Simulation based priority rules for scheduling of a flow shop with simultaneously loaded stations

Professor Dr.-Ing. Frank Herrmann
Hochschule Regensburg - University of Applied Sciences Regensburg
Innovation and Competence Centre for Production Logistics and Factory Planning (IPF)
PO box 120327, 93025 Regensburg, Germany
E-Mail: Frank.Herrmann@HS-Regensburg.de

KEYWORDS
Simulation of processing time, scheduling, flow-shop, no-buffer (blocking), no-wait, priority rules, real world application.

ABSTRACT
In this study, a real world flow shop with a transportation restriction is regarded. This restriction reduces the set of feasible schedules even more than the no-buffer restrictions discussed in the literature in the case of limited storage. Since this scheduling problem is integrated in the usual hierarchical planning, the tardiness is minimised. This NP-hard problem is solved by priority rules, because of the addressed dynamic environment (and a high number of jobs) at the company site. Due to the technological restrictions the real duration of a job is significantly larger than its net processing time. By a simulation of the processing time many priority rules are improved and some even significantly. With extensive simulations successful priority rules in the literature are analysed.

1. INTRODUCTION

Specific products are produced by special machines which are often grouped in a flow shop. They have to produce small batches with short response times, so scheduling algorithms are needed to ensure that under the constraint of a high average load of the flow shop, the due dates of the production orders are met. Nowadays, such special designed flow shops often have technological restrictions, which complicate the scheduling. For example in cell manufacturing, buffer could be non-existent due to limited space and storage facilities. So, in recent years, a considerable amount of interest has arisen in no-buffer (blocking) scheduling problems and in no-wait scheduling problems, with makespan as objective criteria. Often these production systems deliver products for other systems as well. Due to the hierarchical planning which is implemented in enterprise resource planning systems (ERP system) (see e.g. Jacobs et al. 2010), the local completion times in one production system in many cases determine the earliest possible starting times in another production system. Thus, the delay of the operations in a production system has an impact on the effectivity of this coordination process. Therefore, tardiness is considered as objective criteria.

2. A REAL WORLD APPLICATION

The problem is a modification of a partly automated production line at Fiedler Andritz in Regensburg to produce filter (baskets) with a lot size of 1. All filters have unified constructions. They differ in varying heights of the baskets and there exist different designs.
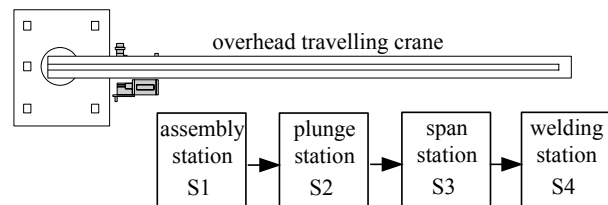

Figure 1: Structure of the production line

The production line consists of 4 stations which are shown in Figure 1. Station 1 assembles 6 single batons (called consoles) on an assembly ground plate to a skeleton of a filter basket. Baton profiles are assembled into the provided slots of the filter basket skeletons. At the plunge station a wire coil is contrived in the device of a lining machine. The lining machine straightens the wire and inserts batons into the slots. To ensure stability, the span station installs span kernels in the case of outflow filter baskets and span belts in the case of inflow filter baskets. Then, the filter basket is lifted from the assembly ground plate and is transported to the welding station, at which the baton profiles are welded on the filter basket skeletons. The accomplished filter basket leaves the production line. Prior to this, the span medium is removed. An overhead travelling crane lifts a filter basket out of a station, transports it to the next station and inserts it directly in this station. This is just possible if this station is free. So, there is no buffer in the production line and each feasible schedule of jobs is a permutation of these jobs. Due to other operational issues the crane can just be moved if all stations are inactive. Since an operation cannot be interrupted, the transport has to be performed after the completion of all operations on the stations in the flow shop. Due to further operational issues this restriction has to be applied also for the first and the last station; note, that the crane loads S1 and unloads S4 as well. In summary, all stations are loaded and unloaded with filters during a common process and this process starts with the last station S4, followed by station S3, S2, until station S1 is reached. It is allowed that a station is empty; then this station is skipped (may be partially) in this process.

There are 10 part types whose processing times are listed in Table 1.

