International Journal of Industrial Engineering Computations 15 (2024) 887–908

Contents lists available at GrowingScience

International Journal of Industrial Engineering Computations

homepage: www.GrowingScience.com/ijiec

An optimization approach for assembly job shop order release based on clearing functions

Liezheng Shen[a], Haiping Zhu[a,b*] and Haiqiang Hao[a]

[a]School of Mechanical Science and Engineering, Huazhong University of Science and Technology, Wuhan 430074, China
[b]National Centre of Technology Innovation for Intelligent Design and Numerical Control, Wuhan 430074, China

C H R O N I C L E

Article history:
Received March 21 2024
Received in Revised Format June 25 2024
Accepted July 9 2024
Available online July 9 2024

Keywords:
Order release
Assembly job shop
Clearing function
Production planning
Workload control

A B S T R A C T

As an integral part of production planning control, order release management is critical to enhance the competitiveness and production efficiency of companies. Previous literature shows limited application of optimization-based models in assembly job shops, primarily due to the intricate nature of product structures and assembly operations. Therefore, based on the idea of the allocated clearing function (ACF) model, we introduce material flow constraints and complex assembly structure constraints during the assembly stage, proposing the assembly job shop allocated clearing function (AACF) model. The performance of the AACF model and the rule-based mechanisms in terms of cost and timing measures are compared through experiments containing 6 factors and 96 scenarios. The results show that the AACF model performs better in terms of cost management, service level and order due date deviation. In addition, a sensitivity analysis of the objective function parameters is performed to confirm the robustness of the AACF model. Finally, a case application in a real assembly shop illustrates the feasibility and validity of the proposed AACF model.

1. Introduction

A common and important problem in order-oriented production companies is developing an appropriate order release plan for orders with known delivery times. An inappropriate order release plan can lead to uneven workloads, work-in-process (WIP) accumulation, early/late delivery, etc., resulting in significant losses. An important task of a production planning control (PPC) system is to make sure that the correct amounts of the correct products are available at the correct time. Different from a general job shop, an assembly shop is more complex, including but not limited to (i) the product has a complex assembly structure and (ii) assembly operations cannot begin until all components are ready. These characteristics require high levels of coordination and synchronization in production planning, presenting a challenge for research in production planning (Stevenson & Silva, 2008), which are also identified as one of the reasons for delays in companies (Mark Stevenson, Huang, Hendry, & Soepenberg, 2011). The concept of workload control (WLC) is important in PPC systems, which include: order entry, order release, and dispatching (Land & Gaalman, 1996), designed specifically for enterprises with order-based production models (Stevenson et al., 2005). At each stage, WLC integrates two mechanisms (Thürer et al., 2016; Thürer, Stevenson et al., 2019): (i) input control and (ii) output control. Order release is a crucial component of WLC. It serves as the interface between the planning and scheduling, determining when orders are released to the shop floor. This has a significant impact on on-time delivery of orders and reducing inventory costs (Hutter, Haeussler, & Missbauer, 2018; Missbauer & Uzsoy, 2022). The mainstream research on order release is divided into two main branches: rule-based mechanisms and optimisation-based approaches (Stefan Haeussler & Netzer, 2020; Pürgstaller & Missbauer, 2012). The rule-based mechanisms determine the release time of orders by a series of rules within a short time planning horizon without an explicit objective function. The optimization-based release models, on the other hand, use mathematical or simulation models to describe the production system and optimize the order release plan with a suitable objective function. A widely used and effective example of such a model is the Allocation Clearing Function (ACF) model (Jakob Asmundsson, Rardin, Turkseven,

* Corresponding author
E-mail haipzhu@hust.edu.cn (H. Zhu)
ISSN 1923-2934 (Online) - ISSN 1923-2926 (Print)
2024 Growing Science Ltd.
doi: 10.5267/j.ijiec.2024.7.003

& Uzsoy, 2009). In previous literature, the order release approach was widely used in various industries (Hutter et al., 2018; Portioli-Staudacher & Tantardini, 2012; Silva, Stevenson, & Thurer, 2015). However, most literatures discuss the order release in general job shops without assembly operations, and little literature has addressed the problem of order release in assembly job shops (Lu, Huang, & Yang, 2011; Thürer, Stevenson, Silva, & Huang, 2012). The optimisation-based order release approach for assembly shops requires further research, which is the motivation of this paper. The main contributions of this paper are as follows:

We study the optimisation-based release approach for assembly job shop that has not yet been studied in the literature. Using the ACF model, we introduce material flow constraints and complex assembly structure constraints in the assembly stage. And we propose an assembly job shop allocated clearing function (AACF) model. The performance of the AACF model is evaluated in 96 scenarios by two groups of simulation experiments. In experiments, the AACF model is compared with rule-based mechanisms in cost and timing measures, and the proposed AACF model performs better in cost management, service level, and due date deviation. Then, the effects of assembly centre loads are analysed in the experiments. In addition, a sensitivity analysis is performed on the parameters of the objective function to demonstrate the robustness of the AACF model. The AACF model is applied to a practical engineering case, and the results show that the proposed AACF model can balance work centre loads and significantly reduce the WIP, order lead time and order early/late delivery.

This paper is organized as follows. Section 2 reviews the related literature. Section 3 introduce the addressed problem and model. In Section 4, we describe computer experiments, including experimental design, parameter settings, etc. Section 5 analyses and discusses the experimental results. Section 6 introduces a real engineering case and analyses the optimization results. Section 7 gives conclusions and recommendations for future research.

2. Literature review

This paper focuses on order release approaches for assembly job shop. Thus, the literature review focuses on the existing mainstream order release approaches and their application in assembly shops. Two main categories are presented: (i) rule-based mechanisms and (ii) optimisation-based approaches.

2.1 Rule-based mechanisms

Rule-based mechanisms can be classified as periodic, continuous or hybrid (Thürer, Land, Stevenson, Fredendall, & Godinho Filho, 2015) depending on when decisions are made. The periodic release mechanisms make decisions at regular intervals, which are mainly different in the calculation of both planned release date (PRD) and load (Land, 2006; Oosterman, Land, & Gaalman, 2000). The main treatments of work centre loads are the aggregated load approach (Hendry & Kingsman, 1991) and the probabilistic approach (Bechte, 1988). Currently, the corrected aggregate load approaches are considered to be the best (Land & Gaalman, 1996; Oosterman et al., 2000), which are developed from the classical aggregate load approach. Different from the periodic release mechanisms, the continuous mechanisms (Melnyk & Ragatz, 1989) make release decisions at any moment by using a workload trigger. When the work centre load is about to exceed an upper bound, the job is prohibited from entering, and when the load is less than a lower bound in a shortage state, the workload trigger admits the job to the shop floor. A representative continuous mechanism is the superfluous load avoidance release (SLAR) approach (Land & Gaalman, 1998). Hendry and Kingsman (1991) proposed a well-known hybrid approach called the Lancaster University Management School Order Release (LUMS OR) approach, which combines periodic release with a continuous starvation avoidance mechanism. The LUMS OR approach was improved by using the corrected aggregate load approaches (Thürer et al., 2011), known as the LUMS COR approach.

Table 1
Rule-based mechanisms in assembly job shop

Classification	Method	Brief description
No control	IMM	Immediate release
Periodic	BIL	Backward infinite loading. The release time of the parts is their PRD. the lead time of the parts is proportional to the number of operations.
	BILA	Backward infinite loading for assembly job shop. Based on BIL rule, the release times of parts are corrected by considering a fixed assembly delay.
	BILWLC	Backward infinite loading with workload control. According to the BIL rule, if the release of a job violates the workload standard of even a single work center, the release of that job will be delayed. The delayed job will be reconsidered at the beginning of the next review period.
	Periodic WLC	Periodic workload control release. Releases parts periodically up to the workload norm. Pool sequencing is based on PRDs.
Continuous	SLAR	Releases parts if a work centre is starving or if there are no urgent parts queuing in front of a workstation. Pool sequencing is based on PRDs.
Hybrid	LUMS COR	Combining periodic release with continuous release: if a work centre is starving, parts are pulled to the shop floor between periodic reviews. Pool sequencing is based on PRDs.

