Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 55-19 (2022) 31–36

Integrating PHM into production scheduling through a Digital Twin-based framework

Elisa Negri*, Laura Cattaneo, Vibhor Pandhare***, Marco Macchi*, Jay Lee*****

**Department of Management, Politecnico di Milano, Economics and Industrial Engineering, Milan, Italy (Tel: +390223993969; e-mail: elisa.negri@polimi.it; marco.macchi@polimi.it)*

*** School of Industrial Engineering, Università Carlo Cattaneo - LIUC, C.so Matteotti 22, 21053 Castellanza, Italy (lcattaneo@liuc.it)*

**** Mechanical and Materials Engineering, College of Engineering and Applied Science, University of Cincinnati, Cincinnati, United States (e-mail: pandhavr@mail.uc.edu; lj2@ucmail.uc.edu)*

Abstract: Production scheduling has a long history of research but still presents open challenges, when considering production systems with uncertainty. The digitization, and in particular Digital Twins, may play a role in progressing the research field. The paper proposes a framework to exploit the Digital Twin synchronization with the field to include health assessment models into the simulation-based optimization of the production system scheduling. The health assessment is based on a modelling of failure modes and monitoring signals to connect the physical production resources to health and operating states in order to have a more accurate prediction of the makespan with respect to the actual makespan of the production system. The scheduling framework is validated through a laboratory application in the Industry 4.0 Lab at the School of Management of Politecnico di Milano.

Keywords: Digital Twin, Equipment Health, Fault Detection and Diagnosis, Production Scheduling, Prognostics and Health Management PHM.

1. INTRODUCTION AND MOTIVATION

Production scheduling is one of the most important activities for manufacturing operations. This is reflected also in the importance that this topic covers in industrial engineering research since long time. Production scheduling is "the allocation of shared resources over time to competing activities" (Hopp and Spearman, 1996) with the aim of optimizing certain criteria, such as times and costs, by deciding the sequence of orders to be produced, when to produce them and with which resources (Baker, 1974). It is clear that a certain optimal scheduling solution may change as a result of disruptions and perturbations that occur in the real world, many of which are generated by the health status of the production resources. For this reason, scheduling approaches started to embed stochasticity and have been developing in the two main streams of robust and reactive scheduling (Aytug *et al.*, 2005; Duenas and Petrovic, 2008). Digitalization and Industry 4.0 technologies open new ways to take decisions and elaborate solutions: production scheduling may also be impacted. In fact, robust production scheduling has been traditionally studied with simulation-based optimization approaches. With the Fourth Industrial Revolution, simulation becomes field-synchronized and connected two-ways in real-time with the physical production resources. This new simulation concept is known as Digital Twin (Negri, Fumagalli and Macchi, 2017; Kritzinger *et al.*, 2018). The real-time capabilities and the simulation potentials of the Digital Twins enable a new scheduling approach that brings together robustness and reactivity to disruptions and perturbations. This new approach manages to react to physical production resources perturbations in real time, but at the same time it bases the optimization on a prediction of uncertainty and on quasi-real time simulations of scenarios, to make results more robust. This research idea aims at integrating Prognostics and Health Management (PHM) of the production resources into the production scheduling optimization based on Digital Twins. This paper proposes a progress with respect to the preliminary results of this research collaboration (Negri *et al.*, 2019, 2021). Some background work is presented in Section 2, gaps and objectives to be pursued in this paper are in Section 3, in which it will be clear how the paper is a progress with respect to the previous works; in Section 4 the proposed framework and its novelties are presented; in Section 5 the experimental setup and the first results are shown; conclusions are reported in Section 6.

2. BACKGROUND ON SIMULATION-BASED SCHEDULING

Most of the production scheduling problems have been long recognized as NP-hard (Neufeld, Gupta and Buscher, 2016) and the effort to compute an optimal solution may increase even more than exponentially with the size of the problem. Research on production scheduling proposed to use meta-heuristics algorithms that search and find sufficiently good solutions, that do not necessarily correspond to the global optima, but that have the advantage to be found in a reasonable computational time (Osman and Laporte, 1996). Among these meta-heuristics, one of the most used for production scheduling is the Genetic Algorithm (GA) (Ruiz and Maroto, 2005; Zandieh, Mozaffari and Gholami, 2010). Typically, the optimization algorithms are used together with simulation, in

