Improving warehouse responsiveness by job priority management: A European distribution centre field study

Abstract

Warehouses employ order cut-off times to ensure sufficient time for fulfilment. To satisfy increasing consumer's expectations for higher order responsiveness, warehouses competitively postpone these cut-off times upholding the same pick-up time. This paper, therefore, aims to schedule jobs more efficiently to meet compressed response times. Secondly, this paper provides a data-driven decision-making methodology to guarantee the right implementation by the practitioners. Priority-based job scheduling using flow-shop models has been used mainly for manufacturing systems but can be ingeniously applied for warehouse job scheduling to accommodate tighter cut-off times. To assist warehouse managers in decision making for the practical value of these models, this study presents a computer simulation approach to decide which priority rule performs best under which circumstances. The application of stochastic simulation models for uncertain real-life operational environments contributes to the previous literature on deterministic models for theoretical environments. The performance of each rule is evaluated in terms of a joint cost criterion that integrates the objectives of low earliness, low tardiness, low labour idleness, and low work-in-process stocks. The simulation outcomes provide several findings about the strategic views for improving responsiveness. In particular, the critical ratio rule using the real-time queue status of jobs has the fastest flow-time and performs best for warehouse scenarios with expensive products and high labour costs. The case study limits the coverage of the findings, but it still closes the existent gap regarding data-driven decision-making methodology for practitioners of supply chains.

Keywords

responsiveness; queuing model; order fulfilment; cut-off operation; flow-shop scheduling

1 Introduction

Intense competition for speedy order fulfilment characterises current retail markets. Responsiveness (Barclay et al., 1996) includes the ability to react purposefully within appropriate time to external environments for securing competitive advantage. Improving order fulfilment responsiveness is a significant challenge for boosting customer satisfaction (Doerr and Gue, 2013) and many firms, such as Amazon Prime, invest substantial capital to propel responsiveness. Though responsiveness hones competitiveness, it often leads to resource misallocation (Vincent, 2011), and improved responsiveness leads for two-thirds of all firms to increased labour cost (Pearcy and Kerr, 2013). Web retailers show responsiveness by advertising 'Place an order before midnight for next-day delivery.' Customers are nowadays accustomed to fast demand satisfaction in online markets and expect commensurate off-line service. Off-line retailers, therefore, attract customers with promises such as: 'Buy online now and pick up in store tomorrow', forcing off-line retail distributors to improve their responsiveness (Denman, 2017).

The overall speed of order fulfilment in off-line markets depends on processing and transportation speeds from manufacturers through warehouses and retail shops to end-users. This paper focuses on speedy order fulfilment in warehouses, in particular, original equipment manufacturer (OEM) warehouses delivering to retailer warehouses. Their order fulfilment process includes the inbound processes of receiving products and putting them away and the outbound processes of picking, packing, staging and shipping. As OEM warehouses receive products from their manufacturer, the inbound process is easily controlled compared to the somewhat unpredictable consumer demand leading to fast fluctuations of retailer orders. Another characteristic of OEM warehouse is that retailers order relatively large quantities of relatively few products (Bartholdi and Hackman, 2011). This distinguishes such warehouses from those delivering directly to consumers, where order sizes are small and range over a much broader product assortment. Whereas picking is usually the crucial stage for the latter type, in OEM warehouses the packing stage is often the most demanding one. As the receiving retailer warehouses differ in capacity and layout and trucks should be loaded efficiently, re-palletising is a significant task for OEM warehouses. Because of the large order volumes, the re-palletising activities of unpacking, repacking and stacking are relatively labour intensive.

Responsiveness of OEM warehouses is measured by their flexibility to dispatch products ordered by retailers as fast as possible. To mitigate the effect of demand spikes, most OEM warehouses limit their fulfilment liability by daily order cut-off time agreements with their clients to ensure sufficient slack for order fulfilment by the earliest dispatch day (Van den Berg, 2007). To improve responsiveness, these warehouses try to postpone the cut-off time and to handle the same order volume with less slack. Since orders typically have different fulfilment deadlines and processing times, priority-based job scheduling offers the key to efficient solutions. Flow shop scheduling (Johnson, 1954) has notably reduced waste from over-production and waiting times in the manufacturing field. It is genuinely new to apply flow shop scheduling for improving responsiveness in warehouse order fulfilment under cut-off time challenges. Job scheduling timely manages to allocate prioritised tasks to labour resources for chosen goals (Vincent and Billaut,

2006) as the first decision-making in the OEM warehouse. The second decision-making here is how OEM warehouses should strategically choose the proper type of job scheduling to allow later cut-off times for enhancing responsiveness. The goal of this paper is to provide a data-driven method that improves decision-making capability to practitioners in the warehouses.

Warehouse operations are faced with various uncertainties, including dynamic arrival, service and departure times. In particular, unexpected order arrivals with different processing times can yield long delays. There is usually no priority rule that is universally optimal (Lee et al., 1997) in case of these uncertainties. Although there is much research on job-scheduling, it is still tricky for warehouse managers to select the most suitable scheduling rule for their circumstances. This paper presents an integrated decision system for cost-effective job scheduling using flow-shop priority methods to aid warehouses facing postponed order cut-off times. This framework integrates the multiple objectives of low earliness, low tardiness, low labour idleness, and low stocks through processing lanes into a single cost criterion, with weights derived from the cost structure and performance priorities of the warehouse. The methodology supports data-driven decision making by simulating stochastic models (Gong and De Koster, 2011) based on real-life operational data for order arrivals, due times, and service times. The framework assists warehouse practitioners in deciding which scheduling methods perform best under which circumstances. The simulation results presented here advance extant literature for the priority rules by applying a computer-aided, real-time, look-ahead parameter (Kemppainen, 2005) into a well-known priority rule. Warehouse practitioners can incorporate these real-time task-scheduling methods in their warehouse management system (WMS) to create and execute a string of order fulfilment jobs (Van den Berg, 1999; Ramaa et al., 2012).

The main contributions of this paper are as follows:

- A flow shop scheduling problem subject to the responsiveness in OEM warehouses is studied inspired by flow shop research in factory production schedule.
- A priority control with the real-time queue status is devised to deal with prevalent uncertainty in OEM warehouse.
- A decision-making framework for the job priority control is demonstrated by customising the warehouse's business requirements and cost perspectives.

The rest of this paper is structured as follows. Section 2 reviews literature related to responsiveness, warehousing and flow-shop methods. Section 3 describes the operational challenge of responsive order fulfilment for postponed cut-off times. Section 4 presents the priority rules and performance indicators. Section 5 shows simulation results for the case study, and Section 6 discusses some operational implications and conclusions.

2 Literature review

A brief review is given of literature related to the main aspects of the study, i.e., responsiveness, warehouses job schedule, priority-based job scheduling, and selecting criteria in warehouses.

Consumers can nowadays easily use the Internet to compare quality and prices of products across different suppliers. The offered service level remains the primary competitive quality, and warehouse clients perceive responsiveness mainly by the speed of delivery. Shaw et al. (2002) defined a clear hierarchy among the concepts of agility, responsiveness and flexibility. Agility concerns talents for operating 'profitably in a competitive environment of continually, and unpredictably, changing customer opportunities'. It involves both proactive initiatives and reactive responsiveness, and flexibility is one of the conditions enabling responsiveness. The study of Kritchanchai and MacCarthy (1999) identified four factors that determine responsiveness: stimuli, awareness, capabilities, and goals. In our OEM warehouse study, these factors consist respectively of hourly varying demand stimuli, awareness of demand fluctuations, job scheduling opportunities, and the goal of efficient order fulfilment. Giannakis and Luis (2016) also built a conceptual dimension of agility, responsiveness and flexibility. They denote three facets of responsiveness: visibility, rapid detection and reaction, cycle time reduction. Our paper aims to reduce order fulfilment time by adopting a priority rule which rapidly detects external circumstance (i.e. real-time queue status)

Since flow shop research (Johnson,1954) has attracted extensive attention from researchers and practitioner, they still have focused for improving performance (i.e. makespan, idle time) of production line (Gyorgyi and Kis, 2018; Nesello,2018; Yazdani et al., 2017). The labour intensive pallet packing lanes of warehouses are akin to factory workstations or job shops in manufacturing. However, the process of warehouse has more uncertainty in service time and job arrival than the process of factory where follows strictly controlled planning for resources. Whereas unexpected breakdowns of machines (Cui et al., 2018) are considered as one cause of uncertainty in manufacturing, uncertainty is prevalent at warehouses where most of process is still carried by human. Warehouse efficiency studies to overcome the uncertainty focussed mainly on picking strategies (Jarvis and McDowell, 1991; Hall, 1993; Petersen, 1997; Roodbergen and De Koster, 2001; Petersen et al., 2004; De Koster et al., 2007; Chen et al., 2010; De Koster et al., 2012). Proposed strategies include interleaving put-away and picking (Graves et al., 1977), wave picking (Petersen, 2000), and joint order batching (Won and Olafsson, 2005; Van Nieuwenhuyse and De Koster, 2009). The focus on picking is natural for retailer warehouses delivering directly to consumers, as such warehouses typically process large amounts of small orders for a wide variety of products by customer totes via multiple processing lines. Conversely, OEM warehouses delivering to retail warehouses process substantial orders for a comparatively narrow assortment by multiple pallets via few processing lines. The outbound operations constitute three stages: picking, packing and staging. Multiple orders from the same retailer are consolidated for single shipment, which requires customised re-palletising and packing to satisfy dimension restrictions of trucks and retailer warehouses. This makes packing by far the most labour intensive phase of the outbound process in OEM warehouses (Bartholdi and Hackman, 2011). To our best knowledge, this study pioneers the analysis of OEM warehouse outbound processes through job-scheduling methods using priority-dispatching rules to optimise responsiveness.

