Available online at www.sciencedirect.com

ScienceDirect

Procedia Manufacturing 54 (2021) 136–141

Procedia MANUFACTURING

www.elsevier.com/locate/procedia

10th CIRP Sponsored Conference on Digital Enterprise Technologies (DET 2021) – Digital Technologies as Enablers of Industrial Competitiveness and Sustainability

Work in Progress Level Prediction with Long Short-Term Memory Recurrent Neural Network

Viola Gallina[a,*], Lukas Lingitz[a], Johannes Breitschopf[a], Elisabeth Zudor[b], Wilfried Sihn[a,c]

[a] *Fraunhofer Austria, Theresianumgasse 7, Vienna, 1040, Austria*
[b] *Institute for Computer Science and Control (SZTAKI), Eötvös Loránd Research Network (ELKH), Budapest, Hungary, Kende str. 13-17, H-1111 Budapest, Hungary*
[c] *Vienna University of Technology, Institute of Management Science, Theresianumgasse 27, Vienna, 1040, Austria*

* Corresponding author. Tel.: +43-676-888-61-646 *E-mail address:* viola.gallina@fraunhofer.at

Abstract

Since the reliability of production plans drops largely within several days after plan creation, production control faces huge challenges, when trying to foresee the work in progress (WIP) level at bottleneck machines and trying to react appropriately. Whereas several researchers applied artificial intelligence to predict lead times or transition times to improve the planning reliability, only small efforts have been taken on time series prediction in the field of production control, especially on the topic WIP prediction. In this paper univarate times series approaches are used for predicting the work in progress for a bottleneck machine for one and more step ahead. Long short-term memory recurrent neural networks, LSMT models show higher accuracy than classical methods. For more step ahead forecasting four different approaches are investigated. Systematical model tuning and comparison of various error measures are presented for a real industrial use case from the steal manufacturing industry.

Keywords: capacity planning; WIP; prediction; time series; LSTM

1. Introduction

Globalisation leads to higher competition for manufacturing companies [1]. They try to cope with this challenge by focusing on the customer needs and offering highly individualized products, leading to a complex product and production structures [2]. Still the logistical performance is of highest importance to customers, which is why short lead times and high adherence to delivery dates are key factors for long term economic success [3]. Production planning and control (PPC) needs to handle this ever growing complexity to ensure a high logistical performance and is therefore crucial for the companies success [2, 4, 5]. In many industries backward planning based on the planned lead time as sum of transition and processing times from the ERP system, is state of the art. The ERP system calculates the release date from the production order based on the customer due date and the planned lead time [6]. Moreover, the so calculated production plan shows the capacity needs at different stations for the future. High prediction accuracy of the workload guarantees an efficient allocation of capacities and ability to meet planned lead time. However, Günter et al. [7] stated that the reliability of production plans can drop to 25% in the first three days after planning, which results in low predictability of the workload. The results can be capacity shortage or material shortage at bottlenecks leading to poor on time delivery. Therefore the work in progress (WIP) level at work centers is of high interest for production controls and is controlled via capacity allocation and order release.

Industry 4.0 offers new possibilities to overcome the challenges described above and to realize the concept of cyber-physical-production systems also influencing the PPC [4, 5, 8]. Dombrowski and Dix as well as Bueno et al. [5, 9] identified that scheduling, capacity planning, and manufacturing control benefit most. Still they came to the conclusion, that further research in the area of smart production planning and control is needed. In this paper a multistep univariate prediction model is presented to forecast the WIP for a particular work center more accurately. The analysed machine is one of the bottlenecks of the production system. The industrial company favours a ten

2351-9789 © 2021 The Authors. Published by Elsevier B.V.

10.1016/j.promfg.2021.07.047

days ahead prediction as the production plan is always fixed on Thursday until the end of next week. In this study one to ten steps ahead prediction models (one step means one day in the use case) are developed with the following models: exponential smoothing, Holt-Winters model and Long Short-Term Memory (LSTM). To the best of the authors' knowledge, univariate LSTM has not yet been applied for regression of the WIP level. Our novel one to ten step ahead prediction model is demonstrated to offer the highest prediction accuracy for the analysed industrial use case, offering a viable solution to the vital problem of timely exploitation of available information for decision support in manufacturing processes.