Table 1: Processing times for all part types in minutes

Part type	Station				Sum of times
	S1	S2	S3	S4	
P1	100,5	50	53,5	9	213
P2	256,5	50	53,5	9	369
P3	122	135	90	75	422
P4	256,5	50	267	9	582,5
P5	182	200	135,5	140	657,5
P6	100,5	300	53,5	300	754
P7	223	250	196	220	889
P8	223	250	206,5	220	899,5
P9	100,5	300	267	300	967,5
P10	256,5	300	267	300	1123,5

At the company's production site, the jobs for filters are generated by an SAP system and produced filters are stored before they are assembled into other products or sold directly to customers. Therefore, all jobs with a release date after the beginning of a period are released at the beginning of this period. One period consists of one day with three 8 hour shifts. For this investigation, sequences of jobs of filter types with lot size 1 are randomly generated for each period t by an generating algorithm which has been designed in accordance with the proceeding in (Russel et al. 1987) and in (Engell et al. 1994): An additional filter type F, released in period t, consumes capacity on each station during the time between t and its due date; the calculation for the capacity just uses the net processing time and does not regard the dependencies between the jobs (released so far). F is accepted as long as this consumed capacity on each station is below a maximal load level, otherwise it will be skipped to the next period. A maximal load level is an (intended) average load $\left(L_0(S)\right)$ plus 0, –30% and +30% of $L_0(S)$. Over the first 5, 10, 15 etc. consecutive periods, the load level variations average to zero.

In reality at the company, there are large numbers of periods with a low number of late jobs and large numbers of periods with a high (or even very high) number of late jobs. To achieve a comparable situation for this investigations, due dates are determined in a way so that scheduling with the FIFO rule (first-in-first-out) causes a specific percentage of late jobs. The company confirmed that job sequences with 30%, 50%, 70% and 85% of late jobs by scheduling with the FIFO rule (called time pressure) are comparable to the ones which occurred in the real operation. As a result of the generating algorithm's calculations the mean difference between the due dates and the release dates are between 2 and 3,5 days with a standard deviation of 0,5 days. Andritz Fiedler has confirmed that such timeframes for processing jobs are representative.

The time needed for loading and unloading a station is negligible compared to the duration of the operation itself. In addition this task is independent from the allocation (or loading) of other stations and the required time is included in the duration of the operation.

The general scheduling problem consists of M stations and a pool of N jobs, which may change at any time, with known earliest possible starting times for release dates a_i $(1 \le i \le N)$ and due dates f_i $(1 \le i \le N)$ respectively. Also there is the duration $t_{i,j}$ of operation $\left(o_{i,j}\right)$ j $(1 \le j \le M)$ of job i $(1 \le i \le N)$, which is being worked on station j. As performance criteria average tardiness $\left(T_{Mean}\right)$ and standard deviation of the tardiness $\left(T_\sigma\right)$ are primarily analysed.

The time between two consecutive executions of the load process is determined by the maximum of the duration of the operations on the stations in the flow shop. This is called cycle time. This "load"-restriction, the no-buffer condition and the capacity of the stations are the main restrictions. Setup times are relatively small compared to operation times and are included in those.

The no-buffer condition means a relaxation of the scheduling problem with the (above) "load"-restriction. Scheduling problems with the no-buffer are proven to be NP-hard in the strong sense for more than two stations; see e.g. (Hall and Sriskandarajah 1996), which contains a good survey of such problems.

3. LITERATURE REVIEW
As mentioned earlier, the real application is close to the class of no-buffer (blocking) scheduling problems. Solutions for the no-buffer (blocking) scheduling problems are published in various papers. In (McCormick et al. 1989) a schedule is extended by a (unscheduled) job that leads to the minimum sum of blocking times on machines which is called profile fitting (PF). Often the starting point of an algorithm is the NEH algorithm presented in (Nawaz et al. 1983), as it is the best constructive heuristic to minimize the makespan in the flow shop with blocking according to many papers, e.g. (Framinan et al. 2003). Therefore, (Ronconi 2004) substituted the initial solution for the enumeration procedure of the NEH algorithm by a heuristic based on a makespan property proven in (Ronconi and Armentano 2001) as well as by the profile fitting (PF) of (McCormick et al. 1989). (Ronconi 2005) used an elaborated lower bound to realise a branch-and-bound algorithm which becomes a heuristic since the CPU time of a run is limited. Also for minimising makespan, (Grabowski and Pempera 2007) realised and analysed a tabu search algorithm. As an alternative approach, (Wang and Tang 2012) have developed a discrete particle swarm optimisation algorithm. In order to diversify the population, a random velocity perturbation for each particle is integrated according to a probability controlled by the diversity of the current

