment associated with the project itself. Even in case a high-quality and sufficient amount of data is available for the analysis, preliminary statement of the outcomes is often impossible to make as the nature of the processes and complex relationships among various factors may influence the target metrics. Hence, it might happen that the available data cannot contribute to bring the expected results. This can lead to defensive corporate strategy, namely that the management hesitates or beware of starting an analytics project. The lack of confidence is mostly associated with the uncertain return on investment, while it is important to highlight that the project efforts can be high, considering both the expertise and time factors.

This paper aims to introduce a reference architecture for industrial data analytics, providing a guideline to achieve the expected results, increasing both the effectiveness, as well as the return on investments in analytics projects. The presented methods contribute to obtain information from process-level (low level in the hierarchy) data, and provides some ways to elevate this data onto a higher level in the corporate decision making hierarchy, supporting e.g., the material replenishment, network planning or investment decisions through shop-floor data analytics.

* Corresponding author. Tel.: +36-1-279-6139
E-mail address: david.gyulai@sztaki.mta.hu (Dávid Gyulai).

10.1016/j.procir.2020.01.016

2. Data analytics in the manufacturing industry

Recently, an increasing number of industrial analytics applications are integrated in IoT platforms. They enable to implement workflows and pipelines from the source of the data towards its utilization, in the form of descriptive, predictive or even prescriptive applications [9]. The latter is considered to be the future of corporate decision making, as it provides the opportunity to optimize processes with a foresight on its future state, by predicting the outcomes of various scenarios with the analytics models [3]. IoT platforms are essential elements of connected factories, as they provide the opportunity to collect data from multiple sources, and store and analyze them in a common system [12]. There are several platforms available for industrial deployment (most of them are domain independent), and both commercial and open-source tools available. Considering the industrial use, typical application cases include decision support for predictive maintenance, predictive quality control, inventory forecasting and logistics analytics [5, 10, 14], however, there are only a few use cases that consider the direct application of the analytics models in production planning and control. Furthermore, data analysis and modelling are often done in the IoT platforms without a feedback link to the enterprise IT systems, e.g. towards the ERP or MES systems. The platforms are typically used for descriptive and predictive purposes, and they only support the decisions indirectly. Conclusively, there are only a few cases when the data collected from the shop-floor is directly used for achieving the managerial objectives through pipelines not only along the IoT platforms, but also across the various systems both in the forward and backward directions. The paper aims at providing some techniques how this gap between the decision making and industrial analytics can be bridged by real prescriptive decisions for production planning and control.

3. Reference architecture for industrial data analytics

The contribution of the paper is brought by the presentation of a reference model for industrial data analytics projects. Naturally, it is almost impossible to set-up an all-round architecture that can be universally applied in every industrial domains. However, the model is built-up on the solid basis of several, diverse analytics projects from the manufacturing industry. A common initial state of these works was the existence of process-related data, and only a blurry corporate vision of the applicability of data in enterprise-wide decisions. Accordingly, many questions were unable to answer in the initial states, however, the common goal was to increase the high-level effectiveness and performance metrics on a data analytics basis.

3.1. Overview of the architecture

The proposed reference model has a hierarchical structure, composed of four main layers that elevate the information from the sensor-level data collection source all the way up to the enterprise level analytics and application (Fig. 1). The `Data Collection Layer` is responsible for sourcing the data from the processes, and transferring it to the `Data Lake` via the proper communication channels and protocols. On the latter layer, poli-base, either permanent of temporary data storage is implemented, with the help of relational or non-relational database technologies. Considering a single main storage in the form of the `Data Lake`, the `Data Processing Layer` includes all the algorithms that enable to access, filter and manage the data, and to obtain some decision-related information from it. On the top of the hierarchy, the `Data Analytics Layer` takes place that includes all the models and software libraries that enable to make the entire system smarter by the data-driven processes.

