Reducing mean tardiness in a flexible job shop containing AGVs with optimized combinations of sequencing and routing rules

Heger, Jens; Voß, Thomas

Published in:
Procedia CIRP

DOI:
10.1016/j.procir.2019.03.281

Publication date:
2019

Document Version
Publisher's PDF, also known as Version of record

Available online at www.sciencedirect.com

ScienceDirect

Procedia CIRP 81 (2019) 1136–1141

www.elsevier.com/locate/procedia

52nd CIRP Conference on Manufacturing Systems

Reducing mean tardiness in a flexible job shop containing AGVs with optimized combinations of sequencing and routing rules

Jens Heger[a], Thomas Voss[a]*

[a] *Leuphana University Lueneburg, Volgershall 1, 21339 Lueneburg*

* Corresponding author. *E-mail address:* thomas.voss@leuphana.de

Abstract

The complexity of flexible job shop problems increases significantly when using autonomous guided vehicles (AGVs) for material handling. In this study, priority rules - commonly known for their simplicity and small computation time - for sequencing operations, routing jobs and dispatching vehicles are applied. Based on a discrete event simulation study with stochastic inter-arrival times, an artificial neural network is trained to learn interaction effects between the combination of different rules for sequencing, dispatching, routing, and the resulting system performance. Based on the trained network the combination of rules is optimized, reducing the mean tardiness of the jobs under varying system performance.

Keywords: AGVs; Neural network; Regression; Job shop;

1. Introduction and problem description

To manufacture goods economically, every company has to act within certain restrictions where short lead times, low levels of storage, high utilization, and high on-time delivery stand in contrast to each other. Within the last decade, the increasing demand for individual products and the possibility to configure a product just the way the customers want it has led to a high number of variants. Enabled by new functions and features as well as different regional requirements and specifications, each customer wants a unique product that exactly suits his/her need [1]. This fact is leading to a batch size one production, based on unique customer experience. Due to uncertainty and changing customer needs and demands, static grouping and formation of machines in the production systems might no longer be suitable. Since rearranging machines has several disadvantages such as reconfiguration cost and loss of profit, interaction between the cells may enable alternative sequences, increasing the cost for material handling [2]. By now, large enterprises and manufacturers are rethinking their factory concepts, taking into account the uncertain future. This change from mass production to mass customization leads to manufacturing concepts such as flexible job shop systems (FJS). Multiple dissimilar machines with different capabilities are combined in one area. Still, all machines have the same requirements regarding tolerances or capacity. The advantages range from reduced material handling and setup times to increased utilization [3].

To keep the cost of material handling low and still provide a robust and dynamic supply to the shop floor, autonomous guided vehicles (AGVs) can be used. The AGVs are used not only for transferring material from one central warehouse location to all work centers but also between the workstations. Enabled by technology, all assets in the production environment are able to communicate with each other, becoming smart machines called cyber-physical systems (CPS). When machines, storages, AGVs, and workers are connected with each other, the system is able to share

10.1016/j.procir.2019.03.281

knowledge, trigger reactions, and control its assets autonomously. Such an environment leads to new chances and possibilities in the area of production planning and scheduling [4].

The interaction effects and the impact of flexibility on routing jobs through a manufacturing system, dispatching AGVs and sequencing operations in front of machines in flexible manufacturing systems has rarely been considered up to now. Furthermore, the usage of performance forecasts has rarely been considered with this combination of problems.

2. State of the art

Given the fact that scheduling most real-world job shop instances is usually a NP-hard problem [5], it cannot be solved optimally within reasonable time frames. Optimal solutions to problems with AGVs have been presented, mostly considering central approaches such as mathematical modeling or graph-based algorithms. Although, these approaches result in an optimal solution, they can only be applied to small scenarios [6]. For that reason, other methods need to be derived. Heuristic methods, which can find a near optimal solution in approximate real-time have to be developed. Kousi et al. [7] presented a solution where the plan generation is based on time and inventory level driven criteria. Bearing in mind that dynamic systems, characterized by continuously arriving new orders or unplanned interruptions, cannot be scheduled optimally after all, the idea of calculating a near optimal solution grows in importance. Dynamic aspects such as machine break-downs, change in priority and other uncertainties require a change of the calculated schedule. This forces former schedules to be recalculated. Even if an optimal schedule could be calculated for the scenario, the necessary recalculations lead to a non-optimal result in the end.