population. Again, based on the NEH algorithm, (Wang et al. 2011) described a harmony search algorithm. First, the jobs (i.e. a harmony vector) are ordered by their non-increasing position value in the harmony vector, called largest position value, to obtain a job permutation. A new NEH heuristic is developed on the reasonable premise that the jobs with less total processing times should be given higher priorities for the blocking flow shop scheduling with makespan criterion. This leads to an initial solution with higher quality. With special settings as a result of the mechanism of a harmony search algorithm, better results are achieved. Also (Ribas et al. 2011) presented an improved NEH-based heuristic and uses this as the initial solution procedure for their iterated greedy algorithm. A modified simulated annealing algorithm with a local search procedure is proposed by (Wang et al. 2012). For this, an approximate solution is generated using a priority rule specific to the nature of the blocking and a variant of the NEH-insert procedure. Again, based on the profile fitting (PF) approach of (McCormick et al. 1989), (Pan and Wang 2012) addressed two simple constructive heuristics. Then, both heuristics and the profile fitting are combined with the NEH heuristic to three improved constructive heuristics. Their solutions are further improved by an insertion-based local search method. The resulting three composite heuristics are tested on the well-known flow shop benchmark of (Taillard 1993), which is widely used as benchmark in the literature.

To the best of my knowledge, only a few studies investigate algorithms for the total tardiness objective (for flow shops with blocking). (Ronconi and Armentano 2001) have developed a lower bound which reduces the number of nodes in a branch-and-bound algorithm significantly. (Ronconi and Henriques 2009) described several versions of a local search. First, with the NEH algorithm, they explore specific characteristics of the problem. A more comprehensive local search is developed by a GRASP based (greedy randomized adaptive search procedure) search heuristic.

4. HEURISTIC SOLUTION BY PRIORITY RULES
The real application operates in dynamic environments where real time events like station failure, tool breakage, arrival of new jobs with high priority, changes of due dates etc. may turn a feasible schedule into an infeasible one. A feasible assignment of a job is achieved by a priority rule like earliest due date (EDD), because a priority rule orders a queue in front of a station quasi immediately. So, priority rules are still analysed in many studies on scheduling; one example of a recent one is (El-Bouri 2012).

Due to the "load"-restriction the processing time of a job A on the flow shop depends on the other jobs processed on the flow shop at the same time. Therefore, its processing time can be significant larger than the sum of the processing times of its single operations

(t_A), called net processing time of job A. A realistic processing time for a job A is achieved, if the processing on the flow shop is simulated with respect to the jobs processed on this flow shop at the same time. After four cycles $(\tau_1,...,\tau_4)$, job A leaves the flow shop. So, these cycle times depend on the three jobs (X1, X2, and X3) on the flow shop as A is assigned to station 1, and the three jobs (B1, B2, and B3) following A in the sequence; this is illustrated in Figure 2. So, $\tau_1 + \tau_2 + \tau_3 + \tau_4$ is the (total) processing time of A (tt_A). This processing time of job A is only correct, if the tail is (structurally) identical with the three jobs which follow job A in the final permutation. Since these are not known a deviation normally occurs between this calculated processing time and the processing time which will really occur if job A is chosen.
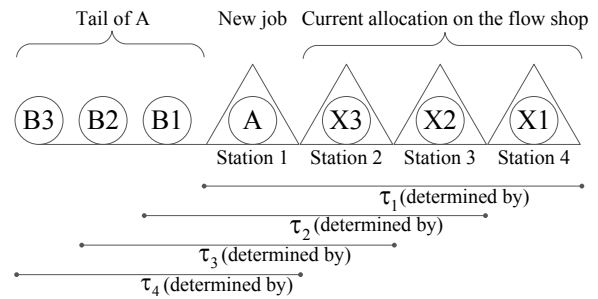