The remainder of the paper is structured as follows. Section 2 summarizes the related work on prediction of production control KPIs with machine learning methods. Section 3 presents the applied research methodology for one and more steps ahead predictions with model tuning and error measures applied. Section 4 describes the use case, investigates the most suitable model structure and reports the major findings of the paper. Section 5 concludes the results, discusses the limitations and possible extensions.

2. State of the art

Various methods can be used to increase the reliability of production plans. Majority of the research work in the area focuses on time prediction. The simplest analytical time prediction method is Little's law. It states that the average number of items in a queuing system equals the average arrival rate of items to the system multiplied by the average lead time of an item in the system [10].

Although it exists for decades and is proven to be very efficient, in recent dynamic manufacturing systems more sophisticated prediction methods are needed. The application of data analytics and machine learning algorithms has become very popular in the era of cyber-physical production systems. Lead time prediction is one of the most intensively investigated area: regression trees [11, 12], support-vector machines [13], deep neural networks [14] and linear regression models [15] have already demonstrated the effectiveness of these methods. Time series data mining has appeared as a new approach to time predictions – after classical regression data mining models [16]. Lingitz et al. emphasised the importance of using historical data (ERP/MES/SCADA) – instead of the typical approach using simulation model [17].

Predicting the WIP levels on the shop-floor has not yet been so thoroughly investigated as time related forecasts, just a few papers can be found discussing this topic. Misrudin and Foong use simulation for providing a one week output forecast for two work centers [18]. Vazan et al. combine simulation with classification and regression for predicting the behaviour of the manufacturing system and for forecasting particular production goals [19]. The best results were gained with neural networks for each production goal – from which one concentrates on the WIP level. Subramaniyan et al. employ ARIMA method to predict throughput bottlenecks in the production system for future production run from a large set of machine data [20]. The proposed algorithm reaches 62% accuracy compared to the naive method of 24%. The authors analyzed the amount of past data needed for forecasting as well. Choueiri et al. propose a hybrid predictive model for remaining time prediction based on transition-systems and statistical regression – considering WIP levels as one of the most important influencing factors [21]. Scholl et al. presented an implementation of a simulation based short-term lot arrival forecast conducted in Infineon Technologies, Dresden [22]. Their approach considers data from the past 60 days and focuses on one day ahead forecasting for a particular work center. Mosinski et al. introduced the Alternative Forecast Method for daily delivery predictions and bottleneck identification [23]. This research work has a forecast horizon of up to 14 days, is based exclusively on MES data and uses statistical calculations in different databases. Chian et al. stated in their research work that vehicle traffic arrival forecasting exhibited the closest similarity to the forecast of WIP arrival [24]. They apply LSTM for classifying the WIP level into the groups low, normal or high. Their purpose was to predict the WIP level for the next three weeks in three consecutive seven-day long steps. 70% of the 90 day-long historical data with hourly resolution (2160 data entries all together) were used for training. Our methodology shows the closest similarity to this approach, however in our case regression is done and different methods for more step ahead prediction are investigated and compared.

3. Research methodology

The methodological approach applied in our use case orientates on a widely accepted and standard approach for data mining, namely on the CRISP-DM model [25]. The first three phases, business understanding, data understanding and data preparation are consolidated in the section *Dataset description*. The next two sections contain the description of one and more step ahead prediction models, while the applied error measures used for comparing the different models are presented in the *Evaluation* section. The deployment part of the CRISP-DM model will be addressed later on in the research project.