Peer review under responsibility of International Federation of Automatic Control.
10.1016/j.ifacol.2022.09.180

a so-called "simheuristics" approach to production scheduling (Juan *et al.*, 2015; Gonzalez-Neira *et al.*, 2017; Hatami *et al.*, 2018). In particular, the role of the metaheuristics is to search for alternative solutions, that are then evaluated within a simulation model, that represents the dynamics and the interactions in time of the production resources, evaluates the effects of modifications on the production resources and schedules, before implementing them in the physical world, and embeds uncertainty prediction through stochastic modelling. The outputs of the simulations are sent back to the metaheuristic algorithm that is then able to compute the fitness function based on them. This work uses a simheuristic approach to production scheduling, based on GA and Digital Twin and integrating the PHM of the production resources. Robustness in production scheduling is defined as the capability of finding solutions that remain (close to the) optimum also in presence of small perturbances or disruptions of the production resources (Vieira *et al.*, 2017). Robustness is achieved by the simheuristics approach. The reactive approach to uncertainty instead is achieved through the real-time monitoring and elaboration of field data, obtained through the synchronization of the Digital Twin, and through the possibility to react in very short time to perturbances and disruption proposing a new scheduling solution to be implemented in the production system (Zhang *et al.*, 2021; Zhang, Tao and Nee, 2021). A deep review of the previous works on metaheuristics-based scheduling optimization and robust production scheduling may be found in (Negri *et al.*, 2021). In particular, it highlighted the role of simulation to model uncertainty to develop robust approaches to scheduling, but it also pointed out how the paradigm encountered in simulation-based optimization for scheduling problems has been rarely coupled with Digital Twins. Also, the research works found in literature investigated how the uncertainty impacted the search of the optimum solution by the optimization algorithm and the research progress in robust scheduling was not based on becoming more "reactive". Finally, although the research works deal with production scheduling, the vast majority investigated and validated the scheduling techniques in not much "realistic" systems. Most of the validations were done neither in production systems as real setting, nor in laboratories resembling real-life conditions. Consequently, field-synchronization to collect shop floor data was not proposed in literature works.

3. ADDRESSED GAPS AND PROPOSED CONTRIBUTIONS

The current work is framed within a wider investigation on the integration of PHM with production scheduling based on Digital Twins. The present paper aims at addressing the first progresses with respect to the mentioned paper (Negri *et al.*, 2021), contributing to answer the two gaps reported below. In particular, the objective of the paper is to demonstrate a robust production (re)scheduling framework, which leverages on the Digital Twin simulation-based metaheuristics optimization empowered with a PHM model describing the health conditions of the production resources. The gaps found in literature and the relative objectives that this work aims to address are the following:

GAP 1: The works proposed in literature deal with simulation-based scheduling optimization, with traditional simulation: Digital Twins are not fully exploited as synchronized simulations to describe the real-time conditions of the production asset/system that is scheduled.
OBJECTIVE 1: The proposed framework integrates the Digital Twin simulation providing real-time information for the degradation and failure statistical models for a fully dynamic rescheduling.

GAP 2: The health model of production resources within simulations for scheduling is typically framing a healthy state (i.e. resource able to operate) and unhealthy state (i.e. resource not able to operate, in a breakdown or under maintenance) (Von Hoyningen-Huene and Kiesmüller, 2015; Wu, Sun and Xiao, 2018).
OBJECTIVE 2: This work provides a more complex and realistic health model, that additionally includes a degraded state, meaning that the resource is able to operate at a slower pace or lower efficiency.