Figure 2: Calculation the processing time of a job A

As known, tardiness is improved by assigning jobs with a small slack; the slack for job i is defined by $f_i - t - tt_i$, where t is the current time and f_i is the due date of job i. Investigations by the author (see also (Engell et al. 1994)) show, that for many job shop problems the rules $CR+SPT = \begin{cases} \dfrac{f_i - t}{tt_i}, & f_i - t - tt_i > 0 \\ tt_i, & f_i - t - tt_i \leq 0 \end{cases}$

(tt_i is the shortest processing time (SPT), which determines the SPT rule), ODD (which is identical with the EDD-rule for the class of problems regarded here) and $SL/OPN = \dfrac{f_i - t - tt_i}{M}$ (note: a low value is always preferred), which is identical with the slack rule (SL) are Pareto optimal to the average, the standard deviation and the maximum tardiness. This explains why SL/OPN and CR+SPT are often used as benchmark; according to (Raghu and Rajendran 1993) other combinations deliver worse results for flow shop problems. As known from several investigations SPT based rules have the tendency to small T_{Mean} at the expense of large T_σ and slack based rules tend to a small T_σ at the expense of a high T_{Mean} ; SPT and SL are representatives for this effect.

In addition, more recent rules are adapted to the class of problems regarded here. One is RR, see (Raghu and

Rajendran 1993), which seeks to minimize both mean flow time and mean tardiness of jobs and has originally been defined for job shop problems, see (Raghu and Rajendran 1993). For the adaption the remaining work content is tt_i and the probable waiting time of the successor of this operation (in job i) at the (next) station, which seems to be less effective in flow shops for minimising mean flow time due to (Rajendran and Holthaus 1999), is integrated in the calculation of the cycle times. Thus, the priority index is $\left(f_i - t - tt_i\right) \cdot e^{-\eta} + e^{\eta} \cdot tt_i$, with utilisation level η of the entire flow shop defined by $\eta = \dfrac{b}{b+j}$ with b being the busy time and j being the idle time of the entire flow shop; the job with the lowest priority index will be processed next.

Using the properties of an optimal solution for the single-station weighted tardiness scheduling problem, (Rachamadugu and Morton 1982) developed the weighted slack-based scheduling rule RM. According to the analysis of this rule performed in (Rachamadugu and Morton 1982), it can be assumed that it delivers near optimal results to the one station problem. (Rachamadugu 1987) shows that this rule can be used as an effective pruning device in enumerative methods. Modifications of this rule deliver very good results for flow shop and job shop problems with weighted tardiness criteria (see (Vepsatainen and Morton 1987)). The rule was successfully adapted to resource constrained project scheduling problems (RCPSP) in (Voß and Witt 2007). The priority index for an operation j of a job i is: $\dfrac{w_{i,j}}{t_{i,j}} \cdot e^{-\frac{k}{\bar{t}} \cdot \max\{sl_{i,j}, 0\}}$, where $sl_{i,j}$ is the slack, $w_{i,j}$ is a weight, k is an empirically determined "look-ahead" parameter and $\bar{t}$ is the average duration of all jobs processed at the same station; a low value is always preferred. For the real world application jobs get priorities, no weighting is used, so that $w_i = 1$ for each job i, $\bar{t} = \dfrac{1}{N} \sum_{i=1}^{N} tt_i$ and the slack is $f_i - t - tt_i$ as usual.

Especially, due to the relatively recent paper of (Voß and Witt 2007) it is motivated to integrate resource costs π_i, so that in total the priority index is: $\dfrac{1}{\pi_i} \cdot e^{-\frac{k}{\bar{t}} \cdot \max\{f_i - t - tt_i, 0\}}$. As local processing time costing $\pi_i^l = tt_i$ is used, called RM local, and as global processing time costing $\pi_i^g = \sum_{i \in U_t} tt_i$, where U_t is the set of unfinished jobs in the pool of orders, excluding job i, is used, called RM global; in (Voß and Witt 2007) job i is included, because a RCPSP is regarded, but here this would mean that for all jobs i waiting in front of the

flow shop have the same π_i^g. (Lawrence and Morton 1993) considered additional resource costs which do not fit to our problem or have already been integrated (in other parameters); e.g. a (static) bottleneck resource cost is not applicable because it is assumed that each station is temporary a bottleneck.