3.2. Data collection and storage layers

The `Data Collection Layer` implements the link between the physical and cyber parts of the system, by the application of sensor networks and even manual data collection interfaces. In this way, detailed data about the products, processes and various assets can be collected even in near-real time. The data collection is done through various communication channels and network protocols that meet the industry standards. Recently, wireless communication is gaining more attention from the industry that enables a flexible way of making the system and products smarter. Among others, the major requirements towards the wireless sensor communication protocols are *compactness, scalability, robustness* and *low energy consumption*. Typical IoT communication and messaging protocols include WLAN, ZigBee, Bluetooth, MQTT etc. [6].

The `Data Lake` is responsible for storing all the data collected from various sources, applying a raw storage format. In contrast to the data warehouses, structured (relational), semi-structured (XML, CSV) and unstructured data formats can be all applied. They are stored in a poli-base structure, including relational databases (accessible with SQL techniques), non-relationl (NoSQL) databases and also files. Having data from multiple sources available in a common layer, cross-functional business analytics can be performed to obtain real business value by analytics.

3.3. Data processing and analytics layers

The `Data Processing Layer` implements the functions that are needed to ”catch” the proper data from the lake in the correct, predefined format. For this purpose, database queries and file parsing approaches are used, and in many cases, various filtering and preprocessing methods are needed to get real useful information from the raw data. Such methods implement data cleansing, data transformation functions, and in many cases, aggregation functions are also applied to reduce the volume of the data that is managed in the latter stages of the analytics process chain.

In order to utilize the data collected an processed so far, analytics, simulation, machine learning and visualization models are deployed on the `Data Analytics Layer`. These models directly provide results for the end users to get insight and fore-

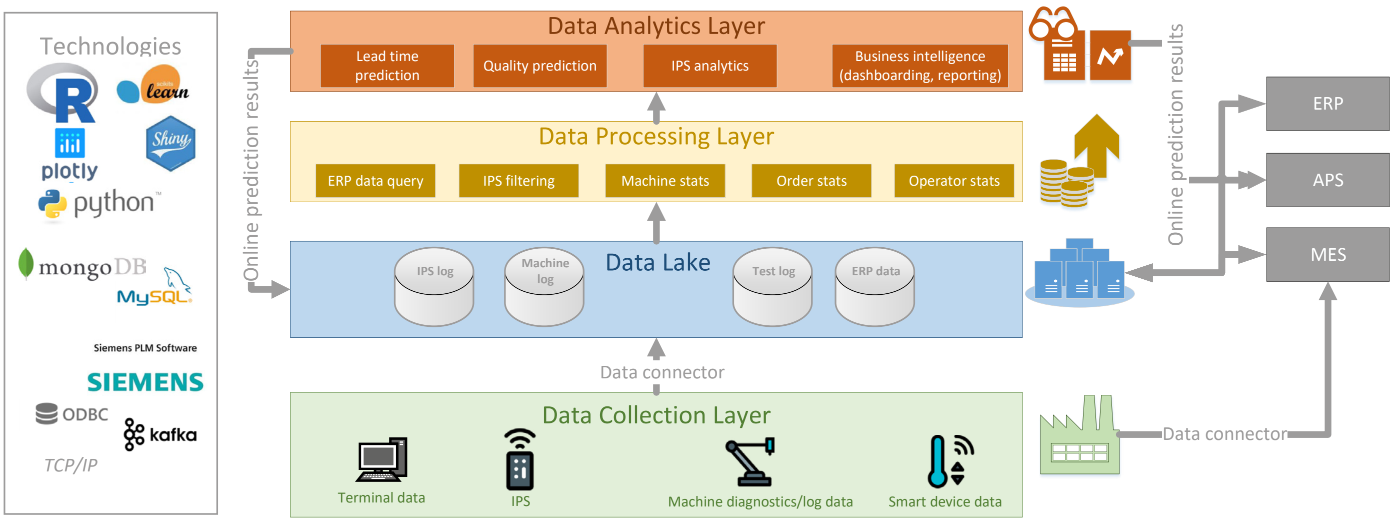

Fig. 1. Reference model for industrial data analytics with exemplar items on the various layers and technologies applied in the presented use case.