In contrast to central approaches, priority based rules can be used to assign tasks to members of the shop floor very fast. Instead of assigning each and every player an operation in advance, decentral approaches neglect information up and down the process chain and consider only local knowledge, which reduces computation times and enables quick reactions. Especially if different members on the shop floor interact with each other under dynamic circumstances, the different rules can influence the system performance drastically. To improve the response to these interactions and other dynamic aspects on the shop floor, a regression model will be generated to forecast system performance based on system state and utilization. To find a suitable model, multiple methods will be compared against each other.

Given the fact that the projected system performance will turn out with a forecast error, the heuristic solution assigning operations needs to be adjusted to the system performance. For all further experiments regarding the heuristic solution and the parameter studies, an evaluation model is needed; in this case a discrete event simulation (DES) model considering the machines, AGVs, the routing of jobs as well as the sequencing of operations.

2.1. Priority based rules for sequencing, routing and dispatching

Single attribute dispatching rules such as First In First Out (FIFO) or Shortest Processing Time (SPT) are commonly known in industrial applications due to fast and simple execution. Panwalkar et al. [8] conducted an extensive study on priority-based rules for sequencing orders in front of a machine. Given the fact, that single attribute rules tend to behave well only in very specific parameter combinations, the usage of composite dispatching rules, combining multiple terminals, has become more important.

Additionally priority rules can be used for routing jobs through a manufacturing system as well. When two or more machines are available, which could possibly handle the task, they compete for one operation, and hence a routing decision has to be made. The degree to what extent this is possible is called routing flexibility [9] and is one of the major contributors to complexity in job shop scheduling. Routing flexibility can lead to higher utilization and balance machine workload [10]. Kumar [11] analyzes the effect of different levels of routing flexibility using the rules presented. The conclusion states a significant decrease in make span due to the proper balancing of workload. In this work, three different routing rules are used. These commonly known routing rules are also listed in [12] and can be generalized into:

- Work In Next Queue (WINQ) - considering the smallest workload of jobs in queue, also called as Least Waiting Time (LWT)
- Number In Next Queue (NINQ) - considering the minimum number of jobs in queue, this can also be called Shortest Queue (SQ)
- Least Utilized Machine (LUM) - considering the machine with least utilization, leveling out the workload of all machines

Priority based rules are used not only for the sequencing and the routing but also for the dispatching of vehicles. Bilge et al. [13] conducted studies for dynamic weights multi-attribute dispatching rules based on system characteristics and showed that dynamic rules are very robust. Guan et al. [14] show that multi attribute dispatching rules can balance system load and make efficient use of system resources. In this study the rules: least utilized vehicles (LUV) and shortest travel time (STT) will be used for the dispatching of the vehicles.

In this contribution, a decentral priority rule based approach for autonomous guided vehicles used for material handling in a flexible job shop is evaluated. Single attribute priority based rules, as presented above, are dynamically combined with each other for routing, dispatching, sequencing under dynamic job arrival, and system utilization. Preliminary results show that a dynamic adjustment can improve the results (reduce the mean tardiness) significantly.