3.1. Dataset description

Based on the procedural model of CRISP-DM, the first stage of the proposed approach focuses on the problem definition and on challenges faced in the industrial use-case. This involves the detection and understanding of bottlenecks as well as the selection of the relevant machine for WIP prediction. From a data perspective, described in the second level of CRISP-DM, production confirmation data is identified as the basis data source for WIP calculation representing the only feature in the proposed approach for prediction. Confirmation data was available from 1.1.2015. This is structured as usual and contains information about the most important data relating a production lot such as order id, date and time, customer, working station, workers, number and dimensions of metal sheets. The yearly average 360.000 data entry for the six year long period resulted in more than 2 million data entries during the whole time period that was reduced to 2190 because of the daily view. However, in the underlying use-case data confirmation is not an entirely automated process, leading sometimes to irregularities and confirmation errors in the collected data set. Thus, the presented approach comprises data preparation and pre-processing steps which correspond to the third stage of CRISP-DM. For accomplishing error detection and data preparation the authors utilize the wide range of available features of the confirmation data set by applying consistency as well as correction rules defined with the company's domain experts. The corrections include adaptions in the sequence of confirmed operations as well as the omission of irrelevant confirmations such as quality checks. The corrected records finally serve as foundation for calculating the WIP time-series of the selected equipment by considering the temporal development of accumulated input and output amounts.

3.2. Prediction with classical univariate time series

In order to accomplish WIP prediction and capacity planning of the regarded bottleneck work center, the authors considered alternative univariate approaches using the WIP timeseries as the only input feature. Starting with the simplest, exponential smoothing is based on the concept of weighted average by assigning exponentially declining weights to past values. Thus, the more recent an observation, the higher the impact on the prediction. Being α the weight value and y the observation series, the one step ahead prediction of simple exponential smoothing is given by:

$$\hat{y}_{t+1} = \alpha y_t + \alpha(1-\alpha)y_{t-1} + \alpha(1-\alpha)^2 y_{t-2} + ... + (1-\alpha)^t \hat{y}_1$$

The Holt-Winters approach extends the approach of exponential smoothing by introducing a level l, trend t and seasonal component s with corresponding smoothing factors α, β and γ. Based on the seasonal behaviour the Holt-Winters approach can be applied in two different variations, expressing either a rather constant seasonal pattern or linear changing seasonality. In the underlying use-case the former so called additive method is applied [26]:

$$\begin{aligned}
\hat{y}_{t+h|t} &= l_t + hb_t + s_{t+h-m(k+1)} \\
l_t &= \alpha(y_t - s_{t-m}) + (1-\alpha)(l_{t-1} + b_{t-1}) \\
b_t &= \beta(l_t - l_{t-1}) + (1-\beta)b_{t-1} \\
s_t &= \gamma(y_t - l_{t-1} - b_{t-1}) + (1-\gamma)s_{t-m}
\end{aligned}$$

3.3. Prediction with univariate LSTM models

LSTM model was developed by Hochereiter and Schmidhuber in 1997 [27]. LSTM is a modified version of a recurrent neural network, that is not able to capture long-term dependencies [28]. In LSTM forget gates are introduced in order to control the information flow. For more information on LSTM see eg. [27] or [29]. In the current analysis exclusively univariates models were developed and just the own historical WIP values were considered.

The model architecture with input, hidden and output layers is very similar to a feedforward neural network. However, data preparation requires more effort, because not just data scaling (between 0 and 1 on our case) but special data formatting is needed as well. After splitting the data set into training and test sets the neurons of the input and output layer have to be arranged in a special array format. The number of historical values used by the training in the input layer is the so called time window, that equals to the number of columns of the input array. The result of the prediction is given in an output array, where the number of columns is identical with the number of steps to be forecasted – one to ten columns for one to ten step ahead predictions, respectively. The number of rows both in the input and output arrays depends on the given data set and the ratio of training-test splitting. The time window value is an additional parameter of the neural network that must be investigated and optimized by the model tuning, besides the usual parameters (eg. number of hidden layers, number of neurons, batch size,...). The best combination of model parameters that results in a high prediction accuracy is desired. This can be achieved with a systematical model tuning, where the authors followed the guideline proposed by Boehmke and Greenwell [30].