Order release approaches for assembly job shop require the coordination of the completion time of each assembly component. Currently, the order release approach has been used less frequently in the assembly job shop and mainly extends the existing rule-based mechanism. The review of the assembly shop release rules is as shown in Table 1 (Liu et al., 2023; Lu et al., 2011; Paul et al., 2015; Thürer et al., 2012, 2013). Among these methods, Thürer et al. (2013) identified the LUMS COR approach as the best-performing approach for assembly shops. Liu et al. (2023) considered the workshop status and redesigned the release mechanisms and scheduling rules according to the dynamic coordination principle to improve the coordination of the assembly job shop.

2.2 Optimisation-based approaches

The essence of the optimisation-based approach is to extend the well-known production smoothing model based on linear programming, by formulating mathematical or simulation models that correlate release decisions with WIP, flow time, and capacity in a production system. Queuing theory suggests there is a nonlinear relationship between lead time and workload, implying that it is better to use different lead times for different workload scenarios. To measure the relationship between the workload and output of work centres in a stochastic production system, the concept of the Clearing Function (CF) was introduced (Graves, 1986). It relates the expected output during a planning period to the WIP levels at the work centre during that period. Asmundsson et al. (2006) defined the CF as the functional relationship between the WIP at a work center and the maximum output during a planning period. Based on the concept of CF, they proposed the ACF model to mathematically model multi-stage, multi-product type production systems. It is possible to interpret CF as a meta-model and avoid an analytical description of the queuing process in a manufacturing system (Haeussler et al., 2020). Kacar et al. (2012) compared the ACF model with the iterative approach proposed by Yi-Feng and Leachman (1996) in an equally scaled-down wafer facility scenario and found that the ACF model outperformed the iterative approach. Kacar et al. (2013) applied the ACF to a large wafer fab scenario and compared it with the fixed lead time model, showing that the ACF performed better. Kacar et al. (2016) compared the performance of LP models with and without non-integer lead times to ACF models by performing simulations in a large wafer fab. Pürgstaller and Missbauer (2012) compared optimisation-based models with a traditional rule-based mechanism and showed that the optimisation-based models are superior to the traditional workload rule-based release mechanism. Haeussler et al. (2020) compared the input-output control model with the ACF model under the same scenario and showed that the ACF model performed better, with lower inventory levels and shorter shop floor throughput times. Haeussler and Netzer (2020) compared the ACF model and the LUMS COR approach with different stochastic demand scenarios, in terms of cost, time, and load balance. The results indicate that in all cases where demand follows an exponential distribution, the ACF model outperforms rule-based mechanisms. However, under high-utilization seasonal demand scenarios, the LUMS COR method performs better in terms of total backlog cost, inventory holding, and load balancing measures. Gopalswamy and Uzsoy (2019) proposed a simulation iteration-based method for the estimation of scavenging function. Schneckenreither, Haeussler, and Gerhold (2021) proposed a flow time estimation procedure to set lead times dynamically using an artificial neural network. Zhang et al. (2023) developed an effective simulation optimisation framework to solve the re-entrant mixed-flow workshop product releasing and routing assignment problem.

In conclusion, order release approaches are widely used in job shops, and a few rule-based mechanisms are extended to the assembly job shop. However, there is a lack of application and research on optimisation-based approaches, which motivates us to study the optimisation-based approaches for the order release problem in the assembly job shop.

3. Order release model of assembly job shop

3.1 Problem description

We extend a general job shop model to a two-stage assembly job shop, which is a generalisation of the common assembly job shop structure in practice (Komaki & Kayvanfar, 2015). The assembly job shop consists of two stages: the processing stage and the assembly stage. There are n work centres that are dedicated to processing the parts in processing stage. And there are m assembly centres that put the parts together to obtain the final products in assembly stage. There are buffers between the processing stage and the assembly stage for storing parts waiting to be assembled, and unlimited buffer capacity in front of each work centre. A product is assembled from several parts, and the parts are processed in the processing stage, with different routes for different parts. Each part is processed at most once per work centre. Each product is assembled only once at the assembly centre. Orders with known delivery times need to be scheduled for a release plan. The decision is to determine the number of different parts to be released in each planning period to meet customer demands and minimise production costs. We define an order contains a product, and the meaning of the order and product is the same. In a real situation, although an order may contain multiple products, we can decompose an order into multiple "sub-orders", a "sub-order" contains a product, so that the problem is still compatible. We define the following notation:

Indices:

i order index $i \in \{1,2,\dots,I\}$

j part type index $j \in \{1,2,\dots,J\}$

p product type index, $p \in \{1,2,\dots,P\}$

t	period index, $t \in \{1,2,\dots,T\}$
n,k	work centre index, $n,k \in \{1,2,\dots,N\}$
m	assembly centre index, $m \in \{1,2,\dots,M\}$
l	operation index for the part in the processing stage, $l \in \{1,2,\dots,L_j\}$
s	linear segments index of the clearing function, $s \in \{1,2,\dots,S\}$

Variables:

X_{jt}^{n}	output of part j at work centre n in period t
R_{jt}^{n}	amount of part j released to work centre n at the start of period t
W_{jt}^{n}	WIP of part j before work centre n at the end of period t
I_{jt}^{n}	inventory of part j after work centre n at the end of period t
Y_{jt}^{nk}	amount of part j transferred from work centre n to work centre k in period t
Z_{jt}^{n}	fraction of output from work centre n allocated to part j in period t
PB_{jt}^{np}	amount of part j transferred from work centre n to buffer p at the end of period t
BW_{jt}^{p}	WIP of part j in buffer p at the end of period t
BA_{pt}^{m}	amount of product p (set of assembled parts) transferred from buffer p to assembly centre m at the end of period t
AW_{pt}^{m}	effective WIP of product p before assembly centre m at the end of period t
AX_{pt}^{m}	output of product p at assembly centre m in period t
AI_{pt}^{m}	inventory of product p after assembly centre m at the end of period t
AY_{pt}^{mf}	amount of product p transferred from assembly centre m to final inventory in period t
AZ_{pt}^{m}	fraction of output from assembly centre m allocated to product p in period t
F_{pt}	inventory of product p at the end of period t
B_{pt}	backorder of product p at the end of period t

Parameters:

ω_{jt}^{n}	WIP holding cost per unit in work centre n for part j in period t
π_{jt}^{p}	WIP holding cost per unit for part j in buffer p for part j in period t
ρ_{pt}^{m}	WIP holding cost per unit in assembly centre m for product p in period t
ϕ_{pt}	backlog cost per unit for product p
χ_{pt}	inventory holding cost per unit for product p
pt_{jt}^{n}	operation time of part j at work centre n in period t
at_{pt}^{m}	operation time of product p at assembly work centre m in period t
K_{pj}	number of units of part j required to assemble one unit of product p
D_{pt}	demand of product type p in period t
α_{ns}^{P}	intercept of the sth linear segment of the CF of work centre n
β_{ns}^{P}	slope of the sth linear segment of the CF of work centre n
α_{ms}^{A}	intercept of the sth linear segment of the CF of assembly centre m
β_{ms}^{A}	slope of the sth linear segment of the CF of assembly centre m
L_j	number of operations for part j in processing stage
O_{jl}^{n}	If the lth operation of part j can be processed on work centre n, $O_{jl}^{n} = 1$, otherwise, $O_{jl}^{n} = 0$
T_{jnk}^{P}	If part j can be transferred from work centre n to work centre k, $T_{jnk}^{P} = 1$, otherwise, $T_{jnk}^{P} = 0$
T_{jnp}^{B}	If part j can be transferred from work centre n to buffer p, $T_{jnp}^{B} = 1$, otherwise, $T_{jnp}^{B} = 0$
T_{pm}^{A}	If product p can be assembled on assembly centre m, $T_{pm}^{A} = 1$, otherwise, $T_{pm}^{A} = 0$.

3.2 AACF model

In the ACF model, the planning horizon is divided into planning periods $t = 1,\dots,T$. The material flow is represented by inventory balance equations for WIP at each work centre and for final products (Stefan Haeussler & Missbauer, 2014). The ACF model is mainly for the job shop scenario and does not apply to the assembly job shop scenario.