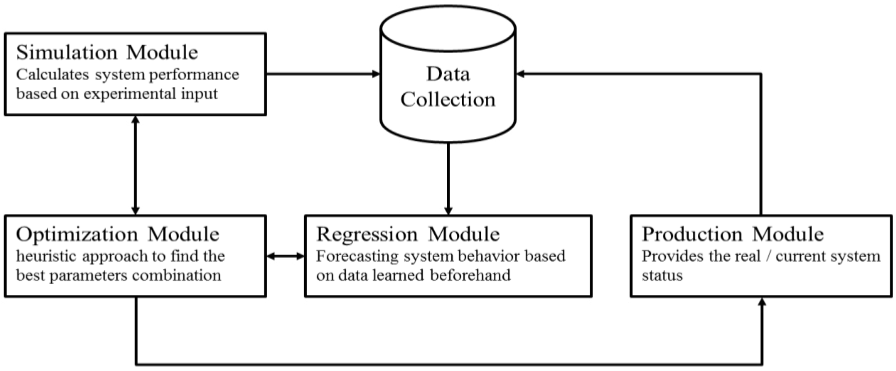

Fig. 1. The model shows the approach using four different modules and one central data collection unit. I should probably label the arrows and what information is flowing where.

2.2. Using regression to forecast system performance

Comparing the time to calculate a key performance indicator (KPI) such as tardiness based on a specific system load, discrete event simulation takes much longer than a regression model. Furthermore, not every combination can be calculated with a simulation model. Based on prior studies and the state of the art, a linear regression and a neural network are tested in this paper [15–18].

Knowing that the KPIs of a FJS are directly related with the utilization of the machines and the AGVs, a linear regression model is calculated to estimate the system performance. The linear regression can be calculated with a function $y = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k + \epsilon$ which calculates the value for all the input parameters x_k with the factor β_k, adding an error ϵ [19]. Linear regression works best if the estimated variable is continuous and steady. The moderation effect, representing the context factors such as the rule combination, has to be evaluated.

The usage of artificial neural networks (ANNs) is nowadays ubiquitous in various fields of science. Their main advantage lies in their ability to classify and evaluate data based on previously acquired data. Although most of ANNs' usage are found in image and speech recognition contexts, in this work they will be applied on simulated data in logistics [20]. Learning in this context refers to the ability of ANNs to derive and depict variations in the input data, and estimate an output for another given input based on the behavior of the original data. Depending on the structure of outer and hidden layers as well as the number of used neurons in the system the performance of the outcome can vary massively [21].

3. Proposed concept

In a previous paper [20], we have used a simulation study to calculate a certain number of simulation runs based on a versatile combination of parameters. The simulation model is used to generate a large data set, representing multiple system loads. The values in this set generate a database to train the regression model offline to forecast the system performance based on varying influences such as utilization. The trained model can be used to estimate the system performance according to used rule combination and machine- /AGV-utilization and forecast the behavior between the trained points. Fig. 1 shows the general concept. An optimization module will be used to find the best combination of rules for the given system parameters in interaction with either the simulation module and / or the forecast module. Even though part of the concept and presented in the figure, the optimization module is not part of this work.

4. Simulation model

In this section, the simulation model used for the experiments is described. The generic model presented in this work is related to [22] and has been published before. Modification have been made to be able to run a DES. The job types are set with a uniform distribution ranging between all four job types. The jobs are inserted into the system with a Poisson distributed inter arrival time. Tables 1 to 4 provide processing time, operation sequence, skills of each machine, and travel times for the scenario, respectively.

Table 1. Processing time for each product type based on the machine used. The processing of a type T1 product requires 60 time units on machine M1.

Type \ Machine	M1	M2	M3	M4	M5	M6
T1	60	60	120	20	20	140
T2	200	200	160	100	100	80
T3	180	180	220	100	100	80
T4	100	100	100	100	100	80

The required sequence of skills needed to complete the job for each of the four job types is presented in Table 2 below.

Table 2. Required skill sequence for each product type. The first operation of job type 1 requires skill 3, the second operations requires skill 1 and so on.

Type \ Skill needed	O1	O2	O3	O4
T1	3	1	2	4
T2	2	3	1	4
T3	1	4	3	2
T4	2	1	3	4

Since machines in a FJS can have different capabilities, the machines are assigned designated skill sets. In combination with Table 1 and Table 2 above, the idea of multi-purpose machines becomes more obvious and the skill set can be deducted. In Table 3, it can be seen that machine 1 (M1) can process the operations 2 and 3 of the type 1 product. This represents the skills 1 and 2. Whereas machine 4 can process operations number 1 and 2 resulting in the skills of 1 and 3. Due to the consideration of multi-purpose machines, where machines have unequal processing times, the impact of the routing rule in combination with the other rules cannot be assed prior to simulation.