In the first part of the analysis the focus was on *one step ahead prediction*. For finding the optimal number of historical data used in the models a grid search was done for time window size with a medium size LSTM model architecture. After calculating the average error measures of five runs the window size was fixed. With the given time window a model tuning was executed in the following steps in order to minimize the mean absolute error of the validation data set : i) model capacity ii) batch normalisation iii) regularization iv) learning rate adjustment. Based on the results the model architecture can be finalised and the focus of the analysis can be shifted to *more step ahead prediction*. The following four approaches can be found in the literature for more step ahead predictions:

- multiple output,
- recursive,
- direct multistep,
- direct recursive hybrid.

In case of the *multiple output* model both the training as well as the prediction are done in one step. The output array has as many columns as many steps ahead the prediction is to be done. The general form of a prediction for the next time periods with the multiple output approach can be formulated as follows:

$$pred_{t+1} = model(obs_t, obs_{t-1}..., obs_{t-w+1}), \tag{1}$$

$$pred_{t+2} = model(obs_t, obs_{t-1}..., obs_{t-w+1}), \tag{2}$$

where t is the current time and w is the time window applied. *Recursive* models need one training step and N prediction step. In the training phase the next value right after the given time window is learned. An N step ahead prediction is done in N recursive steps, where the oldest value is left and the newest forecast value is added, with the following logic:

$$pred_{t+1} = model(obs_t, obs_{t-1}..., obs_{t-w+1}), \tag{3}$$

$$pred_{t+2} = model(pred_{t+1}, obs_t..., obs_{t-w+2}). \tag{4}$$

The *direct multistep* approach requires N steps, where in each step the training and the prediction is done after each other. The prediction models have different weights in the steps because the output of the learning model differs:

$$pred_{t+1} = model_1(obs_t, obs_{t-1}..., obs_{t-w+1}), \tag{5}$$

$$pred_{t+2} = model_2(obs_t, obs_{t-1}..., obs_{t-w+1}). \tag{6}$$

The *direct recursive hybrid* approach combines the previous two methods. In each step the oldest values are updated with the newest prediction values and a new training is completed – resulting in N different models and weights for the time units to be forecasted.

$$pred_{t+1} = model_1(obs_t, obs_{t-1}..., obs_{t-w+1}), \tag{7}$$

$$pred_{t+2} = model_2(pred_{t+1}, obs_t..., obs_{t-w+2}). \tag{8}$$

3.4. Evaluation

Exponential smmoting, Holt-Winters and LSTM models were developed for one step ahead prediction and LSTM models with the aforementioned approaches were implemented for ten step ahead prediction. The accuracy of the models was measured with three various error measures, namely with mean absolute percentage error (MAPE), R-squared (R^2) and normalized root mean squared error (NRMSE) – using the minimal and maximal WIP values for the normalisation. The models were developed with *keras* and *tensorflow* in *R Studio* [31].

4. Use case results

Figure 1 depicts the WIP of the investigated bottleneck machine for the analyzed period. No typical component of a time series can be clearly observed (trend or seasonality), implying unsatisfactory accuracy results of classical time series approaches. The first 70% of the data was used for training, while the last 30% for test purposes.

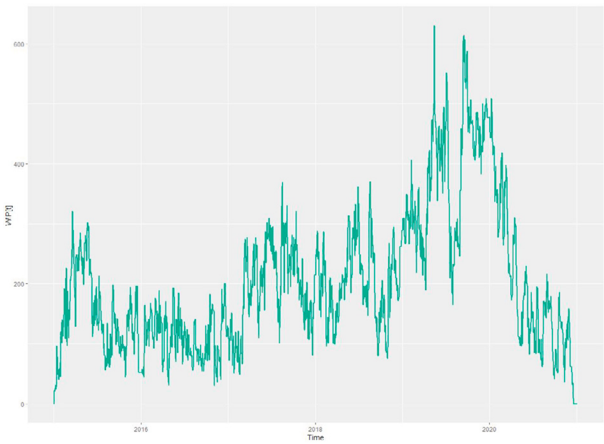

Fig. 1. WIP levels in the analysed time period

In case of an LSTM after data splitting, normalisation and formatting into the array format needed, the model architecture must be defined. The number of neurons with varying model size applied in the model tuning is summarized in Table 1.