5. COMPUTATIONAL RESULTS
The real world application is realised in the simulation tool "Plant Simulation" together with an implementation of the above mentioned hierarchical planning as realised in ERP systems, used in industrial practise. Average tardiness $\left(T_{Mean}\right)$ and standard deviation of the tardiness $\left(T_\sigma\right)$ reach a steady state by a simulation horizon of 5000 periods.

A large number of preliminary studies show that the parameter k has a significant impact on the performance of the local and the global RM rule and the best results are achieved with $k = 1$. In (Voß and Witt 2007) k has a negligible influence, but they regard an RCPSP.

The impact of the simulated processing time is analysed first. Due to its implementation, its concrete values depend from the (3) jobs on the flow shop and the (3) jobs following. So, the values should be independent from the priority rules and the time pressure. But, it could be that individual rules prefer permutations of jobs in a cycle. This is observed in experiments with the (above explained) priority rules. In these experiments the mean, the standard deviation, the minimum and the maximum of the simulated processing times are measured. The measured values for all part types and all rules are between those for the SPT rule and the SL rule which are listed in Table 2. These results demonstrate a huge deviation from the net processing time.

Furthermore, the experiments show a significant impact by the tail. The values in Table 2 are representative for many tails. Exceptions occur if each tail only consists of jobs with a small net processing time, as with part type P1, or a high one, as with part type P10. In the first case the simulated processing times have small mean values, e.g. for part type P1 901,1 minutes and for part type P10 1067,6 minutes, and in the second case the mean values are large, e.g. 1248,5 minutes for part type P1 and 1269,2 minutes for part type P10. The standard deviations are in the first case huge e.g. for part type P1 173,1 minutes and for part type P10 187,6 minutes, and the second case low e.g. for part type P1 77,4 minutes and part type P10 61,3 minutes.

As indicated by these results, the performance of a priority rule is influenced by the concrete tail. Instead of regarding all 1000 possible tails, the study is limited on tails whose part types have similar net processing times and those whose part types have significantly different net processing times. The experiments for each priority rule show both significant and minor deviations. These

significant deviations are not exceptions and the reduction is almost one-third (or even more). Results which are close to the best possible results are achieved, if each tail consists of jobs of part type P4 only (i.e. P4, P4, P4); this tail is used in the following. Experiments show that an accidently tail (i.e. the part types for a tail are randomly selected) is a very good alternative.

Table 2: Net processing time (NET) and simulated processing time in minutes for the rules SPT and SL

Part type	SPT				
	Mean	Net	Standard deviation	Mini-mum	Maxi-mum
P1	1116,5	213	151,8	885	1349.5
P2	1159,8	369	100,45	1061	1349.5
P3	1088,6	422	138,1	929	1349.5
P4	1151,7	582,5	96,4	1061	1349.5
P5	1162,9	657,5	88,5	1063,5	1349.5
P6	1233,3	754	95,3	1098,5	1376
P7	1228,1	889	60,1	1164,5	1349.5
P8	1225,5	889,5	57,8	1164,5	1349.5
P9	1237,2	967,5	93,9	1098,5	1376
P10	1322,8	1123,5	27,9	1098,5	1376
	SL				
	Mean	Net	Standard deviation	Mini-mum	Maxi-mum
P1	1179	213	126,5	885	1349.5
P2	1192,6	369	112,3	885	1349.5
P3	1179,3	422	118,1	885	1349.5
P4	1181,7	582,5	111,5	885	1349.5
P5	1184,6	657,5	106,5	885	1349.5
P6	1198,3	754	106,1	885	1376
P7	1203,8	889	101,7	885	1376
P8	1207,9	889,5	98,2	885	1376
P9	1215,9	967,5	98,1	885	1376
P10	1225,9	1123,5	97,5	885	1376