It can be seen that the processing time is roughly 4 times as large as the transport time. This is called as the PT-ratio [23], which could be analyzed further in the future. For the processing time being lager than the transportation time it can be estimated that choosing a bad sequencing rule will have a larger impact than a bad dispatching rule.

Table 3. Processing times for each machine based on the skills needed for the operation.

Part Type	Operation	Processing Machines					
		M1	M2	M3	M4	M5	M6
1	O11				100	20	
	O12	60	60		100		
	O13	60		120			
	O14			120			140
2	O21	180		160			
	O22				100	100	
	O23	200	200		100		
	O24			160			140
3	O31	180	180		100		
	O32			220			80
	O33				100	100	
	O34	180		220			
4	O41	100		100			
	O42	100	100		100		
	O43				100	100	
	O44			100			100

The travel times are symmetric and all machines can be directly accessed with a straight line. Vehicles can move freely and are not bound to any loop or network.

Table 4. The transportation times between the machines. The time are symmetric and do not consider loops or road networks.

	M1	M2	M3	M4	M5	M6	IN	OUT
M1	0	27	21	17	27	26	26	27
M2	27	0	30	14	6	8	30	3
M3	21	30	0	19	27	22	25	29
M4	17	14	19	0	15	9	21	11
M5	27	6	27	15	0	9	30	8
M6	26	8	22	9	9	0	28	10
IN	26	30	25	21	30	28	0	29
OUT	27	3	29	11	8	10	29	0

In Fig. 2, the scenario is roughly sketched, showing the machines as well as the AGVs. The skills of each machine are represented with the small bubbles on every station.

5. Experiments

Two types of experiments have been conducted: First, the simulation experiments to generate the data to forecast the system performance based on the utilization and the given rule combination, and second, experiments regarding an adequate regression model.

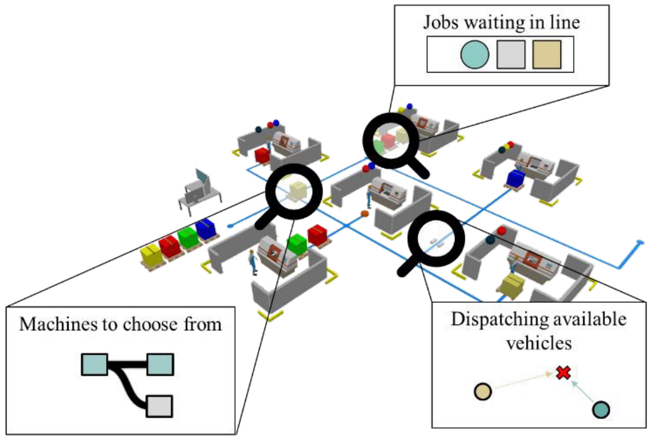

Fig. 2. The scenario and the aspects routing, sequencing and dispatching being considered.

5.1. Simulation experiments

To find a good combination of weights for the dispatching and sequencing rules, multiple scenarios have to be tested. For that reason the weights of the terminals, represented by continuous variables, have been tested as binary variables from zero to one. With 3 different rules for sequencing, 3 rule for routing and 3 rules for dispatching being considered, resulting in 27 possible combinations. Taking into consideration the different utilizations of the machines and the vehicles, 2 more variables have to be considered. The number of vehicles ranging from two to five to vary the AGV utilization. Furthermore, the inter arrival time ranging from 70 to 110 time units in steps of 10.