4.1. One step ahead prediction

The average error measures of five runs with a medium size model with 2 hidden layers for different time window values were calculated. The batch size (32), the optimizer (adam) and the objective function (MSE) were kept fixed. The highest accuracy was achieved with time window 15. After fixing the time window a model tuning was executed. Based on the model capacity results medium or large models with one hidden layer or medium size 3 hidden layer models seemed promising. Batch normalisation did not bring any improvement in the mean absolute error of the validation data set. A widespread regularization method, i.e. dropout, increased the prediction accuracy of the model. With learning rate adjustment no significant advancement was achieved. After the identification of the most promising model tuning possibilities, a grid search was done for one hidden layer model with medium or large node size. With the help of this systematical search the optimal model architecture – with node size 64 and 256; and with dropout rate 0.25 – could be proposed. After fixing the LSTM model structure the prediction was done with two classical univariate methods as well, and the results are summarized in Table 2.

Table 1. Model capacities assessed represented as number nodes per layer

	Number of hidden layers		
Size	1	2	3
Small	16	16, 8	16, 8, 4
Medium	64	64, 32	64, 32, 16
Large	256	256, 128	256, 128, 64

Table 2. One step ahead prediction results on the test data set with the analysed methods

	Exponential smoothing	Holt-Winters	LSTM
MAPE	15.1	22.7	9.8
R^2	0.95	0.93	0.96
NRMSE	4.3	5.5	18.9

4.2. Ten step ahead prediction

As two out of the three applied error measures showed higher prediction accuracy with LSTM, just this one was applied for more step ahead prediction. For ten step ahead prediction all four approaches were developed and tested with a window size 2 and 15. The results are summarized in Tables 3 and 4. Figure 2 depicts the real and predicted values for the test set with the multiple output univariate LSTM approach – with time window 15.

Based on the accuracy results on the test part of the given data set it could be concluded that the most promising approach is the multiple output approach. In case of time window 15 higher prediction accuracy was achieved revealing the importance of including more historical data during the training phase. However, test runs with higher time window value showed lower accuracy and pointed out the optimal length of historical data.

Comparing the error measures of the multiple output model with the other ones it can be stated that the development of the accuracy results has a revised tendency. The multiple output approach is the only one that is able to consider each data entry regarding all the different step ahead prediction values and therefore the training of the network is optimized for the whole

Table 3. Accuracy results on the test data set with time window 2

	Multiple output			Recursive			Direct multistep			Direct recursive hybrid		
	MAPE	R^2	NRMSE	MAPE	R^2	NRMSE	MAPE	R^2	NRMSE	MAPE	R^2	NRMSE
1 step ahead	12.0	0.97	18.1	11.8	0.96	18.8	10.6	0.96	19.9	10.1	0.97	18.3
2 steps ahead	13.8	0.94	25.0	15.3	0.93	27.0	14.4	0.92	28.3	13.9	0.93	26.6
3 steps ahead	15.9	0.90	32.0	18.8	0.87	35.8	17.7	0.85	38.5	16.5	0.90	31.4
4 steps ahead	17.7	0.86	37.9	21.5	0.81	43.9	22.3	0.69	55.8	18.4	0.87	35.9
5 steps ahead	19.3	0.81	43.7	23.9	0.73	51.9	20.8	0.77	48.1	20.1	0.84	39.6
6 steps ahead	20.9	0.75	50.0	25.9	0.64	59.9	22.3	0.69	55.2	22.7	0.82	42.3
7 steps ahead	21.9	0.70	54.3	28.0	0.54	68.1	27.8	0.40	77.3	22.6	0.81	44.0
8 steps ahead	23.3	0.62	61.6	30.1	0.41	76.7	26.1	0.46	73.6	34.2	-0.74	131.8
9 steps ahead	24.2	0.59	64.2	32.0	0.26	85.7	25.4	0.51	69.9	34.9	-0.72	131.0
10 steps ahead	25.2	0.51	69.8	34.0	0.09	95.4	26.2	0.42	75.9	42.1	-2.66	191.2