Pagh and Cooper (1998) offered an overview of postponement and speculation(P/S) strategies that manufacturing and logistics managers can take into consideration for timely and cost-effective operation. Leung et al.(2018) demonstrated a benefit of postponement strategy in e-order fulfilment process by intelligently grouping small lot-sized e-orders in distribution centres. However, cut-off rules such as when to group or release are still open to a manual decision due to various consideration points of business. This study gives guidance to postpone order cut-off times to obtain better responsiveness in terms of flexibility. The cut-off rules induce order peaks just before the cut-off time, thus causing imbalanced workloads. Huang et al. (2006) showed that these imbalances could lead to the 'self-contradiction of hands shortage and idleness' within the day. Such tightly available order fulfilment time from imbalances lead to job prioritising according to rules which optimise objectives of operations. To prioritize order fulfilment process, Kiran and Smith (1984) demonstrated dynamic job-shop scheduling by computer simulation, Lee et al. (1997) incorporated machine learning techniques, Shahrabi et al. (2017) enhanced the performance of the scheduling methods by using reinforcement learning and Freiheit and Wei (2016) conducted a case study to investigate imbalance effects on flow-shop performance. Kemppainen (2005) presented an extensive comparison of various priority scheduling rules and their use in integrated order management. These rules were classified into three groups based on the information used: static (fixed on entry), dynamic (updated by stage), and look-ahead (adapted by probing). First one uses static information such as due date and processing time. Second one uses dynamic information depending the status of order and system such as slack. Third one uses look-ahead parameter by forecasting based on historical data. The advance of IT (i.e. WMS) enables real-time job instruction to individual labourers. This paper investigates dynamic prioritising rules that incorporate dynamic factors and look-ahead factors into labour task instructions.

The benefits of priority-based job scheduling can be evaluated in terms of operational and financial performance criteria. The choice of which priority rule to employ involves a trade-off among multiple performance attributes of the outcomes, for example, handled volume, service level and operational cost (Chen et al., 2010). A popular method to assist this choice is data envelopment analysis (Hackman et al., 2001; De Koster and Balk, 2008). Treleven and Elvers (1985) assessed performance in terms of mean queuing times, mean earliness and percentage of late jobs. Ramasesh (1990) categorised performance in terms of idle machines, stalled promises, work-in-process inventories, and average value added in the queue. Although contract terms often involve earliness and tardiness penalties (Baker and Scudder, 1990; Elsayed et al., 1993), Vincent (2011) noted that most production cost models neglected just-in-time principles. This study incorporates them 'en bloc' since warehouses face penalties both for tardiness because they have to meet carrier schedules and for earliness because pallets staged for loading occupy costly storage space. Cakici et al. (2012) theoretically demonstrated an approach to provide Pareto-optimal solution for a multi-objective supply chain scheduling problem which is regarded as NP-hard. However, such an exact dynamic algorithm can only be available in small-sized problems whereas most of the real-operation case has multiple objects and uncertainty. Thus, either heuristics or metaheuristics are focused by researchers (Deng and Gu, 2012; Wang and Liu, 2013; and

Kheirandish, Tavakkoli-Moghaddam, and Karimi-Nasab, 2015). This paper aims to deliver a stochastic decision system for the real-life environment rather than an exact solution which is only available with the unrealistic experimental environment. Therefore, practitioner can select a scheduling rule that is stochastically evidenced as a best-fit rule for their operational parameter and objective.

3 Model and case study

The research question of central interest is how job priority scheduling can help OEM warehouses to improve their responsiveness to meet current trends of postponed daily order cut-off times for next-day delivery. As customers adapt their ordering policy by spiking demand briefly before the cut-off time, warehouses are confronted with order peaks that have to be processed faster when response times become shorter. OEM warehouses usually dispatch retailer orders by trucks on agreed pick-up times on the next working day. These pick-up times are spread across the day so that incoming orders have different due times that help job prioritisation. As suggested by Van den Berg (2007), workload imbalances can be alleviated by distinguishing can-ship orders from must-ship orders and by shifting the former from busier to quieter hours. Therefore, instead of processing orders on an FCFS basis, the workflow can be balanced by postponing less pressing jobs that have relatively late due times. Balancing the workload has several operational advantages, including reduced overtime and absenteeism reported in the empirical study of De Leeuw and Wiers (2015). The balancing effect of job priority management is illustrated graphically in Figure 1 Through postponing part of the jobs, stemming from demand peaks, the hourly workload becomes smoother with less peaks and troughs compared to FCFS scheduling.

<< Insert Figure 1 about here. >>

Ideally, the workload should be constant across the day as this dramatically simplifies warehouse planning and operation. The incoming order arrival process is irregular so that this ideal situation cannot be achieved in reality. The performance of alternative scheduling strategies is investigated by a simulation study based on actual operational data of a case study warehouse. The methodology to improve order fulfilment responsiveness for postponed cut-off times consists of four steps:

(1) Building a stochastic simulation model of order fulfilment that includes the following operational aspects: arrival distributions, order peaks, due time distribution, service time distributions per operation, and a set of priority rules to schedule remaining jobs for each queue.

(2) Constructing a cost objective function that incorporates penalties for earliness, tardiness, idleness, and work-in-process stock.

(3) Simulating the stochastic model under various cut-off scenarios and determine the costs resulting from each priority rule.

(4) Evaluating the relative performance of these priority rules for the various scenarios and identify which rule performs best under which circumstances.

For the case study warehouse, the simulation model of step (1) above has the following characteristics. The order fulfilment process is modelled as a tandem queue (Burke, 1956) with three service stages: picking, where a pallet or box is moved from storage to the packing lane; packing, where pallets are cubed according to customer requests; and staging, where pallets are moved from the packing lane to the staging zone. The definition of the tandem queue is satisfied here such as a finite chain of queues where each job must visit each queue in order. Figure 2 illustrates this tandem queuing process, where the three stages are linked without diversion and each stage consists of a set of servers with queues of unlimited capacity. Picking and staging are carried out by a single worker with automatized pallet handling equipment. Packing is the most labour intensive manual stage, with a group of workers per pallet. Packers perform re-palletising and wrapping tasks to satisfy customer warehouse pallet size restrictions; they check that orders cubed as one pallet are complete, and they register it into WMS before staging.

<< Insert Figure 2 about here. >>

As order arrival rates vary over the day, the arrival process at the picking stage is modelled as a non-homogeneous Poisson process with time-varying mean value for rates per hour of the day. Service times are modelled by simple exponential distributions with rates that differ for each of the three services of picking, packing and staging. The service rates for picking and packing depend on the customer order structure, with a distinction between relatively simple single-item pallets (SIP) with faster rates (pallet per hour) and complex multi-item pallets (MIP) with slower rates (pallet per hour). When the customer order is a single-line order, a pallet picker needs to visit only one location, and packers need simple cubing without complex stacking patterns. In contrast, if the customer order is a multi-line order, a pallet picker has to visit several locations for each item with a pick cage, and packers also should consider complex stacking patterns to optimise into a pallet. However, staging is the same for single-line and multi-line orders as it only involves moving a finished pallet from the packing area to the staging area. For given service and order type, the service rate is assumed to be constant per worker and per hour of the day. This assumption ignores ergonomic factors like fatigue, but the warehouse employs a refined measurement system for labour productivity per task per worker that indicates that this simplification is not unreasonable. All workers are directed independently via WMS instructions transmitted by hand-held terminals and they work per pallet without any knowledge of job priorities or shipment structures. The picking process is modelled as an M(t)/M/c queue with non-homogeneous Poisson arrival process(M(t)), packing follows a G/M/c queuing model with arrivals determined by departures from upstream picking, and staging also follows a G/M/c queuing model with arrivals determined by upstream packing. The final phase of the order fulfilment process involves waiting, and the waiting time of pallets is defined as the length of time they stay at the staging zone after packing and before shipping. All the above models have job service times according to an each an exponential distribution(M).