sight in relation with the shop-floor processes. Furthermore, the analytics models should be linked with the enterprise IT system, therefore, decisions made by using the ERP, MES and APS systems can be driven by data analytics. This data link is essential part of the reference model, and often disregarded by general purpose IoT reference architectures. Having proper models and analytical pipelines implemented through the various layers of the reference model, the analytics models and the linked IT systems can support decisions relying on multiple criteria, various factors and always on the actual state of the production system, so as implementing a situation-aware decision making process. To this aim, two major model classes are deployed: batch and stream analytics models. The former models are periodically updated/trained with the latest stage of the system, while the latter models rely on chunks of data, streamed continuously. Both model types have pros and cons, and the selection is always based on the business objective and the production environment.

4. Application case

The following sections provide some insights how the different layers can be implemented in a common analytics framework of a connected factory. Different data-driven approaches are integrated in the framework, all implemented in industrial projects. The purpose of the use cases is to justify the idea that the proposed model enables the analytics of shop-floor data, and capable of utilizing it in higher level decision making processes to obtain real business value through analytics.

4.1. Description of the production system and collected data

For demonstration purposes, a single production environment is considered, in the form of a simulation model. It represents the characteristics of the real production systems, and in order to prove the effectiveness of the methods, the original numerical test results are also provided. It is highlighted that the target level of KPIs can be achieved via advanced data analytics and machine learning methods, utilizing the data that is continuously collected about the processes.

The production environment of the test case consists of three main segments: a machinery job shop, a surface treatment area and an assembly segment. In the machinery, 29 CNCs from three types are available to perform the cutting processes, conclusively, alternative resources are considered when preparing the schedule of the system. The machines are automated, and each job visits multiple machines, depending on its cutting tasks. In contrast to the machinery, products are moving in batches through the surface treatment segment into a mid-progress buffer, where semi-finished goods are stored. Four identical, manually operated flexible assembly lines are available to assemble the finish goods, regarding the product complexity, 12 main finish good types are assembled using 6 main parts that are machined in the first segment.

As for the collected data, a sensor network responsible for monitoring the system is completely implemented in the `Data Collection Layer` of the reference model. In the machining job-shop system, the machines are equipped with sensors that are capable for measuring the main cutting parameters, e.g., cutting forces, torques, vibration etc. The sensor signals are collected and processed continuously, so as implementing an online monitoring system. While it is relatively easy to collect data about machining processes, manual assembly processes are often hard to be monitored efficiently. Among the advanced technologies of the industrial digitalization, indoor positioning systems (IPS) have received higher attention from the manufacturing industry, as they provide the opportunity of tracking and tracing assets in shop-floor environments more efficiently than previous related solutions. IPSs can be used for locating almost any kind of physical asset in a production environment; typical examples are the tracing of products, tools and fixtures. The assembly lines in the use case are all equipped with IPS, in order to track the product movements, and derive KPIs from the location data.

The main data that an IPS system can provide include the time-indexed spatial coordinates of the assets equipped with sensor tags. The tag locations are captured by anchors installed on the ceiling of the shop-floor. The raw sensor data is filtered, and triangulation algorithms are applied to calculate the X-Y coordinates of a tag, with a certain (typically 10-20 cm) accuracy. Despite the filtering algorithms—typically *Kalman-filters* [2]—deployed on the edge devices (gateways or anchors), the physical environment usually disturbs the positioning process, resulting in outlier positions in the time series of the tracked coordinates. As for the data collection, tag locations are streamed over TCP/IP sockets with a 5 seconds sampling frequency, in JSON format, including the ID of the tracked tag, its raw (unfiltered) X and Y coordinates and the corresponding timestamp.