Table 4. Accuracy results on the test data set with time window 15

	Multiple output			Recursive			Direct multistep			Direct recursive hybrid		
	MAPE	R^2	NRMSE	MAPE	R^2	NRMSE	MAPE	R^2	NRMSE	MAPE	R^2	NRMSE
1 step ahead	22.1	0.70	54.5	11.6	0.97	18.4	10.0	0.97	17.1	9.3	0.97	17.0
2 steps ahead	21.8	0.71	53.4	16.4	0.90	31.7	17.3	0.85	39.0	16.8	0.87	35.3
3 steps ahead	20.9	0.73	51.9	20.5	0.79	45.2	19.9	0.74	50.6	20.6	0.77	47.6
4 steps ahead	20.5	0.72	52.2	24.0	0.64	59.7	22.7	0.66	58.3	20.9	0.78	46.4
5 steps ahead	19.9	0.72	51.9	27.1	0.43	75.1	21.9	0.71	54.0	22.5	0.70	54.5
6 steps ahead	19.8	0.70	54.7	29.9	0.15	91.9	26.6	0.29	83.9	22.8	0.67	57.2
7 steps ahead	19.9	0.71	53.0	32.6	-0.22	110.4	24.2	0.58	64.5	26.4	0.36	79.7
8 steps ahead	18.8	0.69	54.9	35.1	-0.73	131.3	26.9	0.28	84.8	28.0	0.26	86.0
9 steps ahead	18.9	0.67	57.1	37.4	-1.40	154.6	28.9	-0.00	100.1	32.0	-0.23	110.8
10 steps ahead	18.2	0.69	54.9	39.7	-2.28	180.9	30.2	-0.16	107.5	35.0	-0.64	127.9

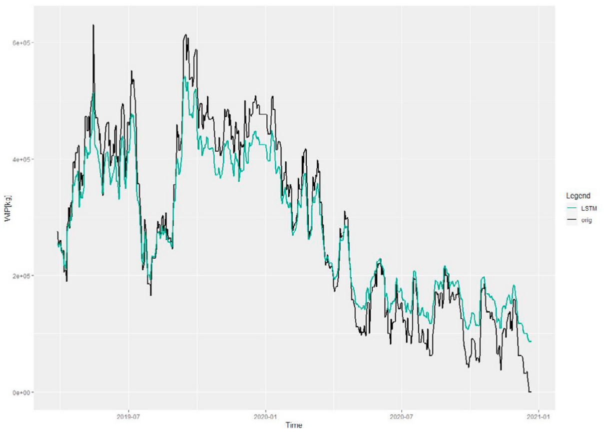

Fig. 2. Ten step ahead prediction with multiple output LSTM - comparison of real (black) and predicted (green) values of the test dataset

time frame. This results in poorer accuracy results on short term and better performance on long term.

5. Future research agenda

In this paper, predicting the WIP levels with classical and LTSM approaches has been presented. Univariate one to ten step ahead LSTM prediction models were developed in order to forecast the WIP for a given bottleneck machine. In the analysed industrial use case the highest prediction accuracy was achieved with the multiple output univariate LSTM model. Applying LSTM for predicting important production control KPIs (WIP, lead time or waiting time) might be interesting for practitioners without deeper knowledge about neural networks. Our intention was to present LSTM and demonstrate the various accuracy results for more step ahead predictions with different approaches. Our results might motivate practitioners to test this method in their own production environment.

There are several ideas how to widen the scope of this research. Firstly, the possibilities of a rolling prediction with LSTM will be addressed. Secondly, the focus of the analysis will be extended to other machines or processes as well. Thirdly, multivariate LSTM models could be used. An alternative approach might be the prediction of lead times. Last but not least, the prediction accuracy might be increased if the forecast is done not for all of the products in general but for given product types or customers in various steps.

Acknowledgements

We extend our sincere thanks to the Austrian Research Promotion Agency (FFG) for supporting the research project MLinPPC under the grant agreement No.877446. The authors would also like to acknowledge the financial support of the European Commission through the H2020 project EPIC (https://www.centre-epic.eu/) under grant No. 739592.

Appendix A. Open science

All readers that are interested in the R-Studio scripts are referred to the open source repository WIPwithLSTM [32]. This repository contains the developed models for one and *n* steps ahead predictions – with all 4 approaches.