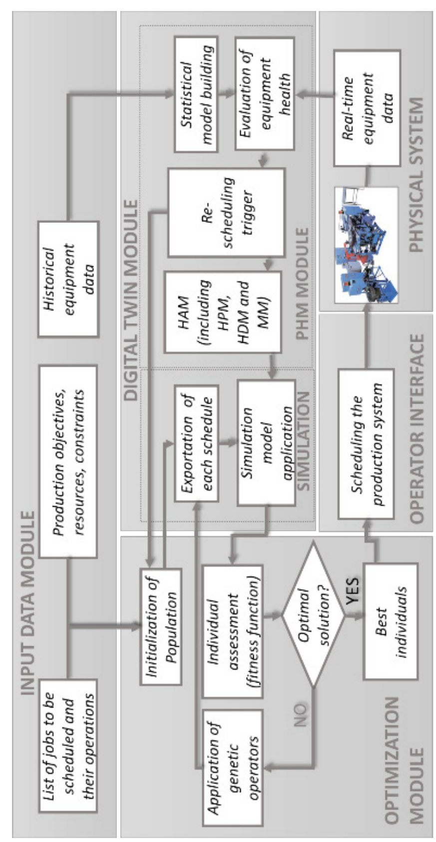

Figure 1. Proposed Digital Twin-based framework

4. PROPOSED FRAMEWORK

Figure 1 presents the proposed scheduling framework. It is composed of various modules: one collecting input data, one performing the optimization, one with the simulation and the resource-synchronized PHM that together compose the Digital Twin. The framework also includes a human-machine interface and the physical production resources of the system. The present article presents a progress with respect to the previous work, where it is possible to find a detailed description of each module (Negri *et al.*, 2021). In fact it improves the scheduling optimization by leveraging on a more precise real time estimation of production asset/system health, thanks to the Digital Twin capabilities.

4.1 Prognostics and Health Management module

The modelling of the states of the production resources includes three possible health states and three possible operating states, as reported in Figure 2. The health states are: *(i) Healthy*, the production resource is able to work at nominal efficiency (e.g. the processing times are standard); (ii) *Degraded*, the production resource is able to work at lower efficiency; (iii) *Failed*, the production resource is not able to work. The three operating states are: 1) *In Operation*, the production resource is currently working; 2) *Idle*, the production resource is able to work but not required to at the moment; 3) *Under Maintenance*, the production resource is not able to work and currently undergoing maintenance actions.

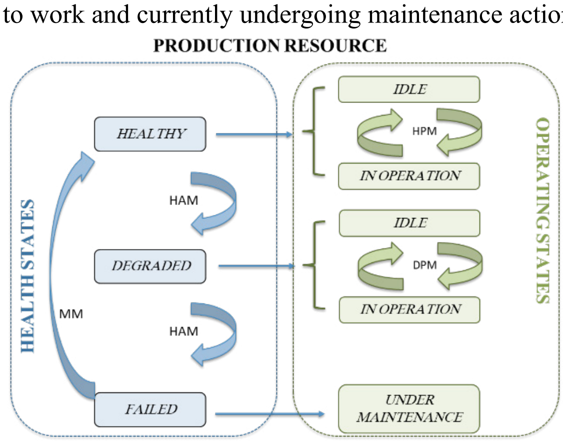

Figure 2. Production Resource State Map

Since the health and operating states affect the productivity of the production system, also the production capacity and the schedule are impacted. They are carefully modelled in the Digital Twin simulation with the synchronization of field data, in this way the evaluation of alternative schedules results more robust. To this end, the Digital Twin module embeds several mathematical models to describe the transition from one production resource state to another one, which represents one of the contributions of this work.

In particular, the Health Assessment Model (HAM) characterizes the transitions from a health state to another one (from healthy, to degraded, to failed, to healthy again, as expressed in Figure 2). The HAM outputs the time the production resources spend in each state, through the use of three models:

a) the Healthy Production Model (HPM) identifies when the resource is in healthy state and if it is in idle or operating state;
b) the Degraded Production Model (DPM) identifies when the resource is in degraded state and if it is in idle or operating state;
c) the Maintenance Model (MM) identifies when the resource is back to healthy state.

As in Figure 3, the HAM assesses the Conformance Value (CV) of a production resource through monitoring signals of the physical resource, which come from synchronized controllers, accelerometers and other sensors of the system following a series of analyses that includes data preprocessing, model training and model testing. Based on the evaluated CV, the Digital Twin will use the processing time of the resource computed through the HPM or the DPM. In case the CV indicates the resource is failed, the repair time is computed through the MM. The HAM considers several possible degradation modes at once, as a progress with respect to the previous mentioned work in which only one failure mode was considered. Since the authors assumed normal processing times, the parameters of the normal distribution should be adapted to the health and operating state in which the production resources are currently in and must be estimated based on field data. Field data are segmented through pre-processing methods into two broad categories of features: time-based features and condition-based features. In particular, during healthy and degraded operating states, time-based features are collected to build the HPM and DPM. The condition-based features are collected during all resources states and are used to build the HAM.