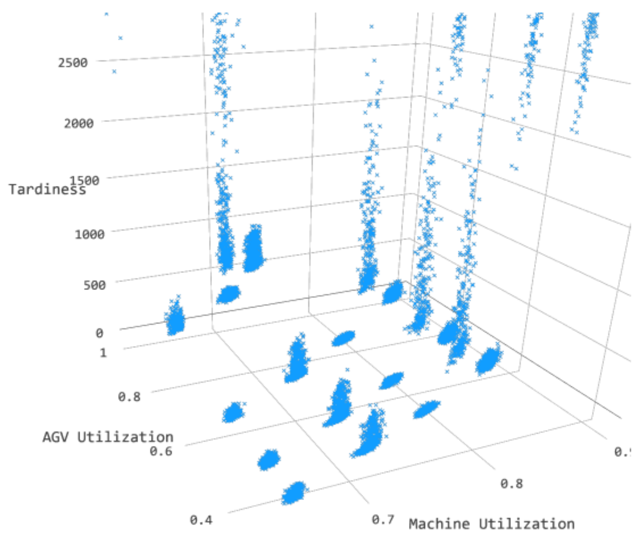

Fig. 3. Every marker represents a single simulation run. In this picture, 16929 markers are represented.

The different rule combinations and system load level result in a large dataset, which can be used to train the regression method. In Fig. 3, the single simulation runs are represented as scattered points. The simulation parameters are presented in Table 5.

5.2. Forecasting the tardiness

To be able to consider various system utilization levels, regression models have been generated, estimating the system performance based on utilization as well as the weights used in the combination.

Table 5. Simulation parameters, which have been used in the different runs.

Number of machines:	6
Number of AGVs:	2 - 5
Overall system utilization:	65 % to 99 %
Job types:	4 (uniform distr.)
Operation per job:	4
Inter-Arrival time:	Poisson distr.
Processing time:	Static
Due-date:	TWK method
Seq. rules:	FIFO, SPT, EDD
Disp. rules:	LIV, LUV, STT
Routing rules:	LUM, LWT, SQ
Warm up:	500 jobs
Run length:	2,000 jobs
Replications:	10
KPIs:	Mean Tardiness (MT)

Based on the simulation output, the training input data, containing 14 variables including 4 continuous variables, is preprocessed to ensure ideal performance for both the models. The continuous variables, AGV- and machine-utilization range from zero to one. The flow time as well as the tardiness are float numbers. The categorical variables have different categories so they are one-hot encoded. Before training the model, the input data is split into train and test data in the ratio of (56:44) and the data points are shuffled before training.

A linear regression model (LR) is implemented based on the formulation in section 2. The plane describes the fit for all utilizations with one rule combination, denoted by a specific value for the 18 input parameters given.

A very short preliminary study conducted by the authors showed, that a neural network (NN) with 3 layers, 32 Neurons each, produces feasible results. The default activation function RELU, is used. The hyper parameter configuration has not been optimize to fit the dataset. For visual validation, the utilization of machines and AGVs is varied in step size 1 % and plotted against the mean tardiness forecasted with the neural network (NN). It can be seen, that the behavior of the model in Fig. 4 seems to be plausible.

For comparison of both models, the Mean Absolute Percentage Error (MAPE), Mean absolute error (MAE) and Mean Square Error (MSE) are used to evaluate the models and presented in Table 6.

Table 6. Performance indicators for the comparison of the two regression models. The neural network shows superior behavior regarding the MAPE.

Regression Method	MSE	MAE	MAPE
LR	150,915	240	171 %
NN	14,559	53	26 %

Given the idea, that the interaction effect of routing, dispatching and sequencing has a high degree of complexity, this proves the neural network to be suitable to represent the relation between these parameters.

5.3. Finding the presumably best combination

Given the trained model, the mean tardiness can be forecasted for every combination of system load and rule combination. An excerpt from the data is presented in Table 7, showing the particular rule combination of (STT / FIFO / SQ).

Table 7. Forecasted mean tardiness for one reference rule combination in time units. Based on the simulation data and the trained regression model the mean tardiness is forecasted for different system loads.

Reference STT / FIFO / SQ	90 % AGV Utilization	65 % AGV Utilization
90 % Machine Utilization	2,968.53	3,015.90
70 % Machine Utilization	350.87	215.54
20 % Machine Utilization	-45.38	-46.31

In comparison to the reference rule the mean tardiness as well as the improvement in (%) for the combination (LUV / SPT / SQ) is presented in Table 8.