References

[1] D. Spath, E. Westkämper, and H.-J. Bullinger, H.-J.and Warnecke, *Neue Entwicklungen in der Unternehmensorganisation*. Springer, 2017.

[2] H. ElMaraghy, G. Schuh, W. ElMaraghy, F. Piller, P. Schönsleben, M. Tseng, and A. Bernard, "Product variety management," *CIRP Annals*, vol. 62, no. 2, pp. 629–652, 2013.

[3] H. Lödding, *Verfahren der Fertigungssteuerung: Grundlagen, Beschreibung, Konfiguration*, 3rd ed., ser. VDI-Buch. Berlin and Heidelberg: Springer Vieweg, 2016. [Online]. Available: http://dx.doi.org/10.1007/978-3-662-48459-3

[4] H.-P. Wiendahl and H.-H. Wiendahl, *Betriebsorganisation für Ingenieure*, 9th ed. Hanser Fachbuchverlag, 2019.

[5] U. Dombrowski and Y. Dix, "An analysis of the impact of industrie 4.0 on production planning and control," in *Advances in Production Management Systems. Smart Manufacturing for Industry 4.0*, ser. IFIP Advances in Information and Communication Technology, I. Moon, G. M. Lee, J. Park, D. Kiritsis, and G. von Cieminski, Eds. Cham: Springer International Publishing, 2018, vol. 536, pp. 114–121.

[6] N. Duffie, J. Bendul, and M. Knollmann, "An analytical approach to improving due-date and lead-time dynamics in production systems," *Journal of Manufacturing Systems*, vol. 45, pp. 273–285, 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0278612517301358

[7] G. Schuh and C. Fuß, "Ergebnisbericht des bmf-verbundprojektes prosense. hochauflösende produktionssteuerung auf basis kybernetischer unterstützungssysteme und intelligenter sensorik." Aachen, Germany, Tech. Rep., 2015.

[8] L. Monostori, B. Kádár, T. Bauernhansl, S. Kondoh, S. Kumara, G. Reinhart, O. Sauer, G. Schuh, W. Sihn, and K. Ueda, "Cyber-physical systems in manufacturing," *CIRP Annals*, vol. 65, no. 2, pp. 621–641, 2016.

[9] A. Bueno, M. Godinho Filho, and A. G. Frank, "Smart production planning and control in the industry 4.0 context: A systematic literature review," *Computers & Industrial Engineering*, vol. 149, 2020.

[10] J. D. Little, "Or forum-–little's law as viewed on its 50th anniversary," *Operations research*, vol. 59, no. 3, pp. 536–549, 2011.

[11] A. Pfeiffer, D. Gyulai, B. Kádár, and L. Monostori, "Manufacturing Lead Time Estimation with the Combination of Simulation and Statistical Learning Methods," *Procedia CIRP*, vol. 41, pp. 75–80, 2016. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S2212827115010975

[12] A. Öztürk, K. Sinan, and N. E. Özdemirel, "Manufacturing lead time estimation using data mining," *European Journal of Operational Research*, vol. 173, no. 2, pp. 683–700, 2006.

[13] A. Alenezi, S. A. Moses, and T. B. Trafalis, "Real-time prediction of order flowtimes using support vector regression." *Computers & Operations Research*, vol. 35, no. 11, pp. 3489–3503, 2008.

[14] C. Wang and P. Jiang, "Deep neural networks based order completion time prediction by using real-time job shop rfid data," *Journal of Intelligent Manufacturing*, vol. 30, no. 4, pp. 1–16, 2019.

[15] I. Sabuncuoglu and A. Comlekci, "Operation-based flowtime estimation in a dynamic job shop," *Omega*, vol. 30, no. 6, pp. 423–442, 2002.