4.2. Data storage, processing and transformation methods

Based on the real time monitoring of the machinery and assembly segments, data needs to be stored permanently or temporarily in the `Data Lake` and then, it is fed into predictive machine learning models. The IPS tag coordinates are stored in the `Data Lake` as time series (timestamp and the corresponding spatial data) in a NoSQL database, provided in the form of *MongoDB* collections [4]. The analyzed assembly lines have a dimension of cca. 15x15 meters, and the workstations have a cca. 1x1 meters size. The IPS system has an accuracy of around 30 cm, however, some position errors are observed (∼100 cm), due to the disturbances resulted by the physical environment. Similarly to the IPS logs, machinery data is also stored in a NoSQL data collections, while in contrast, MES data—e.g. job release and completion times—are managed in relational databases *(MySQL)*. Considering the operation of the `Data Lake`, the volume of the stored data increase sharply together with the operation of the system, and gigabytes of data are collected in every hour of production, resulting in a massive dataset to be processed and analyzed.

Essential part of the `Data Analytics Layer` is the IPS noise filtration, in order to obtain the real production parameters in the assembly area. As a first step of spatial data cleansing a *Savitzky-Golay (S–G)* filter [11] is applied to remove the noise from the position logs. The S–G digital filter can be applied to a set of data points for the purpose of smoothing, that is, to increase the precision of the data without distorting the signal tendency. This is achieved in a process known as convolution by fitting successive subsets of adjacent data points with a low-degree polynomial by the method of linear least squares (Fig. 2). Having a smoother position log obtained, a knowledge-based filter is applied to match the refined coordinates with workstations (zones), utilizing the technological data of the processes. Matching the observed spatial data with a predefined routing consists of two parts: first, the smoothened data must be dragged onto the theoretical routing of the product, and then a probability-based correction is applied. As a result, the spatial time series is converted into a series of workstations indexed with time, which enables to calculate the cycle times, failure and utilization rates.

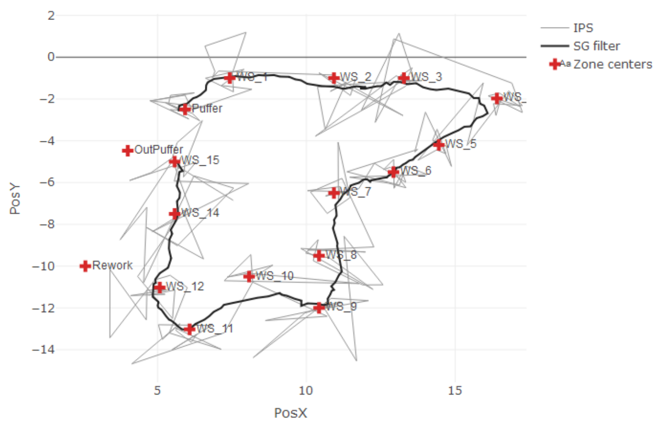

Fig. 2. Position coordinates before and after the filtering.

4.3. Analytics and prediction models, dashboarding

Within the use-case, different data-driven decision support functions are implemented that enable to achieve higher level managerial objectives with the help of manufacturing data. In addition to these functions, a central visualization and dashboarding tool is also part of the `Data Analytics Layer`, supporting the users to perform cross-function analytics, interact with the tools and with the system itself. The dashboard is implemented in *R Shiny* reactive, web-based environment, and provide access to all functions and databases of the connected factory through *Node-RED* flows. Various views and visualization tools of the interface help to understand to processes in near real time, and interact with the system, e.g., through the rescheduling events triggered by certain events. The interface of the web-based central dashboard is illustrated in Fig. 3.

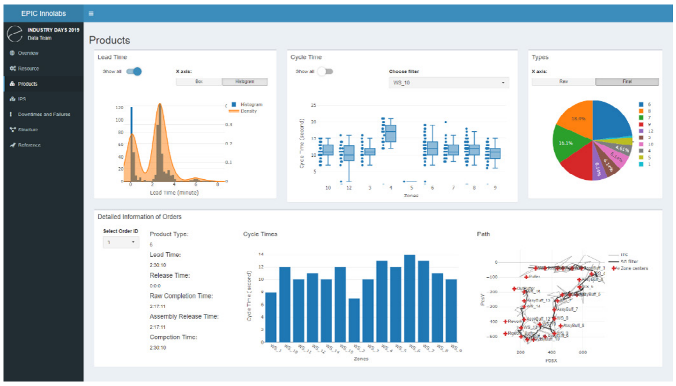

Fig. 3. The web-based, reactive interface of the central dashboard.