Table 8. Forecasted mean tardiness for LUV / SPT / SQ in time units. The mean tardiness can be reduced up to 9 % under specific system load.

Reference LUV / SPT / SQ	90 % AGV Utilization	65 % AGV Utilization
90 % Machine Utilization	3,280.47 (-11 %)	2,833.88 (6 %)
70 % Machine Utilization	322.34 (8 %)	195.85 (9 %)
20 % Machine Utilization	53.45 (-218 %)	- 35.57 (-23 %)

In Fig. 4, multiple rule combinations are drawn in one plot. The machine utilization as well as the AGV utilization are presented on the x and y axis. Each color represents a specific rule combination. Detailed analysis prove that, in comparison to a reference rule combination (STT / FIFO / SQ), the mean tardiness (represented on the z axis) can be reduced up to 15 % choosing a suitable rule combination base on system load.

Based on a threshold assigned by the intersecting surfaces and a cyclic calculation of the system utilization, a new combination could be set. This results in the lowest tardiness at the given system load. Using this method leads to a gird, showing under which system load a certain rule combination behaves best.

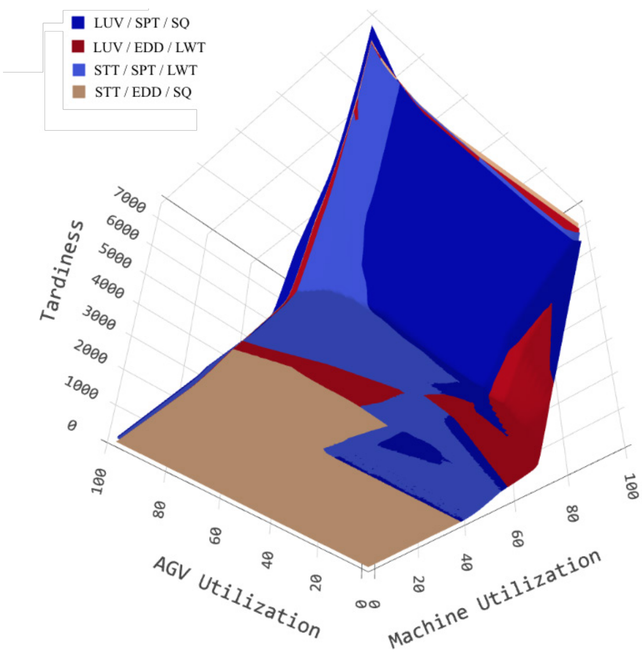

Fig. 4. Multiple neural networks drawn together in one plot. Each color represents a certain rule combination. Whenever two surfaces intersect, a rule change is recommended.

6. Conclusion

Due to their simplicity, priority-based rules have been used for sequencing operations, routing jobs and dispatching vehicles. In a flexible manufacturing system as the one presented, these rules tend to interact with each other in a highly complex relation. To be able to find a good combination, suiting the actual system load, the simulation study provides a large set of possible combinations of system performance and a resulting system performance measure (tardiness). Still, not all possible combinations of the system load can be simulated. For that reason, the system performance based on the rule combination has to be forecasted. An accurate neural network has been trained to do so, providing a forecast of the mean tardiness of the jobs at the given system status. Comparing different rule combinations, utilization specific sets providing the lowest possible tardiness are identified. In the future, the consideration of multi attribute rules for sequencing as well as routing could be considered. Furthermore, the scenario has to be adjusted to suit real world problems especially in size. Finally, to be able to assess the quality of the presented approach, a central optimization approach should be compared.

Acknowledgments

The authors would like to thank Sughanthan Sekar who was involved in the coding and supporting the work. Furthermore, the authors want to thank the editor and reviewers for their valuable comments and suggestions, which helped to improve the paper.