[16] G. Schuh, A. Gützlaff, F. Sauermann, and T. Theunissen, "Application of time series data mining for the prediction of transition times in production," *Procedia CIRP*, vol. 93, pp. 897–902, 2020. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S2212827120306272

[17] L. Lingitz, V. Gallina, F. Ansari, D. Gyulai, A. Pfeiffer, W. Sihn, and L. Monostori, "Lead time prediction using machine learning algorithms: A case study by a semiconductor manufacturer," *Procedia CIRP*, vol. 72, pp. 1051–1056, 2018. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S2212827118303056

[18] F. Misrudin and L. C. Foong, "Digitalization in Semiconductor Manufacturing- Simulation Forecaster Approach in Managing Manufacturing Line Performance," *Procedia Manufacturing*, vol. 38, pp. 1330–1337, 2019. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S2351978920301578

[19] P. Vazan, D. Janikova, P. Tanuska, M. Kebisek, and Z. Cervenanska, "Using data mining methods for manufacturing process control," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 6178–6183, Jul. 2017. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S240589631731460X

[20] M. Subramaniyan, A. Skoogh, H. Salomonsson, P. Bangalore, and J. Bokrantz, "A data-driven algorithm to predict throughput bottlenecks in a production system based on active periods of the machines," *Computers & Industrial Engineering*, vol. 125, pp. 533–544, Nov. 2018. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0360835218301608

[21] A. C. Choueiri, D. M. V. Sato, E. E. Scalabrin, and E. A. P. Santos, "An extended model for remaining time prediction in manufacturing systems using process mining," *Journal of Manufacturing Systems*, vol. 56, pp. 188–201, Jul. 2020. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0278612520300935

[22] W. Scholl, B. P. Gan, P. Lendermann, D. Noack, O. Rose, P. Preuss, and F. S. Pappert, "Implementation of a simulation-based short-term lot arrival forecast in a mature 200mm semiconductor FAB," in *Proceedings of the 2011 Winter Simulation Conference (WSC)*. Phoenix, AZ, USA: IEEE, Dec. 2011, pp. 1927–1938. [Online]. Available: http://ieeexplore.ieee.org/document/6147907/

[23] M. Mosinski, D. Noack, F. S. Pappert, O. Rose, and W. Scholl, "Cluster based analytical method for the lot delivery forecast in semiconductor fab with wide product range," in *Proceedings of the 2011 Winter Simulation Conference (WSC)*. Phoenix, AZ, USA: IEEE, Dec. 2011, pp. 1829–1839. [Online]. Available: http://ieeexplore.ieee.org/document/6147897/

[24] T. Chian Tin, K. Leng Chiew, S. Chee Phang, S. Nah Sze, and P. San Tan, "Incoming Work-In-Progress Prediction in Semiconductor Fabrication Foundry Using Long Short-Term Memory," *Computational Intelligence and Neuroscience*, 2019.

[25] R. Wirth and J. Hipp, "Crisp-dm: Towards a standard process model for data mining," *Computer Science*, p. 11, 2000.

[26] R. Hyndman, A. Koehler, J. Ord, and R. Snyder, *Forecasting with Exponential Smoothing: The State Space Approach*, ser. Springer Series in Statistics. Springer Berlin Heidelberg, 2008. [Online]. Available: https://books.google.at/books?id=GSyzox8Lu9YC

[27] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[28] R. Zhao, J. Wang, R. Yan, and K. Mao, "Machine health monitoring with LSTM networks," *Proceedings of 2016 10th International Conference on Sensing Technology (ICST)*, no. Nanjing, China, pp. 1–6, 2016.

[29] J. Kumar, R. Goomer, and A. K. Singh, "Long short term memory recurrent neural network (lstm-rnn) based workload forecasting model for cloud datacenters," *Procedia Computer Science*, vol. 125, pp. 676–682, 2018. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S1877050917328557

[30] B. Boehmke and B. Greenwell, *Hands-On Machine Learning with R*. Taylor & Francis Group, 2020. [Online]. Available: https://bradleyboehmke.github.io/HOML/index.html

[31] RCoreTeam, "R: A language and environment for statistical computing," 2017, visited on 01.03.2018. [Online]. Available: https://www.R-project.org

[32] V. Gallina, "Predicting the WIP level with long short-term memory recurrent neural network," 2021. [Online]. Available: https://github.com/ViolaGallina/WIPwithLSTM