4.3.1. Adaptive scheduling

As for the first main analytics function, an adaptive scheduling method is provided. Based on real-time sensor signal processing, machine learning (ML) models are applied for predicting the remaining processing times. Using an ML-based prediction of a possible delay in machining, an initial schedule can be adaptively recalculated in order to reduce wastes and increase machine utilization. Here, it is assumed that predefined cutting parameters as e.g. feed rate or spindle speed are adjusted. which will lead to a change in the expected process duration. Furthermore, considering the quality of the machined parts, ML-models can be used whether the product will be within specification after the current process. In case the product is out of tolerance, the product either needs to be reworked or is a scrap part. The rework-operation triggers an adjustment in the schedule. The determination of a scrap part at early stages reduces wastes. Altogether, these changes lead to the disruption of the entire production schedule, which inherently relies on the real processing time of the jobs. The machine learning models are trained incrementally, and used to predict the remaining process times. The know of this parameter is especially important when cutting processes are long, and therefore, changing process conditions may lead to significant lateness.

As the remaining process time is a numerical value, regression models are deployed on the `Data Analytics Layer`, which provides real time estimations about the timespan until the job completion on the basis of the collected sensor data. By using predefined threshold values, a significant deviation from the expected process time can trigger a rescheduling event, influencing the scheduled release time and machine assignment of all jobs that are not yet released (in-progress) or booked already for machining. In this way, the entire target order set can be rescheduled at an early point of time, and the new schedule might deliver less total waste than rescheduling after realizing the total lateness in the end of the given (late completed) process. According to the numerical results achieved in the original industrial test case, adaptive scheduling provided an approximately 10% shorter execution makespan in average, considering the machinery segment only.

4.3.2. IPS data analytics for robust production planning

In order to implement efficient scheduling in the assembly segment as well, accurate information about the real cycle times—as driver parameters in scheduling—is required. In the analyzed use case, raw IPS logs are processed to calculate the actual cycle times and workstation utilization values in the assembly zone, based on the preprocessing techniques described in Section 4.2. The first results of the IPS-based cycle time analysis—applying the above specified architecture—is described in [7]. According to the test results (Fig. 4), this multi-step approach results in more accurate cycle times, than those obtained by methods relying on the raw logged positions (typical in industrial practice). Conclusively, more efficient decisions can be made on the basis of these parameters, e.g., scheduling or process improvement decisions.

Having trustful assembly cycle times obtained from the IPS system in near real time, the emphasis in assembly planning and scheduling was put on the increasing the robustness of the production plans, associated with the flexible, manually operated assembly lines. For the planning itself, a new simulation-based optimization method was developed to manage the stochastic variables and random events in production planning of multi-product assembly lines [8], through the analysis of cycle time and ERP data. To this aim, simulation metamodels were defined to predict the actual capacity requirements of different produc-

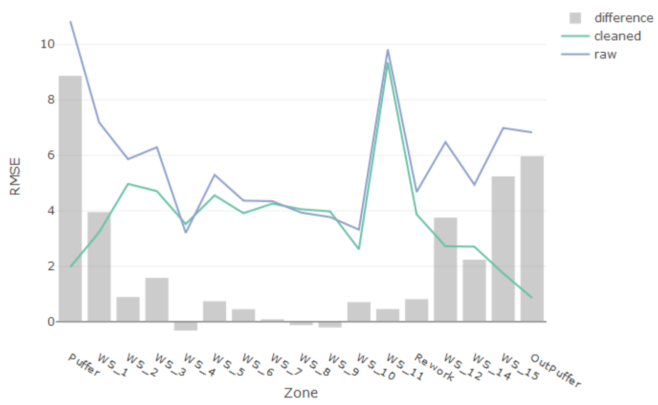

Fig. 4. Root Mean Square Error (RMSE) of cycle times calculated with and without data cleansing and the difference between the two method (green line: cleaned data, blue line: raw data, grey bar: difference).

tion scenarios, instead of calculating the plans according to the idealistic cycle times. The novelty of the planning method is provided by the combination of production data obtained from the IPS and the enterprise resource planning (ERP) systems, facilitating the calculation of robust production plans.