Historical warehouse operational data are used to specify the simulation input parameters for hourly arrival rates (17, one for each hour of the working day from 6 am until 11 pm), service rates (6, one for SIP and one for MIP for picking, packing and staging), and the mix of SIP and MIP orders (with probability 0.77 for SIP and 0.23 for MIP). Due times are uniformly distributed over the 17 hours of the next working day, because the warehouse takes the lead in setting due times due to agreements with carriers to spread truck pick-up optimally during the day. Multiple orders from the same client are consolidated and have the same due time to reduce transport costs.

4 Priority rules and performance criteria

The literature review mentioned some well-known priority rules for job scheduling from flow-shop production theory, which will now be described in more detail. The most straightforward rule is first-come-first-served (FCFS), where jobs that arrive earlier get higher priority. The so-called earliest due date (EDD) rule gives higher priority to jobs with earlier due time. Jackson (1955) proposed this priority rule and showed that it minimises the maximum of job tardiness. In this thesis OEM warehouse case study, the operational due time of dispatch by the carrier is already assigned upon arrival of the order owing to pre-arrangements with the retailers placing the orders. Smith (1956) proposed an alternative priority rule where jobs with shortest processing time (SPT) get the highest priority to get minimal mean flow time, that is, minimal work-in-process inventories. This result is related to Little's law (Little, 1961), which states that in steady state the mean number of units in the system (L) equals the product of the mean arrival rate (λ) and the meantime the unit spent in the system (W), so that $L = \lambda \times W$. An opposite rule gives the highest priority to jobs with the longest processing time (LPT). In our case study, processing times are defined in terms of the expected total service time of all remaining operations, i.e., picking plus packing plus staging for the picking queue; packing plus staging for the packing queue; and staging for the staging queue.

EDD and SPT focus on tardiness performance, but earliness and post-completion costs are also relevant. Berry and Rao (1975) studied the slack time (SLACK) and the critical ratio (CR) rules to improve inventory performance. For given time (t), the slack time (St) of a job with due time (D) is defined as the difference between remaining time ($Dt = D - t$) and (expected) remaining processing time (Pt) with correction factor ($z > 1$) to account for expected queuing and other time losses in the process, so that $St = Dt - z \times Pt$. SLACK gives higher priority to jobs with less slack time and constitutes a trade-off between EDD and LPT, as it assigns a higher priority to jobs with earlier due times that take longer to process. Berry and Rao (1975) showed that this rule averts both inventory surpluses from early replenishment and inventory shortages from late supplier deliveries. Similar to EDD and SPT, the SLACK priority of a job is static in the sense that all priority parameters (due times and expected remaining processing times) are known upon arrival. CR is a dynamic rule and replaces the correcting factor (z) by the expected queuing times that apply during dynamic operation. This rule assigns the highest priority to the job with the smallest value of remaining time until due time (Dt = D – t) divided by the sum of expected remaining processing time (Pt) and currently expected remaining queuing time (Qt), that is, (D – t)/(Pt + Qt). Here Pt depends on the stage of the job; for example, at the packing stage, it involves the expected service times of packing and staging. Qt depends not only on the stage of the job but also on the queues it should still pass. These queues are dynamic, and Qt depends on the expected processing times of all unfinished jobs with higher priority. Putnam et al. (1971) reported that the CR rule reduces uncertainty by trimming tardiness variance. In general, CR is expected to perform better than SLACK because it employs relevant extra dynamic information.

Table 1 provides a summary of the considered priority rules. EDD and SLACK reduce tardiness but may result in longer flow times than the alternatives. SPT and CR aim for short flow times but often lead or lag due dates with resulting weaker just-in-time and tardiness performance. Both SLACK and CR leverage processing times to account for other factors. CR provides dynamic corrections by means of "live" waiting times and is therefore expected to give shorter flow times than SLACK.

<< Insert Table 1 about here. >>

Next, the performance criteria to evaluate OEM warehouse operations is considered. The warehouse outcomes are evaluated in terms of a joint cost criterion that integrates the four objectives of low earliness, low tardiness, low labour idleness, and low work-in-process stocks. The weight of each objective is determined by the associated penalty for failing to reach it, and this cost structure will be case dependent. The cost criterion function for fulfilling a set of orders is given by

$$\text{Cost} = \sum_{i=1}^{n}(w_1 \times \alpha_i + w_2 \times \beta_i) + w_3 \times \gamma + w_4 \times \delta.$$

Here the symbols have the following meaning: 'i' denotes the order; 'n' is the total number of orders; 'α_i' is the earliness cost of job 'i' and involves space costs at the staging zone for awaiting pick-up; 'β_i' is the tardiness cost of the job and consists of demurrage costs for carriers from appointed pick-up time until actual dispatch time; 'γ' is the total idleness cost, the sum total of idle labour costs in the phases of picking, packing and staging; 'δ' is total work-in-process cost, the sum over all 'n' jobs of financial costs from work-in-process inventories during picking, packing, and staging; and 'w_j' (j=1,2,3,4) are selection weights that determine which objectives are incorporated (1 if yes and 0 if no), depending on the business environment. However, this paper has a limitation not to give a value between 0 and 1 which can be topic of future research.

The four objectives and expected performance of alternative priority rules are summarised in Table 2 Earliness penalties favour just-in-time strategies like SLACK by reducing staging buffer space, whereas CR and SPT exacerbate these penalties because of their shorter flow times. Tardiness penalties favour strategies like EDD that aims at early completion. Even though CR and SPT have shorter flow times, they tend to generate some very late jobs with mainly associated tardiness penalties. If favourable business relationships between warehouses and truckers allow rescheduling appointments without cost, then the tardiness penalty may be waived (w_2=0). Idleness and stock penalties, which are linked since curtailed stock-in-process requires less labour, are related to lean production principles (Krafcik, 1988). The law $L = \lambda \times W$ of Little (1961) implies that work-in-process inventories (L) and associated stock penalties are proportional to flow time (W), so that CR and SPT are expected to perform well in this respect. However, if handled products are relatively cheap so that inventory costs are negligible, then stock penalties could be discarded (w_4=0).

<< Insert Table 2 about here. >>

5 Simulation results

The cost performance of alternative job priority rules is investigated by a simulation study, with parameters derived from a case study OEM retail distribution centre of a multinational consumer electronics manufacturer. Figure 3 summarises the interactions of this distribution centre with its manufacturer, sales department, retail warehouses and shops, carriers, and labour provider. The order arrival process is determined by the sales department, and due times for order fulfilment are agreed with carriers.

<< Insert Figure 3 about here. >>

The main question of interest is how to improve responsiveness for postponed daily order cut-off times. Curve A in Figure 4 shows the historical hourly average order pattern for 2012-2014, with a steep demand peak just before the order cut-off time that was fixed at 2 pm during that period. The simulation study considers postponed cut-off scenarios with cut-off time at 3 pm (B), 4 pm (C), or 5 pm (D) but keeps the same due times for all orders. The corresponding demand patterns are simply extrapolated by shifting the base scenario (A) forwards in time while keeping the size of demand peaks and daily totals fixed.

<< Insert Figure 4 about here. >>

Table 3 summarises the input parameters for the simulations derived from historical operational data of the case study warehouse. The sales order desk is open from 8 am until 6 pm, and orders rarely arrive outside these hours, resulting in relatively small means and large standard deviations of arrivals for out-of-office hours. Order arrivals have a 77% chance to be SIP and a 23% to be MIP, and service rates for SIP are higher than those for MIP by factors 2.83 for picking and 1.34 for packing. Weekly idleness costs are obtained by multiplying the average non-utilisation ratio by the weekly sum of total labour costs of €21.93 per hour. The stock-carrying cost of €10.14 per pallet per week for work-in-process stocks is derived from stock value and interest costs. The staging zone space cost of €6.96 per pallet per week is used as earliness penalty because this area can be used flexibly for extra bulk storage during peaks. Time criticality of order fulfilment for this warehouse is shown by high demurrage costs of €75.00 per pallet per hour. Finally, for the correction factor z in the definition of slack ($St = Dt - z \times Pt$) we choose the same value (20) as in the pilot study of FCFS by Kanet and Hayya (1982) to correct machine processing time for queuing times. The average total processing time is 0.197 hours ($1/12.94 + 1/9.40 + 1/73.13$) for SIP and 0.376 hours ($1/4.57 + 1/6.99 + 1/73.12$) for MIP. This corresponds (for $z = 20$) to average fulfilment durations of $20 \times 0.197 = 3.9$ hours for SIP and $20 \times 0.376 = 7.5$ hours for MIP, which reasonably fits experiences in the case study warehouse.