References

[1] ElMaraghy H, Schuh G, ElMaraghy W *et al.* Product variety management. CIRP Annals - Manufacturing Technology 2013; *62*: 629–652.
[2] Aalaei A, Davoudpour H. Revised multi-choice goal programming for incorporated dynamic virtual cellular manufacturing into supply chain management. Engineering Applications of Artificial Intelligence 2016; *47*: 3–15.
[3] Wemmerlöv U, Hyer NL. Cellular manufacturing in the U.S. industry. International Journal of Production Research 1989; *27*: 1511–1530.
[4] Thoben K-D, Wiesner S, Wuest T. Industrie 4.0 and Smart Manufacturing. International journal of automation technology 2017; *11*: 4–16.
[5] Brucker P, Heitmann S, Hurink J *et al.* Job-shop scheduling with limited capacity buffers. OR Spectrum 2006; *28*: 151–176.
[6] Heger J, Voss T. Optimal Scheduling of AGVs in a Reentrant Blocking Job-shop. Procedia CIRP 2018; *67*: 41–45.
[7] Kousi N, Koukas S, Michalos G *et al.* Scheduling of smart intra – factory material supply operations using mobile robots. International Journal of Production Research 2018; *59*: 1–14.
[8] Panwalkar SS, Iskander W. A Survey of Scheduling Rules. Operations Research 1977; *25*: 45–61.
[9] Browne J, Dubois D, Rathmill K *et al.* Classification of flexible manufacturing systems. The FMS magazine 1984; *2*: 114–117.
[10] Sharma P, Jain A. Effect of routing flexibility and sequencing rules on performance of stochastic flexible job shop manufacturing system with setup times: Simulation approach. Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture 2016; *231*: 329–345.
[11] Kumar R. Simulation of Manufacturing System at Different Part Mix Ratio and Routing Flexibility. Global Journal of Enterprise Information System 2016: 10–14.
[12] Doh H-H, Yu J-M, Kim J-S *et al.* A priority scheduling approach for flexible job shops with multiple process plans. International Journal of Production Research 2013; *51*: 3748–3764.
[13] Bilge Ü, Esenduran G, Varol N *et al.* Multi-attribute responsive dispatching strategies for automated guided vehicles. International Journal of Production Economics 2006; *100*: 65–75.
[14] Guan X, Dai X. Deadlock-free multi-attribute dispatching method for AGV systems. The International Journal of Advanced Manufacturing Technology 2009; *45*: 603.
[15] El-Bouri AW, 2016. A score-based dispatching rule for job shop scheduling, in *2016 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, IEEE, p. 774.
[16] Heger J, Branke J, Hildebrandt T *et al.* Dynamic adjustment of dispatching rule parameters in flow shops with sequence-dependent set-up times. International Journal of Production Research 2016; *54*: 6812–6824.
[17] Scholz-Reiter B, Heger J, Hildebrandt T, 2010. Analysis of Priority Rule-Based Scheduling in Dual-Resource-Constrained Shop-Floor Scenarios, in *Machine learning and systems engineering*, Springer Science & Business Media.
[18] Park Y, Kim S, Lee Y-H. Scheduling jobs on parallel machines applying neural network and heuristic rules. Computers & Industrial Engineering 2000; *38*: 189–202.
[19] Fahrmeir, L., Kneib, T., Lang, S., 2009. *Regression*. Springer Berlin Heidelberg, Berlin, Heidelberg.
[20] Heger, J., 2014. *Dynamische Regelselektion in der Reihenfolgeplanung*. Springer Fachmedien Wiesbaden, Wiesbaden.
[21] Nasiri MM, Yazdanparast R, Jolai F. A simulation optimisation approach for real-time scheduling in an open shop environment using a composite dispatching rule. International journal of computer integrated manufacturing 2017; *30*: 1239–1252.
[22] Poppenborg J, Knust S, Hertzberg J. Online scheduling of flexible job-shops with blocking and transportation. European J. of Industrial Engineering 2012; *6*: 497.
[23] Kim CW, Tanchoco JAM, Koo P-H. AGV dispatching based on workload balancing. International Journal of Production Research 1999; *37*: 4053–4066.