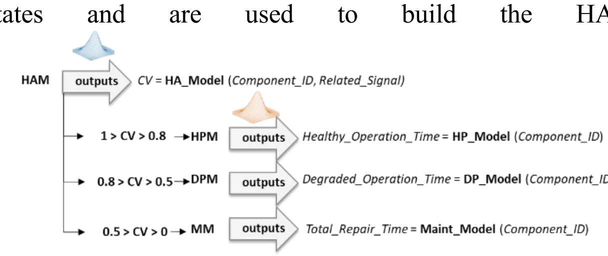

Figure 3. Components of the HAM

4.2 The role of Simulation Model and Genetic Algorithm

The Simulation model is a Discrete Event Simulation model which receives the normally distributed processing times and repair times, with the parameters computed starting from the synchronized field data by the HAM. The uncertainty is then treated by repeating simulation runs various times in a Monte Carlo approach, in order to evaluate schedules and to find a robust solution. When scheduling is triggered, the processing times distributions are input in real-time into the simulation model, achieving the field-synchronization of the simulation model with the current operating and health conditions of the production resources. The other input into the simulation model is an alternative production schedule generated by the GA as one individual of a population. The simulation model simulates the sequence of jobs entering the system and evaluates the makespan of the alternative schedule, repeating several runs in order to get an average and standard deviation of makespan to be compared with the ones of the other alternative schedules of the population. The GA will then evaluate whether the best individual of the current population is the desired “sufficiently good” schedule or if it is necessary to proceed with the generation of a new population of alternatives. This approach has been explained in (Fumagalli *et al.*, 2019).

5. APPLICATION CASE

5.1 Laboratory and problem description

The framework has been applied to a laboratory assembly line installed at the Industry 4.0 Lab at the School of Management of Politecnico di Milano (Luca Fumagalli et al. 2016). The line assembles a simplified smartphone, in seven workstations. The validation of the framework is more realistic with respect to the previous work, as it encompasses more than one degradation modes at a time and also considers more than one critical workstations. In particular, the most critical workstations, whose degradation modes are considered in the HAM, are the drilling station (where holes are drilled on the front cover) and the back cover station (where the back cover of the phone is positioned over the other assembled components). Three possible degradation modes in the drill station are considered in the experiments, as follows: (i) Drilling tool wear; (ii) Lubrication wear on handling axis; and (iii) Track wear on handling axis. On the back cover station the degradation mode consisted of wear on the axis used in locking of the cover mechanism. In order to have more frequent conditions recognized as failures or degraded health states, manipulations during the operations are artificially induced by a vibration shaker on the drilling axis, by manipulating the pressure valve of the handling axis and by manipulating the valve on the cover mechanisms. Monitored signals received in real-time from the field (i.e. from the PLC of the two critical workstations through the OPC-UA architecture) are electricity, power, air flow rate and air pressure. The scheduling problem formulated for this validation aims at minimizing the total makespan to produce eight jobs in the assembly line, each of which is different from the others in terms of variations in the tasks to be done at the workstations of the line. The optimization module and the parameters of the GA did not change from the previous work (Negri *et al.*, 2021).

5.2 Processing of data and training of models

Samples are collected in healthy and non-healthy states. Preprocessing of data is done through a segmentation process which extracted 66 features (frequency domain features for signals from the accelerometer and time domain features for the rest of signals). Features are normalized and Principal Component Analysis (PCA) is implemented to observe clusters in the healthy, degraded and faulty states and PCA T^2 test is used to detect failures (Wise *et al.*, 1999). As a result of the data preprocessing and the training of the models, the different distributions of the processing times in the drill station and the back cover station related to the different degraded modes are clearly identified, see Figure 4.

5.3 Description of the experiment

The validation experiments are ten, distributed in two scenarios (see Table 1).

A) The first scenario corresponds to always-healthy production resources. Six experiments are conducted to compare the predicted with the actual makespan when the workstations of the assembly line are always healthy (no component degradation is present in these experiments).

B) The second scenario corresponds to the scenario where a production resource is operating in a degraded state. Considering the scope of this work, a fault is artificially induced (as explained in Section 5.1) only in the drill station for track wear on handling axis. Four experiments are repeated with the same degradation mode for the drill station for robust investigation. In all four experiments, the fault is induced in between a production order after three jobs have already been produced from the order. At this time, the re-scheduling is automatically triggered as a result of the real-time condition-monitoring by the Digital Twin and its HAM. The remaining sequence of five jobs is re-optimized in real-time taking into consideration the new conditions of the production system. A new predicted makespan is compared with the actual makespan obtained after the failure.