<< Insert Table 3 about here. >>

This study uses the Matlab Simulink (2018) tool to build the simulation model. Every single simulation run corresponds to one week of warehouse operations with hourly order arrivals, order types, and order service times. A week consists of five days of 17 hours each (85 hours in total) with expected total arrival orders of around 3,200 pallets (average 630 pallets x 5days). To process the order arrivals, the simulation model uses 4 pickers, 5 packing lanes (with 4 persons per lane), and 1 stager for each of the four cut-off scenarios. One standard set of 1,000 simulation runs is employed to study the outcomes of the five considered priority rules for each of the four cut-off scenarios (A-D). To evaluate the mean value of each scenario by t-test with a normal distribution assumption, the replication of simulation should be more than 30. Each of these twenty scenarios is evaluated in terms of operational performance. To improve performance, this study applies the priority rules at each operation's queue separately, which is also called an operational due dates strategy (Kanet and Hayya, 1982). The flow time of a job is the total time it spends in the shop, that is, the time elapsing between arrival and completion. Earliness is defined as the difference between completion time and due time, so that negative values correspond to timely completion. For smooth operation, it is preferred to have not only small mean but also small variation of flow times and earliness, and therefore both the mean and the standard deviation of these two characteristics are considered across the set of jobs within a given simulation run, that is, a given week of warehouse operations. Tardiness occurs if earliness is positive, that is, if jobs are completed after the due time limit. Maximum tardiness is defined as the maximum value of (positive) earliness across all jobs within a given simulation run.

The operational outcomes of 1,000 simulation runs (weeks of order fulfilment) are summarised in Table 4 and Figure 5. Table 4 shows that postponed cut-off times lead, as expected, to shorter flow times, less earliness and more tardiness. FCFS does not perform well across all performance dimensions and has the worst tardiness outcomes, especially for tight cut-off scenarios. Of the five priority rules, CR performs the best in terms of flow time, whereas EDD and SLACK have excellent tardiness results as none of their jobs have positive earliness. Figure 5 shows some outlying tardiness results for CR, both in the benchmark cut-off scenario (A, 2 pm) and in the most ambitious scenario (D, 5 pm). Table 4 shows SLACK and EDD perform roughly similar, but because SLACK amplifies the weight of processing times it has highest earliness mean value and highest flow times mean value of all priority rules. Compared to these two methods, SPT has shorter flow times but more tardiness. The outcomes in Table 4 are in line with those in Table 1 because CR and SPT have shortest flow times, EDD and SLACK have lowest tardiness, and SLACK comes closest to just-in-time planning as it has highest earliness.

<< Insert Figure 5 about here. >>

<< Insert Table 4 about here. >>

Table 5 summarises the financial outcomes of the simulation experiments. These outcomes consist of costs associated with earliness, tardiness, idleness, and stock costs. This study considers an integrated cost function that includes all four cost components as well as two modified versions. One version excludes stock costs, which is relevant for warehouses at urban locations with just-in-time planning that have relatively low stock value compared to high storage rental costs. Another version excludes tardiness costs for warehouses that handle expensive goods with high storage rental costs and that have flexible pick-up agreements with carriers to skip tardiness penalties. EDD performs best if all components are included, SLACK is best if there are no stock costs, and CR is best if there are no tardiness costs. These rankings of priority rules do not depend on the cut-off scenario and get more pronounced for tighter scenarios. In scenario A (2 pm), the percentage of simulation runs for which EDD, SLACK and CR are optimal are respectively 46.5, 48.1, and 56.6, and for scenario D these percentages are respectively 93.7, 66.6, and 59.4. The outcomes illustrate that there is no priority rule that is universally best for all business situations, but each warehouse may find a suitable rule by selecting the performance objectives that apply for its specific situation.

<< Insert Table 5 about here. >>

As EDD and SLACK perform roughly similar, a more detailed comparison of these two rules is provided by means of paired t-tests (Welch, 1947) for operational and financial performance for the tightest cut-off scenario (D, 5 pm). The sample size of 1,000 runs far exceeds the usual rule-of-thumb threshold (30). Therefore the conventional standard normal distribution is employed to compute p-values. The results in Table 6 show significant differences between the two methods. In terms of operational performance, SLACK is more 'just in time', and EDD has shorter flow time. From a financial perspective, SLACK requires less staging space, but EDD has higher server utilisation and less work-in-process stocks. The two rules do not show significant differences in tardiness and associated demurrage costs.

<< Insert Table 6 about here. >>

6 Some operational implications and conclusions

In this analysis, performance is distinguished along four dimensions by preventing earliness (staging costs), tardiness (demurrage costs), idleness (labour costs), and work-in-process inventories (stock costs). It depends on the business environment which of these dimensions is relevant. Preventing tardiness, for example, is imperative if delayed delivery spoils all product virtues, whereas it is less relevant if delays can be solved by the penalty-free rescheduling of pick-up times. The latter situation often applies for OEM warehouses that deliver to retailer warehouses and shops. This study's simulation results show that the critical ratio (CR) priority rule performs well in such situations. It offers the shortest flow time with least work-in-process stock, which is valuable for businesses that handle expensive products with high labour costs. It shows a way how to implement data-driven strategies into a real application.

The case study warehouse currently uses the earliest due date (EDD) strategy for sequencing its order fulfilment jobs as it is easy to implement due to the static nature of the job priority (i.e. fixed on entry). The simulation results based on the warehouse-specific cost parameters indicate potential benefits of the CR rule. Compared to the other priority rules for which the priority is either fixed on entry or updated per stage, the CR has the unique property that it adapts priorities by probing the dynamic queuing status of jobs. For this purpose, the simulation study employs a data-driven estimate of queuing times based on expected processing times of jobs with higher priority. Studying actual workflow patterns could refine these estimates by queuing data from the warehouse process and by forecasting queuing times using statistical and machine learning methods. The case study warehouse currently considers incorporating more dynamic job scheduling strategies into task instruction module in WMS. The application of stochastic simulation models for uncertain real-life operational environments contributes to the previous literature on deterministic models for theoretical environments.

Summarising the contributions of this paper, the current retail market trend leverages responsiveness of order fulfilment and forces higher levels of fluidity in supply chains. To improve responsiveness, warehouses try to postpone the cut-off time while handling the same order volume with less slack. In such a compressing situation, Practitioners might waste labour resources if they could not improve data-driven decision-making capabilities for organising fluidity of jobs processing. From this perspective, job scheduling using flow-shop priority rules offers solutions for distribution centres facing cut-off time pressures. By prioritising each job, warehouses can efficiently maintain responsiveness without increasing labour to satisfy compressed order-fulfilment deadlines. However, there is no research about the methodology practitioners can apply the job scheduling theory to their warehouses. The decision making in this paper means two-fold; one is which job should be prioritised among outstanding jobs. The other is which job scheduling rule should be adopted into a warehouse in practice. This paper presents a data-driven decision-making methodology for selecting priority rules by simulating alternative rules evaluating corporate cost objectives that can be tailored to warehouse-specific settings. In stochastic models built from real-life data on order arrivals, due dates, and service times, simulation results indicate good performance of the SLACK rule for just-in-time operations with high storage costs and of the CR rule for high-value product operations with flexible pick-up schedules. This computerised method specifically presents how to transfer the benefit of the academic decision-making theory from academics to practitioners.

Further research is needed to analyse the trade-off between potential revenue gains through better service with postponed cut-off times against increased costs due to tighter processing conditions. It is also of interest to study historical workflow patterns in more detail to refine CR-type priority rules by improving forecasts of remaining processing and queuing times.

References

Baker, K. R., & Scudder, G. D. (1990). Sequencing with earliness and tardiness penalties: a review. *Operations Research*, *38*(1), 22-36.

Barclay, I., Poolton, J., & Dann, Z. (1996). Improving competitive responsiveness via the virtual environment. In *Engineering and Technology Management, 1996. IEMC 96. Proceedings., International Conference on* (pp. 52-62). IEEE.

Bartholdi, J. J., Eisenstein, D. D., & Foley, R. D. (2001). Performance of bucket brigades when work is stochastic. *Operations Research*, *49*(5), 710-719.

Bartholdi, J. J., & Hackman, S. T. (2016). *Warehouse & Distribution Science*: Release 0.91. Retrieved from http://www.warehouse-science.com

Berry, W. L., & Rao, V. (1975). Critical ratio scheduling: an experimental analysis. *Management Science*, *22*(2), 192-201.

Burke, P. J. (1956). The output of a queuing system. *Operations Research*, *4*(6), 699-704.

Cakici, E., Mason, S. J., & Kurz, M. E. (2012). Multi-objective analysis of an integrated supply chain scheduling problem. *International Journal of Production Research*, *50*(10), 2624-2638.

Chen, C. M., Gong, Y., De Koster, M.B.M., & Van Nunen, J. A. (2010). A flexible evaluative framework for order picking systems. *Production and Operations Management*, *19*(1), 70-82.