Table 3 contains the percental changes by using simulated processing times instead of net processing times. Due to the characteristic behaviour of the SPT rule and the SL rule to T_{Mean} and T_σ (see above), both are analysed first. The SPT rule benefits most from a more realistic processing time. In case of the SL rule there are just small improvements but often significant deteriorations. As long as the CR+SPT rule uses CR – which is some kind of slack – there are deteriorations for T_{Mean}. Since the net processing time is much smaller than the simulated processing time, the CR+SPT rule with simulated processing time decides earlier according to SPT as the CR+SPT rule with net processing time which explains the improvement. Both rules RR and RM are a combination of slack and shortest processing time (SPT). The RR rule benefits from a more precise processing time. A smaller (percental) improvement compared to the SPT rule is caused by already better values if net processing is used and the impact of the slack; The impact of the slack is shown most clearly by a time pressure of 85%. Much

better (percental) improvements for T_{Mean} in case of a small time pressure is caused because RR rule prefers critical jobs with positive slack much better than the SPT rule (which does not take due dates into account). The RM rule prefers small jobs if there is no slack and otherwise jobs with small slack. Depending on the degree of influence of the slack on the priority, it can be expected that their changes are between those of the rules SL and CR+SPT. The values in Table 3 confirm this conclusion primarily for the rule RM global. The concrete values depend on the time pressure. In the case of rule RM local the percental improvements for T_{Mean} and for T_σ with small time pressure are comparable to the ones of the SPT rule. In the case of T_σ a better processing time causes an increase of the (absolute) values for this rule on the level of the values of the SPT rule, except for a low time pressure.

Table 3: Change by using simulated processing time instead of net processing time in percent, compared to the result with net processing time.

Rule	Time pressure			
T_{Mean}	30%	50%	70%	85%
SPT	57.6	10.3	12.3	25.8
SL	-41.2	-2.3	0.23	-17
CR+SPT	-4	-10.3	-7.4	8.9
RR	75.3	4.2	3.9	11.9
RM local	48.9	10.4	11.7	21.5
RM global	3.3	-10.9	-4	-5.8
T_σ	30%	50%	70%	85%
SPT	68.65	31.3	27.2	48.11
SL	-20.1	1.2	3.4	-19
CR+SPT	51.4	34.1	37.1	48.6
RR	69.8	14.3	14.4	23.6
RM local	68.3	-21.2	-25.5	-242.8
RM global	50	-10.1	10.4	67.9

The absolute values, listed in Table 4 (just the ones by using simulated processing times), differ partially from the results published in other papers. In order to judge this and the following more detailed results, it should be pointed out that the real world problem in this paper has a very special problem structure, compared to the representative problems usually regarded in the literature. The above mentioned expectation of the SPT and the SL rule – namely small T_{Mean} at the expense of large T_σ by the first rule and the opposite by the second one – is fulfilled. For many job shop and flow shop problems the CR+SPT rule outperforms the best results from the SPT and the SL rule. For this real world application CR+SPT delivers always, with one exception, results, which are much worse than the ones of all other rules. Since the rules RR and RM also combine slack and SPT, the worse results of CR+SPT are caused by a too late switch from preferring jobs with small slack to jobs with small SPT. Compared to other job shop problems this is more significant because just a

misguided decision causes long cycles, which reduces the remaining slack for all other jobs much more than in the case of typical job shop problems. In addition, these long cycles could be very ineffective in terms of large times on some stations without any processing.

Table 4: Absolute performance measures for priority rules with simulated processing time in minutes

Rule	Time pressure			
T_{Mean}	30%	50%	70%	85%
SPT	99.1	323.3	326.2	646.9
SL	161.6	321.5	344.7	1149.8
EDD	103.4	300.2	342.2	948.7
CR+SPT	581.97	575.7	574.6	1032.7
RR	40.7	279.5	313.5	823.1
RM local	55.3	267.1	278.6	626.2
RM global	134.9	346.7	359.2	1008.5
T_{σ}	30%	50%	70%	85%
SPT	314.5	449.2	473.4	826.7
SL	353.4	315.3	326.6	787.6
EDD	243.1	293.1	328.9	643.4
CR+SPT	2023.9	916.4	901.2	1464.1
RR	125.1	282.8	305.7	564.6
RM local	235.9	425.8	454.1	1187.4
RM global	395.5	456.2	464.6	890.04