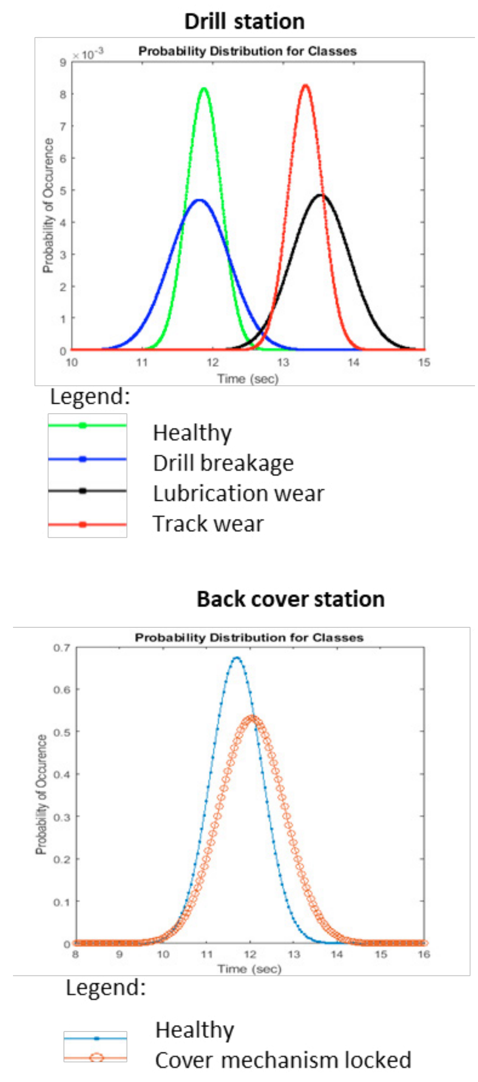

Figure 4. Processing times distributions

5.4 Experimental results

The results of the scenarios are represented in Table 1, in which it is clear how the production schedule follows the real-time conditions of the production system and the predicted and actual makespans are consistent. The predicted makespan is the one predicted by the optimization algorithm of the schedule by predicting the processing times of the production resources starting from the field-data taken in real time from the HAM and its sub-models. The actual makespan is the makespan of the actual physical assembly of the eight jobs, directly measured on the assembly line. It is worth noting that the second scenario has lower makespans compared with the first scenario, because the first scenario considers the whole production schedule made of eight jobs to be assembled. The second scenario considers the predicted and actual makespans of the remaining five jobs, that are re-scheduled (see description B of Section 5.3). The implemented optimal sequences are reported in the last column of Table 1. It is important to note that, even though a single degradation is induced in the second scenario, the HAM is receiving data from all the four signals from both the drill and back cover stations and is able to identify the correct degradation mode and correspondingly choose DPM. The accuracy of the HAM is reported in Fig. 5. The confusion matrix of the HAM here reported demonstrates the capability of detecting the deviation in the monitoring signals, to diagnose the correct failure mode and to estimate the time worsening due to degradation of the production resource accordingly (Fig. 6), so estimating the new total makespan in the degraded scenario in real-time. As shown in the table, the predicted makespan lies always within an error percentage comprised between -3,5% and +3,5% of the actual makespan (where error percentage = [predicted makespan – actual makespan]/actual makespan).

Table 1. Experimental results

	#	Predicted Makespan [sec]	Actual Makespan [sec]	Error %	Sequence
1st SCENARIO	1	208	215	-3,3%	8-2-7-1-5-3-4-6
	2	206	207	-0,5%	1-8-7-2-5-3-4-6
	3	209	213	-1,9%	5-2-6-8-1-7-3-4
	4	207	210	-1,4%	2-8-7-1-5-3-4-6
	5	204	208	-1,9%	2-5-7-8-1-3-6-4
	6	210	212	-0,9%	8-3-2-7-5-1-4-6
2nd SCENARIO	7	177	183	-3,3%	1-5-3-4-5
	8	185	190	-2,6%	1-8-6-3-4
	9	183	181	1,1%	1-7-5-4-6
	10	186	181	2,8%	8-1-5-4-6