Cui, W., Lu, Z., Li, C., & Han, X. (2018). A proactive approach to solve integrated production scheduling and maintenance planning problem in flow shops. *Computers & Industrial Engineering*, 115, 342-353.

De Koster, M.B.M., & Balk, B. M. (2008). Benchmarking and monitoring international warehouse operations in Europe. *Production and Operations Management*, *17*(2), 175-183.

De Koster, M.B.M., Le-Duc, T., & Roodbergen, K. J. (2007). Design and control of warehouse order picking: a literature review. *European Journal of Operational Research*, *182*(2), 481-501.

De Koster, M.B.M., Le-Duc, T., & Zaerpour, N. (2012). Determining the number of zones in a pick-and-sort order picking system. *International Journal of Production Research*, *50*(3), 757-771.

De Leeuw, S., & Wiers, V. C. (2015). Warehouse manpower planning strategies in times of financial crisis: evidence from logistics service providers and retailers in the Netherlands. *Production Planning & Control*, *26*(4), 328-337.

Deng, G., & Gu, X. (2012). A hybrid discrete differential evolution algorithm for the no-idle permutation flow shop scheduling problem with makespan criterion. *Computers & Operations Research*, 39(9), 2152-2160.

Denman, T. (2017). *The omni channel supply chain: From the docks to the doorstep.* Retrieved from http://risnews.com/omnichannel-supply-chain-docks-doorstep

Doerr, K. H., & Gue, K. R. (2013). A performance metric and goal‐setting procedure for deadline‐oriented processes. *Production and Operations Management*, *22*(3), 726-738.

Elsayed, E. A., Lee, M. K., Kim, S., & Scherer, E. (1993). Sequencing and batching procedures for minimizing earliness and tardiness penalty of order retrievals. *The International Journal of Production Research*, *31*(3), 727-738.

Freiheit, T., & Li, W. (2017). The effect of work content imbalance and its interaction with scheduling method on sequential flow line performance. *International Journal of Production Research*, *55*(10), 2791-2805.

Giannakis, M., & Louis, M. (2016). A multi-agent based system with big data processing for enhanced supply chain agility. *Journal of Enterprise Information Management*, 29(5), 706-727.

Gong, Y., & De Koster, M.B.M. (2011). A review on stochastic models and analysis of warehouse operations. *Logistics Research*, *3*(4), 191-205.

Graves, S. C., Hausman, W. H., & Schwarz, L. B. (1977). Storage-retrieval interleaving in automatic warehousing systems. *Management Science*, *23*(9), 935-945.

Györgyi, P., & Kis, T. (2018). Minimizing the maximum lateness on a single machine with raw material constraints by branch-and-cut. *Computers & Industrial Engineering*, 115, 220-225.

Hackman, S. T., Frazelle, E. H., Griffin, P. M., Griffin, S. O., & Vlasta, D. A. (2001). Benchmarking warehousing and distribution operations: an input-output approach. *Journal of Productivity Analysis*, *16*(1), 79-100.

Hall, R. W. (1993). Distance approximations for routing manual pickers in a warehouse. *IIE Transactions*, *25*(4), 76-87.

Huang, H. C., Lee, C., & Xu, Z. (2006). The workload balancing problem at aircargo terminals. *OR Spectrum*, *28*(4), 705-727.

Jackson, J.R. (1955). *Scheduling a production line to minimize maximum tardiness*. Los Angeles, CA: University of California Los Angeles.

Jarvis, J. M., & McDowell, E. D. (1991). Optimal product layout in an order picking warehouse. *IIE Transactions*, *23*(1), 93-102.

Johnson, S. M. (1954). Optimal two-and three-stage production schedules with setup times included. *Naval Research Logistics*, *1*(1), 61-68.

Kanet, J. J., & Hayya, J. C. (1982). Priority dispatching with operation due dates in a job shop. *Journal of Operations Management*, *2*(3), 167-175.

Kemppainen, K. (2005). *Priority Scheduling Revisited – Dominant Rules, Open Protocols, and Integrated Order Management*. PhD Thesis A-261, Helsinki School of Economics.

Kheirandish, O., Tavakkoli-Moghaddam, R., & Karimi-Nasab, M. (2015). An artificial bee colony algorithm for a two-stage hybrid flowshop scheduling problem with multilevel product structures and requirement operations. *International Journal of Computer Integrated Manufacturing*, 28(5), 437-450.

Kiran, A. S., & Smith, M. L. (1984). Simulation studies in job shop sheduling—I a survey. *Computers & Industrial Engineering*, *8*(2), 87-93.

Krafcik, J. F. (1988). Triumph of the lean production system. *Sloan Management Review*, *30*(1), 41.

Kritchanchai, D., & MacCarthy, B. L. (1999). Responsiveness of the order fulfilment process. *International Journal of Operations & Production Management*, *19*(8), 812-833.

Lee, C. Y., Piramuthu, S., & Tsai, Y. K. (1997). Job shop scheduling with a genetic algorithm and machine learning. *International Journal of Production Research*, *35*(4), 1171-1191.

Leung, K. H., Choy, K. L., Siu, P. K., Ho, G. T., Lam, H. Y., & Lee, C. K. (2018). A B2C e-commerce intelligent system for re-engineering the e-order fulfilment process. *Expert Systems with Applications*, 91, 386-401.

Little, J.D.C. (1961). A proof for the queuing formula: L=λW. *Operations Research*, 9 (3), 383–387.

Nesello, V., Delorme, M., Iori, M., & Subramanian, A. (2018). Mathematical models and decomposition algorithms for makespan minimization in plastic rolls production. *Journal of the Operational Research Society*, 69(3), 326-339.

Pagh, J. D., & Cooper, M. C. (1998). Supply chain postponement and speculation strategies: how to choose the right strategy. *Journal of Business Logistics*, *19*(2), 13.

Pearcy, M., & B. Kerr. (2013). *From customer orders through fulfilment: Challenges and opportunities in manufacturing, high-tech and retail*. Retrieved from http://www.fi.capgemini.com/from-customer-orders-through-fulfilment-challenges-and-opportunities

Petersen, C. G. (1997). An evaluation of order picking routeing policies. *International Journal of Operations & Production Management*, *17*(11), 1098-1111.

Petersen, C.G. (2000). An evaluation of order picking policies for mail order companies. *Production and Operations Management*, *9*(4), 319-335.

Petersen, C. G., Aase, G. R., & Heiser, D. R. (2004). Improving order-picking performance through the implementation of class-based storage. *International Journal of Physical Distribution & Logistics Management*, *34*(7), 534-544.

Putnam, A. O., Everdell, R., Dorman, G. H., Cronan, R. R., & Lindgren, L. H. (1971). Updating critical ratio and slack time priority scheduling rules. *Production and Inventory Management*, *12*(4), 51-72.

Ramaa, A., Subramanya, K. N., & Rangaswamy, T. M. (2012). Impact of warehouse management system in a supply chain. *International Journal of Computer Applications*, *54*(1).

Ramasesh, R. (1990). Dynamic job shop scheduling: a survey of simulation research. *Omega*, *18*(1), 43-57.

Roodbergen, K. J., & Koster, M.B.M. (2001). Routing methods for warehouses with multiple cross aisles. *International Journal of Production Research*, *39*(9), 1865-1883.

Shahrabi, J., Adibi, M. A., & Mahootchi, M. (2017). A reinforcement learning approach to parameter estimation in dynamic job shop scheduling. *Computers & Industrial Engineering*, 110, 75-82.

Shaw, A., McFarlane, D. C., Chang, Y. S., & Noury, P. J. G. (2003). Measuring response capabilities in the order fulfillment process. *Proceedings of EUROMA, Como, Italy*.

Smith, W. E. (1956). Various optimizers for single-stage production. *Naval Research Logistics*, *3*(1-2), 59-66.

Treleven, M. D., & Elvers, D. A. (1985). An investigation of labor assignment rules in a dual-constrained job shop. *Journal of Operations Management*, *6*(1), 51-68.

Van den Berg, J. P. (1999). A literature survey on planning and control of warehousing systems. *IIE Transactions*, *31*(8), 751-762.

Van den Berg, J.P. (2007). *Integral warehouse management*. Utrecht : Netherlands Management Outlook Publications.

Van Nieuwenhuyse, I., & de Koster, M.B.M. (2009). Evaluating order throughput time in 2-block warehouses with time window batching. *International Journal of Production Economics*, *121*(2), 654-664.

Vincent, T. K. (2011). Multicriteria models for just-in-time scheduling. *International Journal of Production Research*, *49*(11), 3191-3209.

Vincent, T. K. & Billaut, J. C. (2006). *Multicriteria scheduling: Theory, models and algorithms*. Berlin : Springer Science & Business Media.