The performance of the RR rule for the real world application is primarily compared with the results in (Rajendran and Holthaus 1999), because in (Raghu and Rajendran 1993) an open shop problem is regarded. In (Rajendran and Holthaus 1999) the RR rule delivers better results than the other rules except for one case. In detail, in most of the cases, the improvements are less significant (partially much less) and the sequences of the regarded rules according to both performance criteria, are different. The results of the priority rules in (Rajendran and Holthaus 1999) and in (Raghu and Rajendran 1993) are significantly impacted by the station utilisation levels, which are primarily 80% and 95%. In addition (Rajendran and Holthaus 1999) uses an allowance factor and (Raghu and Rajendran 1993) uses the due date tightness, which both have a much smaller impact than the station utilisation level. In this investigation the time pressure is only increased by using tighter due dates while the load on the stations remains unchanged. Therefore, it can be expected that the time pressure here has a similar effect as the allowance factor in (Rajendran and Holthaus 1999) or the due date tightness in (Raghu and Rajendran 1993). A much larger impact is caused by a significant fluctuation of the load in the periods, so there are some periods where the load is either much higher than 95% and or much lower than 80% (as in (Rajendran and Holthaus 1999) or in (Raghu and Rajendran 1993)). Thus a tighter due date has a more significant effect in periods with relative very high load than in periods with a relative very small load. In total this seems to cause the different amount of improvement.

The poor results of RM global compared with RM local (with one exception) contradicts the results in (Lawrence et al. 1993). The same happens in the investigation of (Voß and Witt 2007). The local processing time costing prefers more often jobs with short processing times than the global processing time costing. This procedure explains the much better results in the case of those time pressures for which the SPT rule delivers a much better T_{Mean} than to the SL rule. In the other cases this procedure is beneficial if many tardy jobs are waiting in front of the flow job. Finally, in the flow shop in (Voß and Witt 2007) with parallel resources and setup states the differences in the results of the rules are smaller than in this investigation and vice versa for the more general problem structure in (Voß and Witt 2007).

Overall, the simulated processing times should be used in the priority rules. Then, the rules RR and RM local deliver the best mean tardiness. RR is beneficial with low and RM local with (very) high time pressure. The RR rule delivers the best standard deviation of the tardiness (for all time pressures).

6. CONCLUSIONS
This paper presents a real world flow shop scheduling problem with more restrictive restrictions than the ones normally regarded in literature. To ensure online scheduling this investigation is restricted to those priority rules, which are considered in literature as being very effective. The substitution of the net processing time, normally used in priority rules, by a simulated one delivers often significant better results. Some tests with optimisation solutions for a small test problem indicate that priority rules do not recognise when to prefer large variances of cycle times or small ones, respectively, and that the schedules of priority rules have outliers in the cycle times, which are usually avoided by optimal solutions. With this and other characteristics of optimal solutions an efficient metaheuristic like local search or genetic algorithm shall be developed next.

Up to now, the workers for the manual tasks are not scheduled. In addition, limited resources, like the available number of coils or assembly ground plates, occur at some company sites. Such requirements are also left to future investigations.

REFERENCES
El-Bouri, A. 2012. "A cooperative dispatching approach for minimizing mean tardiness in a dynamic flowshop". *Computers & Operations Research*, Volume 39, Issue 7 (July), 1305 – 1314.

Engell, S.; F. Herrmann; and M. Moser. 1994. "Priority rules and predictive control algorithms for on-line scheduling of FMS". In *Computer Control of Flexible Manufacturing Systems*, S.B. Joshi and J.S. Smith (Eds.). Chapman & Hall, London, 75 – 107.

Framinan, J.M.; R. Leisten; and C. Rajendran. 2003. "Different initial sequences for the heuristic of

Nawaz, Enscore and Ham to minimize makespan, idletime or flowtime in the static permutation flowshop sequencing problem". *International Journal of Production Research*, 41, 121 – 148.

Grabowski, J. and J. Pempera. 2007. "The permutation flow shop problem with blocking. A tabu search approach.". *Omega*, 35 (3), 302 – 311.

Hall, N.G. and C. Sriskandarajah. 1996. "A survey of machine scheduling problems with blocking and no-wait in process". *Operations Research* 44 (3), 510–525.

Jacobs, F.R.; W. Berry; D. Whybark; T. Vollmann. 2010. "Manufacturing Planning and Control for Supply Chain Management". McGraw-Hill/Irwin (New York), 6 edition.

Lawrence, S. and T. Morton. 1993. "Resource-constrained multi-project scheduling with tardy costs: Comparing myopic, bottleneck, and resource pricing heuristics.". *European Journal of Operational Research* 64, 168 – 187.

McCormick, S.T.; M.L. Pinedo; S. Shenker; and B. Wolf. 1989. "Sequencing in an assembly line with blocking to minimize cycle time". *Operations Research*, 37 (6), 925 – 935.