6. CONCLUSIONS

The paper proposes a progress with respect to a previous published work, in the research investigation that explores the new potentials coming from including a real-time field-synchronized Digital Twin into the simulation-based optimization for production scheduling for a convergence of the benefits of reactive scheduling into a robust scheduling approach. Doing this, the Digital Twin hosts a Health Assessment Model that elaborates the health states and corresponding Production Models that elaborate the processing times of production resources by gathering real-time data from the field, based on failure mode and signal monitoring. For this reason, the prediction of the makespan results in being realistic and connected to the conditions of the physical resources, that leads to three possible health states (healthy, degraded, and faulty) and to three possible operating states (in operation, idle, under maintenance). The paper reports the first experimental results of this framework, demonstrating the capability of the scheduling framework to detect monitoring signals deviations, diagnose the relative degradation mode, and estimate a consistent makespan, compared to the actual makespan of the assembly process, both in the healthy scenarios and in the degraded ones. The experiments have demonstrated that the framework works in a (quasi) real-time fashion, allowing to reschedule within seconds, providing a convincing near-optimum solution to the scheduling problem. This work has both industrial and research implications: representing a step forward to understanding how the Digital Twin may be practically used for industrial decision making. A limitation of this work is that during the re-scheduling activity triggered by a failure or degradation in the production, the resources keep working, so the conditions and the monitoring signals may change, especially for large production orders, making the new optimal schedule obsolete.

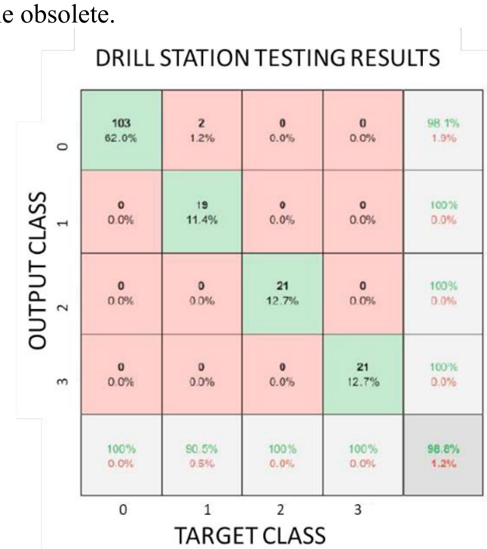

Figure 5. Confusion matrix of the HAM. Class 0 represents the healthy state, while classes from 1 to 3 represent the three degradation modes of the drill station.

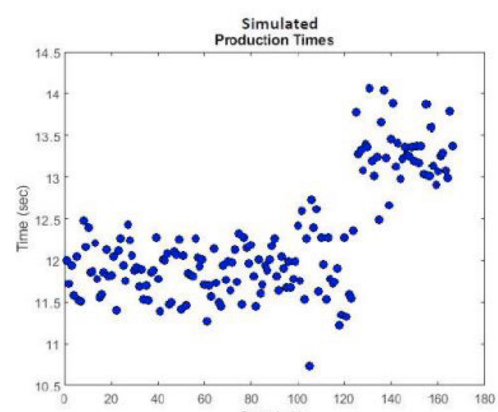

Figure 6. DPM as estimated from the HAM results for the drill station, track wear on handling axis.

Future works may go in the direction of solving this issue, by leveraging even more on the power of synchronization of the Digital Twins, by constantly monitoring and updating the HAM for the new schedule optimization, also during the scheduling activity itself. Moreover, future work will include more experiments to demonstrate the behavior of this scheduling framework under different scenarios, degraded and failed conditions. Last but not least, it is worth remarking the future role of the framework for decision-making support: the framework will guide the implementation of a supervisory control to dynamically adjust the asset performances and the production schedules. It will require both to embed prognostic capabilities on the PHM side, and a predictive workload on the production side. In its entirety, the predictive approach should effectively lead to optimize, as a joint practice, production, and maintenance objectives in the presence of disturbances and perturbations from the shop floor.

REFERENCES

Aytug, H., Lawley, M.A., McKay, K., Mohan, S., and Uszoy, R. (2005). 'Executing production schedules in the face of uncertainties: A review and some future directions', *European Journal of Operational Research*, 161(1), pp. 86–110.

Baker, K. (1974) *Introduction to sequencing and scheduling*. John Wiley & Sons.

Duenas, A. and Petrovic, D. (2008) 'An approach to predictive-reactive scheduling of parallel machines subject to disruptions', *Annals of Operations Research*, 159(1), pp. 65–82.

Fumagalli, L., Macchi, M., Pozzetti, A., Taisch, M., Tavola, G. and Terzi, S. (2016) 'New methodology for smart manufacturing research and education: The lab approach', *Proceedings of the Summer School Francesco Turco*, 13-15-Sept, pp. 42–47.