Wang, S., & Liu, M. (2014). Two-stage hybrid flow shop scheduling with preventive maintenance using multi-objective tabu search method. *International Journal of Production Research*, 52(5), 1495-1508.

Welch, B. L. (1947). The generalization of 'Student''s problem when several different population variances are involved. *Biometrika*, *34*(1/2), 28-35.

Won, J., & Olafsson, S. (2005). Joint order batching and order picking in warehouse operations. *International Journal of Production Research*, *43*(7), 1427-1442.

Yazdani, M., Aleti, A., Khalili, S. M., & Jolai, F. (2017). Optimizing the sum of maximum earliness and tardiness of the job shop scheduling problem. *Computers & Industrial Engineering*, 107, 12-24.

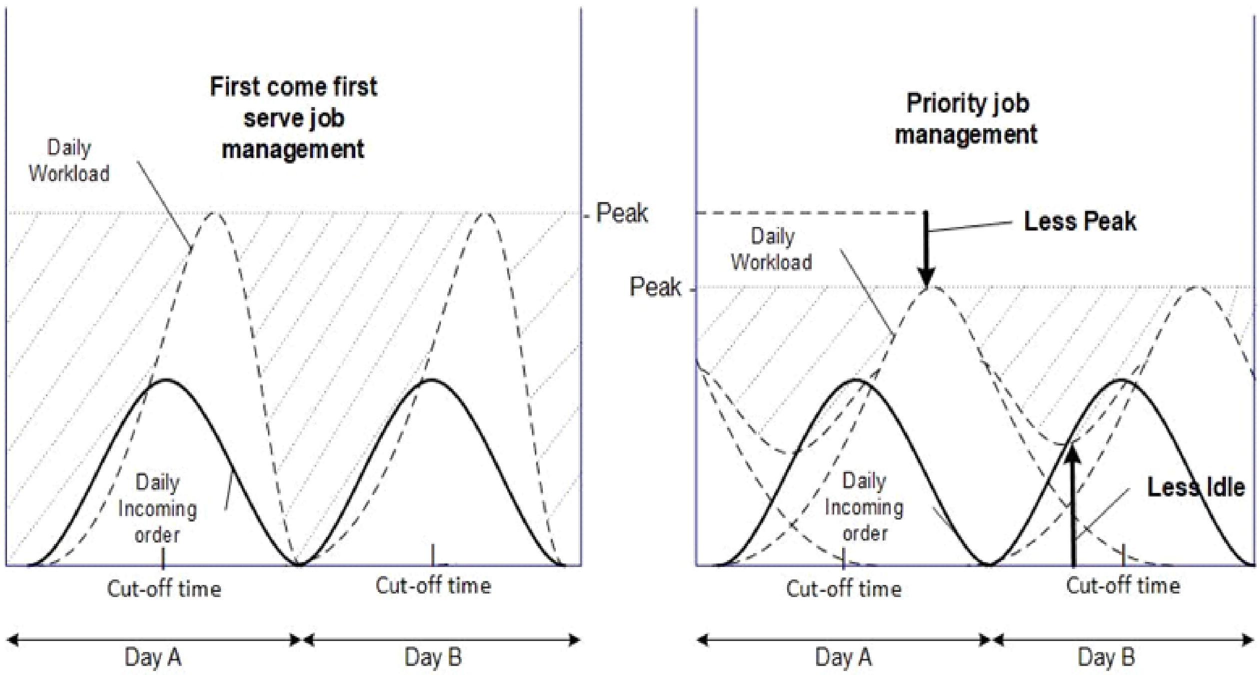

Figure 1: Daily incoming orders and two job management methods.

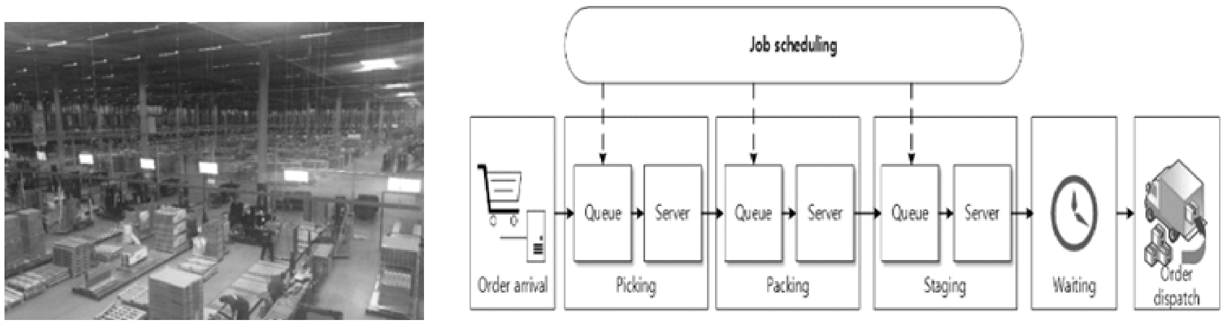

Figure 2: Actual warehouse process (left) and queuing model (right).

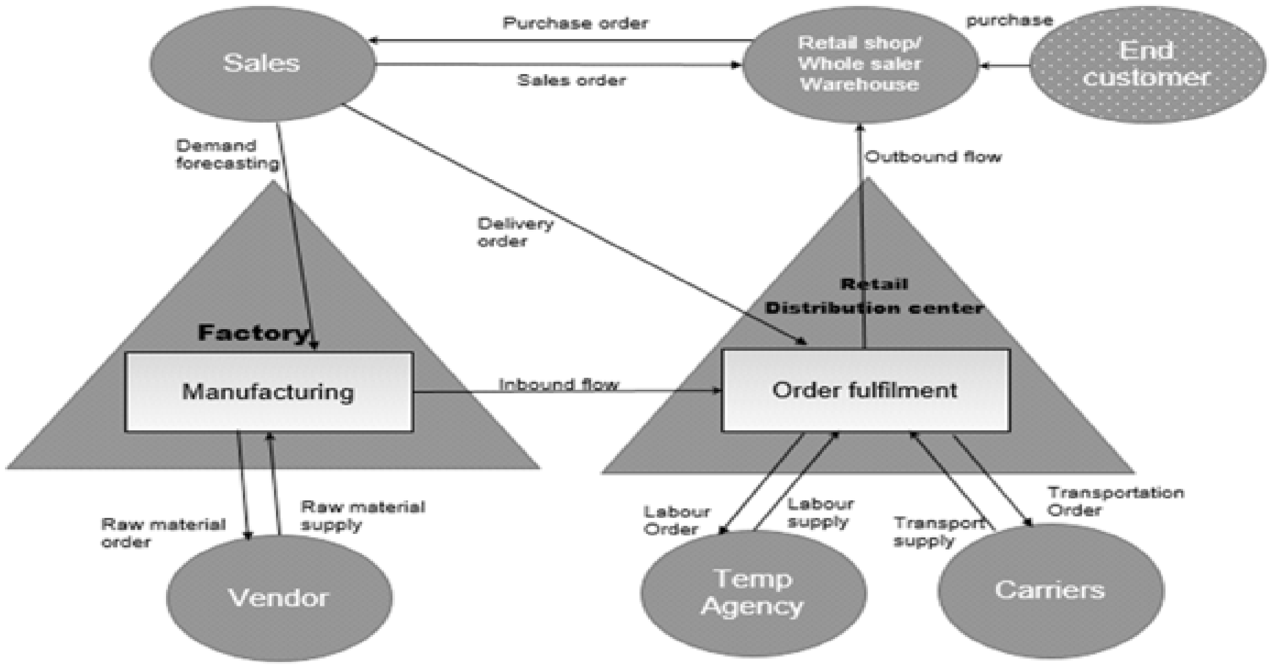

Figure 3: Retail distribution centre (OEM warehouse) and SCM partners.