Nawaz, M.; E.E. Enscore; and I. Ham. 1983. "A heuristic algorithm for the m-machine, n-job flow sequencing problem". *Omega*, 11(1), 91 – 95.

Pan, Q.; and L. Wang. 2012. "Effective heuristics for the blocking flowshop scheduling problem with makespan minimization". *Omega*, 40 (2), 218 – 229.

Rachamadugu, R.M.V. 1987. "Technical Note –A Note on the Weighted Tardiness Problem". *Operations Research*, 35, 450 – 452.

Rachamadugu, R.V. and T.E. Morton. 1982. "Myopic heuristics for the single machine weighted tardiness problem". Working Paper No. 28-81-82, Graduate School of Industrial Administration, Carnegie-Mellon University, Pittsburgh, PA.

Raghu, T.S. and C. Rajendran. 1993. "An efficient dynamic dispatching rule for scheduling in a job shop". *International Journal of Production Economics* 32, 301 – 313.

Rajendran, C and O. Holthaus. 1999. "A comparative study of dispatching rules in dynamic flowshops and job shops". *European Journal of Operational Research*, 116 (1), 156 – 170.

Ribas, I.; R. Companys; and X. Tort-Martorell. 2011. "An iterated greedy algorithm for the flowshop scheduling problem with blocking". *Omega*, 39, 293 – 301.

Ronconi, D.P. and V.A. Armentano. 2001. "Lower Bounding Schemes for Flowshops with Blocking In-Process". *Journal of the Operational Research Society*, 52 (11), 1289 – 1297.

Ronconi, D.P. 2004. "A note on constructive heuristics for the flow-shop problem with blocking". *International Journal of Production Economics*, 39 – 48.

Ronconi, D.P. 2005. "A branch-and-bound algorithm to minimize the makespan in a flowshop with blocking". *Annals of Operations Research*, (138), 53 – 65.

Ronconi, D. and L. Henrique. 2009. "Some heuristic algorithms for total tardiness minimization in a flow shop with blocking". *Omega*, 37 (2), 272 – 281.

Russel, R.S.; E.M. Dar-El; and B.W. Taylor. 1987. "A comparative analysis of the COVERT job sequencing rule using various shop performance measures". *International Journal of Production Research*, 25 (10), 1523 – 1540.

Taillard, E. 1993. "Benchmarks for basic scheduling problems". *European Journal of Operational Research*, 64 (2), 278 – 285.

Vepsatainen, A.P. and T.E. Morton. 1987. "Priority rules for job shops with weighted tardiness costs". *Management Science* 33/8, 95 – 103.

Voß, S. and A. Witt. 2007. "Hybrid Flow Shop Scheduling as a Multi-Mode Multi-Project Scheduling Problem with Batching Requirements: A real-world application.". *International Journal of Production Economics* 105, 445 – 458.

Wang, X. and L. Tang. 2012. "A discrete particle swarm optimization algorithm with self-adaptive diversity control for the permutation flow shop problem with blocking". *Applied Soft Computing*, (12, 2), 652 – 662.

Wang, L.; Q.-K. Pan; and M.F. Tasgetiren. 2011. "A hybrid harmony search algorithm for the blocking permutation flow shop scheduling problem". *Computers & Industrial Engineering*, 61 (1), 76 – 83.

Wang, C.; S. Song, S.; J.N.D. Gupta; and C. Wu. 2012. "A three-phase algorithm for flowshop scheduling with blocking to minimize makespan". *Computers & Operations Research*, 39 (11), 2880 – 2887.

AUTHOR BIOGRAPHY

Frank Herrmann was born in Münster, Germany and went to the RWTH Aachen, where he studied computer science and obtained his degree in 1989. During his time with the Fraunhofer Institute IITB in Karlsruhe he obtained his PhD in 1996 about scheduling problems. From 1996 until 2003 he worked for SAP AG on various jobs, at the last as director. In 2003 he became Professor for Production Logistics at the University of Applied Sciences in Regensburg. His research topics are planning algorithms and simulation for operative production planning and control. His e-mail address is Frank.Herrmann@HS-Regensburg.de and his Web-page can be found at http://www.hs-regensburg.de/index.php?id=2457