Fumagalli, L., Polenghi, A., Negri, E., Roda, I. (2019) 'Framework for simulation software selection', *Journal of Simulation*, 13(4), pp. 286–303.

Gonzalez-Neira, E. M., Ferone, D., Hatami, S., and Juan, A.A. (2017) 'A biased-randomized simheuristic for the distributed assembly permutation flowshop problem with stochastic processing times', *Simulation Modelling Practice and Theory*, 79, pp. 23–36.

Hatami, S., Calvet, L., Fernández-Viagas, V., Framiñán, J.M., and Juan, A.A. (2018) 'A simheuristic algorithm to set up starting times in the stochastic parallel flowshop problem', *Simulation Modelling Practice and Theory*, 86(April), pp. 55–71.

Hopp, W. J. and Spearman, M. L. (1996) *Factory Physics: foundations of manufacturing management*. Homewood, Illinois: Irwin.

Von Hoyningen-Huene, W. and Kiesmüller, G. P. (2015) 'Evaluation of the expected makespan of a set of non-resumable jobs on parallel machines with stochastic failures', *European Journal of Operational Research*, 240(2), pp. 439–446.

Juan, A. A., Faulin, J., Grasman, S.E., Rabe, M., Figueira, G. (2015) 'A review of simheuristics: Extending metaheuristics to deal with stochastic combinatorial optimization problems', *Operations Research Perspectives,* 2, pp. 62–72.

Kritzinger, W., Karner, M., Traar, G., Henjes, J., and Sihn, W. (2018) 'Digital Twin in manufacturing: A categorical literature review and classification', *IFAC-PapersOnLine*. Elsevier B.V., 51(11), pp. 1016–1022.

Negri, E., Ardakani, H.D., Cattaneo, L. Singh, J., Macchi, M., and Lee, J. (2019) 'A Digital Twin-based scheduling framework including Equipment Health Index and Genetic Algorithms', in *13th IFAC Workshop on Intelligent Manufacturing Systems*, pp. 43–48.

Negri, E., Pandhare, V., Cattaneo, L. Singh, J., Macchi, M., and Lee, J. (2021) 'Field-synchronized Digital Twin framework for production scheduling with uncertainty', *Journal of Intelligent Manufacturing*, 32, pp. 1207–1228.

Negri, E., Fumagalli, L. and Macchi, M. (2017) 'A review of the roles of Digital Twin in CPS-based production systems', *Procedia Manufacturing*, 11, pp. 939–948.

Neufeld, J. S., Gupta, J. N. D. and Buscher, U. (2016) 'A comprehensive review of flowshop group scheduling literature', *Computers and Operations Research*. Elsevier, 70, pp. 56–74.

Osman, I. H. and Laporte, G. (1996) 'Metaheuristics: A bibliography', *Annals of Operations Research*, 63, pp. 513–623.

Ruiz, R. and Maroto, C. (2005) 'A comprehensive review and evaluation of permutation flowshop heuristics', *European Journal of Operational Research*. Elsevier, 165(2), pp. 479–494.

Vieira, G. E., Kück, M., Frazzon, E. and Freitag, M. (2017) 'Evaluating the Robustness of Production Schedules using Discrete-Event Simulation', in *IFAC-PapersOnLine*, pp. 7953–7958.

Wise, B. M., Gallagher, N.B., Watts Butler, S., White D.D., and Barna, G.G. (1999) 'A comparison of principal component analysis, multiway principal component analysis, trilinear decomposition and parallel factor analysis for fault detection in a semiconductor etch process', *Journal of Chemometrics*, 13(3–4), pp. 379–396.

Wu, Z., Sun, S. and Xiao, S. (2018) 'Risk measure of job shop scheduling with random machine breakdowns', *Computers and Operations Research*, 99, pp. 1–12.

Zandieh, M., Mozaffari, E. and Gholami, M. (2010) 'A robust genetic algorithm for scheduling realistic hybrid flexible flow line problems', *Journal of Intelligent Manufacturing*, 21(6), pp. 731–743.

Zhang, J., Deng, T., Jiang, H., Chen, H., Qin, S., and Ding, G. (2021) 'Bi-level dynamic scheduling architecture based on service unit digital twin agents', *Journal of Manufacturing Systems*, 60(May), pp. 59–79.

Zhang, M., Tao, F. and Nee, A. Y. C. (2021) 'Digital Twin Enhanced Dynamic Job-Shop Scheduling', *Journal of Manufacturing Systems*, 58(PB), pp. 146–156.