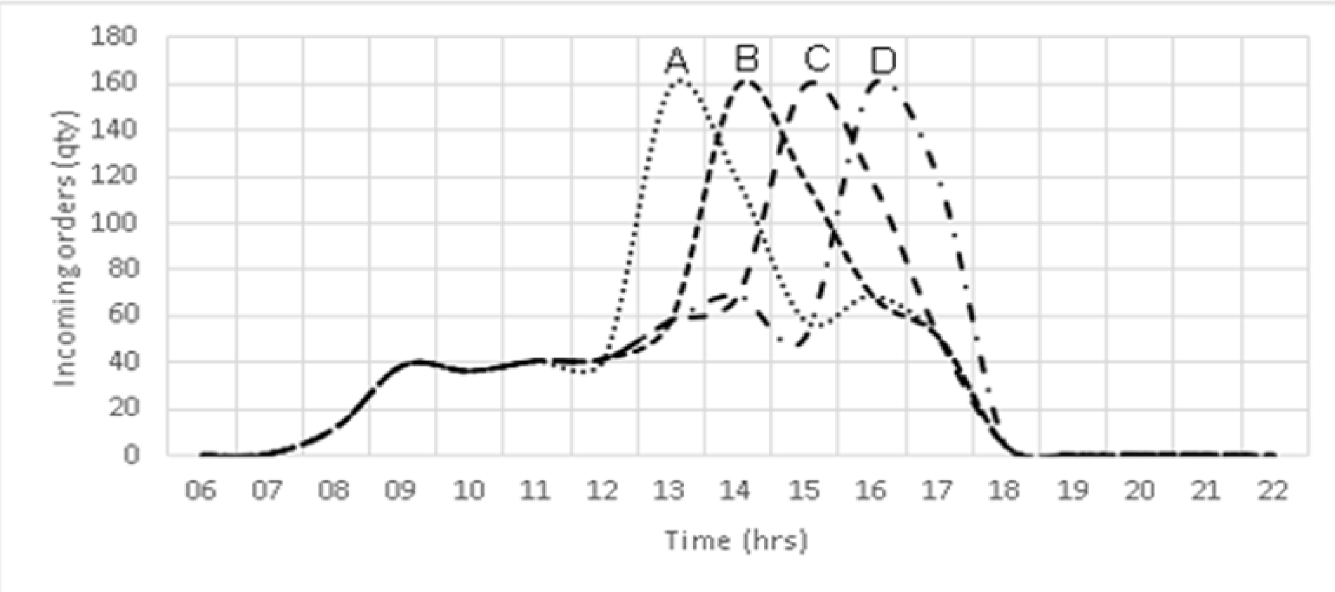

Figure 4: Average hourly incoming orders per for four cut-off scenarios (current is A).

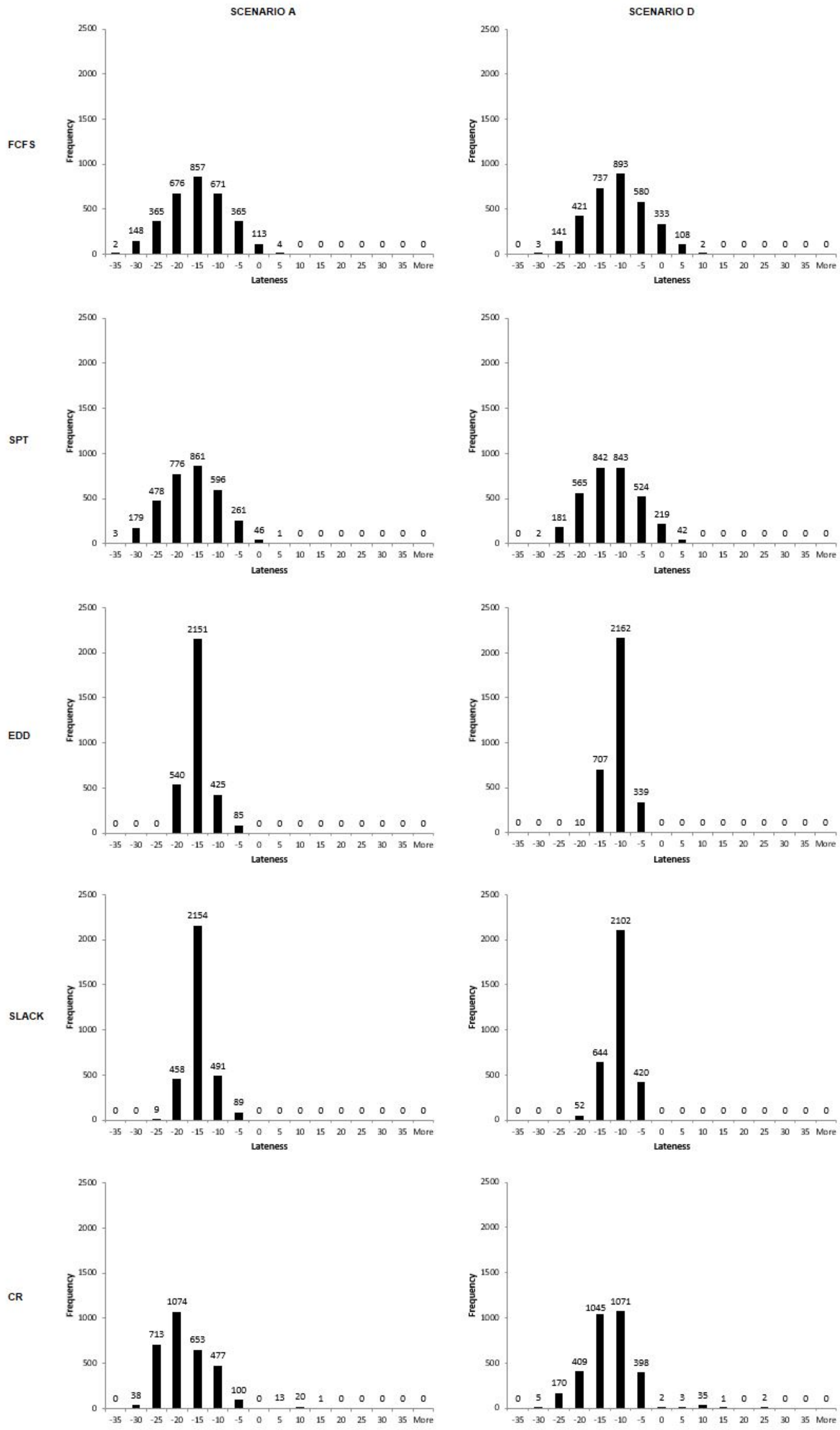

Figure 5: Histograms of simulated outcomes for Earliness (0 on horizontal axis means -4.99 to 0.00)

Priority rule	Source	Performance objectives			
		Low tardiness	Short flow time	Just-in-time	Dynamic
First-come first-served (FCFS)	--	o	o	o	o
Earliest due date (EDD)	Jackson (1955)	+	o	o	o
Shortest processing time (SPT)	Smith (1956)	-	+	-	o
Minimum slack (SLACK)	Berry and Rao (1975)	+	-	+	o
Critial Ratio (CR)	Putnam et al. (1971)	-	+	-	+

Table notes

For each rule, + means advantage, - disadvantage, and o neutral performance for the objective.

Table 1: Performance of five priority rules for a set of four responsiveness goals.

Penalty	Operations	Objective	Penalty Calculation			Priority rule	
			Cost Driver	Count	Unit cost	Advantage	Disadvantage
Earliness	Staging, appointment	Just-in-time	Staging stocks	Max	Storage cost (€ per pallet per week)	SLACK	CR / SPT
Tardiness	Appointment, dispatch	Early in time	Late hours	Sum	Demurrage cost (€ per pallet per hour)	EDD	CR / SPT
Idleness	Picking, packing, staging	Short flow time	Idle hours	Sum	Labour cost (€ per hour)	CR / SPT	SLACK
Stock	Picking, packing, staging	Short flow time	Work-in-process inventory	Average	Inventory value (€ per pallet per week)	CR / SPT	SLACK

Table notes

The sequence of operations consists of picking, packing, staging, appointment, and dispatch.

Table 2: Performance of various priority rules among four cost dimensions.

Parameter		Unit	Specification	Value
Arrival rate		Pallets per hour	6-7 am	0.01 / 0.41
			7-8 am	0.85 / 20.86
			8-9 am	12.00 / 30.59
			9-10 am	38.82 / 54.65
			10-11 am	36.23 / 50.74
			11-12 am	40.70 / 57.53
			12-1 pm	41.46 / 58.94
			1-2 pm (cut-off)	158.84 / 116.53
			2-3 pm	118.00 / 142.31
			3-4 pm	57.88 / 71.95
			4-5 pm	68.68 / 86.70
			5-6 pm	50.64 / 84.02
			6-7 pm	3.94 / 21.62
			7-8 pm	0.34 / 5.17
			8-9 pm	0.53 / 7.94
			9-10 pm	0.43 / 6.90
			10-11 pm	0.01 / 0.33
Service rate	Picking	Pallets per hour per server	SIP	12.94
			MIP	4.57
	Packing	Pallets per hour per lane	SIP	9.40
			MIP	6.99
	Staging	Pallets per hour per server	SIP	73.13
			MIP	73.13
Penalty	Earliness	€ per pallet per week	storage cost	6.96
	Tardiness	€ per pallet per hour	demurrage cost	75.00
	Idleness	€ per hour	labour cost	21.93
	Stock	€ per pallet per week	work-in-process stock	10.14
	Queuing	Scalar value z	Slack = Dt - z×Pt	z = 20

Table notes

SIP and MIP denote respectively single-item pallets (77%) and multi-item pallets (23%).

Reported values are mean and standard deviation for arrival rates, mean for service rates, and financial penalty costs in terms of prime interest rates published by The Wall Street Journal for December 2016.