Data analytics is essential part of the overall method, as simulation metamodeling is a data intensive task. It is aimed at defining regression models on the basis of several simulation experiments, in order to capture the system behaviour in an analytical way, which enables to leave out simulation runs from the planning cycle, so as saving significant amount of time that the simulation experiments would require [1]. The greatest benefit of simulation metamodeling is the opportunity to capture the underlying process dynamics, and represent the stochastic nature of certain parameters implicitly, in the form of simple predictive models. The models are built over datasets that are generated by a series of simulation runs, and in case the target training accuracy is achieved, they can be deployed to be part of the decision making process, e.g., production planning in the use case.

In the use case, manual operations and product reject rates in the assembly segment introduce stochasticity and uncertainty in the planning parameters, so as degrading the robustness of the plans. In case the plans are calculated on the basis of ideal, corporate standard cycle times, the variance of these parameters leads to load imbalances, and result in backlogs and lateness. Therefore, the data analytics based robust planning approach relies on the actual production data obtained from the IPS, MES and ERP systems periodically. Accessing the data from the `Data Lake` enables to parametrize the lines' simulation models to represent their quasi actual state. The simulation model of a given line is deployed on the `Data Analytics Layer`, and experiments are performed by considering several random production (order set) scenarios. This results in a dataset with production orders as input, and capacity requirements as output values. Important to highlight that the simulation model is parametrized by taking into account the variance of the time parameters, fitting distribution functions on the actual values obtained from the MES. With the use of the resulting dataset, multivariate linear regression models are periodically trained in the `Data Analytics Layer`, which guarantees that the models used for production planning will always reflect the latest, quasi-actual state of the system. These models are used in the optimization of the production plan, as the simulation metamodels can provide the actual capacities required to assemble a set of orders. The workflow of the implemented robust production planning method is illustrated by Fig. 5.

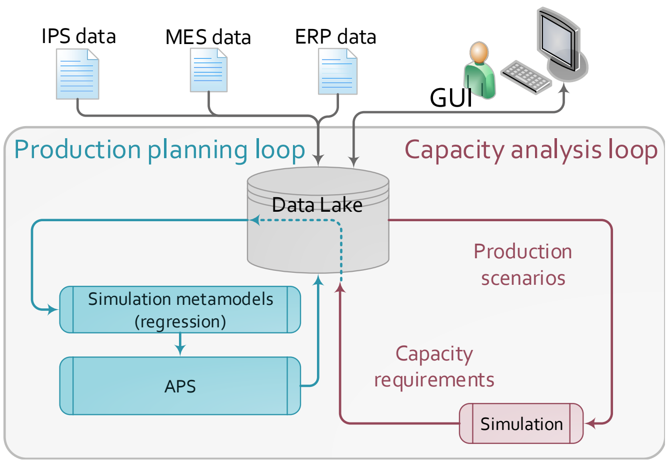

Fig. 5. Data-driven robust production planning method for manually operated flexible assembly lines.

As for the original industrial test case, five days' production was planned on a rolling horizon, considering orders on hand that were known already in the beginning of the horizon, and also those that are placed by the customers during the five days. Nine product types were assembled, of which orders are placed for 36 variants. Comparing the results obtained when applying the robust planning method and the corporate standard method, the resulted total output volumes were quite similar. However, the operators' performance and workload, were substantially better in case of the proposed method, as the average output per operator was 15% higher, compared to the historical values.