Table 3: Operational parameters for the case study warehouse (scenario A).

		Flow time				Lateness				Tardiness			
		Mean		Standard dev.		Mean		Standard dev.		Maximum		Fraction (%)	
Cut-off	Priority	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std
2 pm	FCFS	8.7	2.3	2.9	0.2	-16.6	2.3	7.2	0.2	2.6	2.5	1.1	1.5
	SPT	7.3	2.1	5.0	0.7	-18.0	2.1	6.9	0.2	2.3	2.4	0.6	0.8
	EDD	8.7	2.3	6.9	1.0	-16.6	2.3	3.2	1.0	0.0	0.0	0.0	0.0
	SLACK	9.3	2.4	7.2	0.9	-16.0	2.4	3.3	0.9	0.0	0.0	0.0	0.0
	CR	4.8	0.5	6.9	2.2	-19.9	0.6	6.8	1.5	33.7	24.4	1.0	0.8
3 pm	FCFS	8.5	2.2	2.9	0.2	-14.9	2.2	7.2	0.3	3.9	2.8	2.2	2.5
	SPT	7.1	2.0	5.2	0.7	-16.4	2.0	6.9	0.2	3.6	2.7	1.1	1.4
	EDD	8.5	2.2	7.0	0.9	-15.0	2.2	3.0	1.1	0.0	0.0	0.0	0.0
	SLACK	9.2	2.3	7.3	0.9	-14.3	2.3	3.2	0.9	0.0	0.0	0.0	0.0
	CR	4.5	0.6	6.9	2.3	-18.3	0.6	6.7	1.5	35.2	23.2	1.1	0.8
4 pm	FCFS	8.4	2.3	2.8	0.2	-13.0	2.3	7.3	0.2	5.8	2.9	4.5	4.0
	SPT	6.9	2.1	5.1	0.6	-14.5	2.2	6.9	0.2	5.5	2.9	2.4	2.4
	EDD	8.4	2.3	7.0	0.9	-13.1	2.3	2.8	1.1	0.0	0.0	0.0	0.0
	SLACK	9.0	2.4	7.2	0.8	-12.4	2.4	3.0	0.9	0.0	0.0	0.0	0.0
	CR	4.3	0.6	6.8	2.4	-16.4	0.7	6.6	1.6	36.0	25.5	1.1	0.9
5 pm	FCFS	8.1	2.4	2.7	0.3	-11.7	2.4	7.2	0.2	7.0	3.1	6.8	5.2
	SPT	6.6	2.3	4.8	0.5	-13.2	2.3	6.9	0.2	6.7	3.1	3.9	3.3
	EDD	8.1	2.4	6.7	0.8	-11.7	2.4	2.7	1.1	0.0	0.0	0.0	0.0
	SLACK	8.8	2.5	7.0	0.7	-11.0	2.5	2.9	0.9	0.0	0.0	0.0	0.0
	CR	4.1	0.6	7.0	2.4	-15.0	0.7	6.5	1.6	37.7	25.2	1.2	0.8

Table notes

Underscored mean values are for the best performing priority rule per objective and per cut-off scenario. Flow time, Earliness and tardiness are measured in hours, and fraction of tardiness is measured as percentage.

The standard deviation columns (std) show the variation of outcomes across the 1,000 simulation runs.

Table 4: Simulated performance of five priority methods.

Objective		Earliness (α)		Tardiness (β)		Idleness (γ)				Stock (δ)	Cost specification								
Measurement		Max. stage		Truck penalty		Pick	Pack	Stage	Total	Average	All four (α + β + γ + δ)			No stock cost (α + β + γ)			No tardiness cost (α + γ + δ)		
Unit		Pallet		€		Utilization (%)				Pallet	€			€			€		
Cut-off	Priority	mean	std	mean	std						mean	std	best (%)	mean	std	best (%)	mean	std	best (%)
2 pm	FCFS	687	80	4,708	8,242	97.2	81.2	48.2	82.4	327.4	22,056	11,546	0.1	18,657	10,882	9.4	16,323	890	0.1
	SPT	712	82	2,350	4,220	97.2	81.7	48.6	82.8	274.9	18,687	6,396	46.6	15,831	5,807	2.7	15,766	794	43.3
	EDD	687	81	0.00	0.00	97.2	81.2	48.2	82.4	327.0	16,324	893	46.5	12,927	557	39.7	16,324	893	0.0
	SLACK	668	86	0.00	0.00	97.2	80.9	48.0	82.2	351.5	16,571	916	0.0	12,922	570	48.1	16,571	916	0.0
	CR	782	31	47,841	57,230	97.2	80.6	47.8	82.0	179.4	61,648	52,689	6.8	59,826	52,499	0.1	15,678	633	56.6
3 pm	FCFS	629	79	11,070	18,080	97.3	81.1	48.2	82.4	322.4	27,330	18,948	0.0	24,054	18,266	3.1	15,791	860	0.0
	SPT	657	67	5,614	9,769	97.3	81.7	48.7	82.9	268.1	21,056	10,360	26.2	18,334	9,758	3.3	15,203	811	48.1
	EDD	629	80	0.00	0.00	97.3	81.1	48.2	82.4	321.6	15,780	857	69.4	12,515	581	40.4	15,780	857	0.0
	SLACK	607	87	0.00	0.00	97.3	80.7	47.9	82.1	346.2	16,024	875	0.0	12,508	596	53.2	16,024	875	0.0
	CR	726	29	50,023	62,785	97.3	80.5	47.8	81.9	170.9	62,125	55,679	4.4	60,394	55,495	0.0	15,121	600	51.9
4 pm	FCFS	561	82	26,973	33,274	97.2	81.2	48.3	82.4	316.6	45,424	35,018	0.0	42,088	34,249	0.4	15,366	925	0.0
	SPT	607	52	14,222	18,379	97.2	82.1	49.0	83.2	260.9	30,480	19,675	6.8	27,718	18,973	0.4	14,799	925	42.9
	EDD	561	83	0.00	0.00	97.2	81.2	48.3	82.5	315.9	15,347	928	87.3	12,018	601	35.0	15,347	928	0.0
	SLACK	535	89	0.01	0.25	97.2	80.8	48.0	82.1	340.9	15,583	949	0.1	11,994	610	64.2	15,583	949	0.0
	CR	657	31	53,694	70,056	97.2	80.7	47.9	82.1	163.9	64,066	66,456	5.8	62,415	66,254	0.0	14,610	644	57.1
5 pm	FCFS	509	85	45,756	48,295	97.2	81.2	48.3	82.4	307.0	63,757	48,231	0.0	60,559	47,444	0.0	14,865	926	0.0
	SPT	572	55	24,732	27,716	97.2	82.2	49.1	83.3	250.1	40,158	27,733	0.8	37,544	27,016	0.0	14,335	951	40.6
	EDD	508	87	0.00	0.00	97.2	81.2	48.3	82.4	305.5	14,848	912	93.7	11,661	563	32.9	14,848	912	0.0
	SLACK	480	89	0.06	1.27	97.2	80.8	47.9	82.1	330.8	15,164	1,709	0.1	11,715	1,393	66.6	15,080	941	0.0
	CR	604	33	60,877	68,854	97.2	80.8	47.9	82.1	155.8	75,991	65,710	5.4	74,406	65,517	0.5	14,116	644	59.4

Table notes

Idleness costs are for operation with 25 workers: 4 pickers, 5 packing lanes with in total 20 packers, and 1 stager.

Best (%) shows the percentage of all 1,000 simulation runs where this priority rule has lowest cost across the five considered rules.

The standard deviation columns (std) show the variation of outcomes across the 1,000 simulation runs.

Table 5: Simulation outcomes of five priority rules for three customized performance criteria

			Mean Value			Significance		
Evaluation	Objective	Unit	EDD	SLACK	GAP	t-statistic	p-value	Differ
Operational	Flow time	Hour	8.1	8.8	-0.7	-6.050	0.000	Yes
	Lateness	Hour	-11.7	-11.0	-0.7	-6.034	0.000	Yes
	Tardiness	%	0.0000	0.0002	-0.0002	-1.415	0.157	No
Financial	Max pallet staging (α)	Pallet	508	480	28	6.992	0.000	Yes
	Truck penalty (β)	€ / week	0.000	0.057	-0.057	1.416	0.157	No
	Server utilization (γ)	%	82.4	82.1	0.3	5.710	0.000	Yes
	Stock in progress (δ)	Pallet	306	331	-25	-6.050	0.000	Yes

Table notes

GAP is the difference between EDD and SLACK.

The p-value is based on the two-tailed t-distribution.

The column 'Differ' shows whether EDD and SLACK differ significantly (at 5% level).

Table 6: Welch t-test results for differences between EDD and SLACK priority rules (cut-off scenario 5 pm)