5. Conclusions and outlook

In the paper, a reference model for a connected factory was introduced, highlighting various data-driven decision support functions that aim at increasing the effectiveness of production management. This can be achieved if higher level decisions rely on trustful parameters obtained about the actual state of the system. The situation-awareness requires various models and technologies from data collection through processing until analytics and dashboarding, so as implementing multi-perspective analysis on the basis of various data sources. The use cases justified the idea that analysis of shop-floor data can help to understand the causalities in the process chain, and support decision makers to obtain real business value out of the data.

The future work within this research aims at increasing the technology readiness level of the implemented architecture, and also its scalability to reduce latency and achieve faster response times from the analytics models. General purpose machine learning models will be replaced by new models relying on technologies for Big Data, such as *Apache Spark™* [13]. The functionalities of the architecture will be enriched with methods that are available already as separate applications, e.g., lead time and functional quality prediction tools.

Acknowledgment

The research in this paper was (partially) supported by the European Commission through the H2020 project EPIC (https://www.centre-epic.eu/) under grant No. 739592; and also supported by the GINOP-2.3.2-15-2016-00002 grant on an *"Industry 4.0 research and innovation center of excellence"*.

References

[1] Barton, R.R., 1998. Simulation metamodels, in: 1998 Winter Simulation Conference. Proceedings (Cat. No.98CH36274), pp. 167–174 vol.1. doi:10.1109/WSC.1998.744912.
[2] Benkouider, Y.K., Keche, M., Abed-Meraim, K., 2013. Divided difference Kalman filter for indoor mobile localization, in: International Conference on Indoor Positioning and Indoor Navigation, IEEE. pp. 1–8.
[3] Bertsimas, D., Kallus, N., 2014. From predictive to prescriptive analytics. arXiv preprint arXiv:1402.5481 .
[4] Chodorow, K., Dirolf, M., 2010. MongoDB - The Definitive Guide: Powerful and Scalable Data Storage. O'Reilly.
[5] Civerchia, F., Bocchino, S., Salvadori, C., Rossi, E., Maggiani, L., Petracca, M., 2017. Industrial internet of things monitoring solution for advanced predictive maintenance applications. Journal of Industrial Information Integration 7, 4–12.
[6] Gungor, V.C., Hancke, G.P., 2009. Industrial wireless sensor networks: Challenges, design principles, and technical approaches. IEEE Transactions on industrial electronics 56, 4258–4265.
[7] Gyulai, D., Pfeiffer, A., Bergmann, J., 2019. Analysis of asset location data to support decisions in production management and control, in: 13th CIRP Conference on Intelligent Computation in Manufacturing Engineering, CIRP. *In Print*.
[8] Gyulai, D., Pfeiffer, A., Monostori, L., 2017. Robust production planning and control for multi-stage systems with flexible final assembly lines. International Journal of Production Research 55, 3657–3673.
[9] Keramidas, G., Voros, N., Hübner, M., 2016. Components and Services for IoT Platforms. Springer.
[10] Mourtzis, D., Vlachou, E., Milas, N., 2016. Industrial big data as a result of iot adoption in manufacturing. Procedia Cirp 55, 290–295.
[11] Savitzky, A., Golay, M.J., 1964. Smoothing and differentiation of data by simplified least squares procedures. Analytical chemistry 36, 1627–1639.
[12] Soursos, S., Žarko, I.P., Zwickl, P., Gojmerac, I., Bianchi, G., Carrozzo, G., 2016. Towards the cross-domain interoperability of iot platforms, in: 2016 European conference on networks and communications (EuCNC), IEEE. pp. 398–402.
[13] Zaharia, M., Xin, R.S., Wendell, P., Das, T., Armbrust, M., Dave, A., Meng, X., Rosen, J., Venkataraman, S., Franklin, M.J., Ghodsi, A., Gonzalez, J., Shenker, S., Stoica, I., 2016. Apache spark: A unified engine for big data processing. Commun. ACM 59, 56–65. URL: http://doi.acm.org/10.1145/2934664, doi:10.1145/2934664.
[14] Zhang, Y., Wang, W., Wu, N., Qian, C., 2015. Iot-enabled real-time production performance analysis and exception diagnosis model. IEEE Transactions on Automation Science and Engineering 13, 1318–1332.