In this section, we propose the AACF model by extending the ACF model. The material flow in the AACF model is shown schematically in Figure 1, where the blue part is a part of the ACF model, and the red part adds various constraints in the assembly stage. The objective function (1) mainly includes WIP cost (WIPC), finished goods inventory cost (FGIC), and backorder cost (BOC).

In a general job shop, a part is converted into a final product after all processing operations. In the assembly shop, the parts are still WIP after all processing operations and need to be transferred to a buffer before the assembly centre to wait for assembly. The parts are assembled into final products in assembly centres and then enter the finished goods inventory. Therefore, the cost in the objective function (1) consists of four items: (i) the WIPC generated by the processing stage, (ii) the WIPC generated by the assembly stage, (iii) the WIPC generated by the buffer between the processing stage and assembly stage, and (iv) the FGIC and BOC.

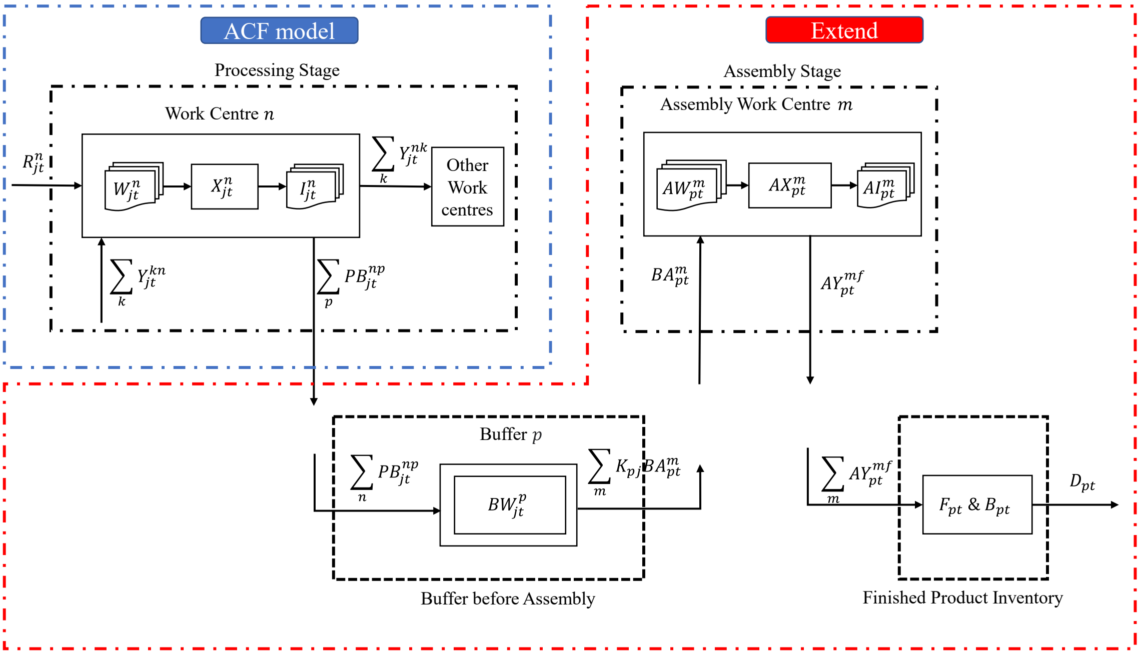

Fig. 1. Schematic of material flow in the AACF model.

Objective:

$$min \left[\sum_{n=1}^{N} \sum_{j=1}^{J} \sum_{t=1}^{T} \omega_{jt}^{n} \left(W_{jt}^{n} + I_{jt}^{n} \right) + \sum_{m=1}^{M} \sum_{p=1}^{P} \sum_{t=1}^{T} \rho_{pt}^{m} AW_{pt}^{m} + \sum_{p=1}^{P} \sum_{j=1}^{J} \sum_{t=1}^{T} \pi_{jt}^{p} BW_{jt}^{p} + \sum_{p=1}^{P} \sum_{t=1}^{T} \left(\chi_{pt} F_{pt} + \phi_{pt} B_{pt} \right) \right] \tag{1}$$

subject to:

$$W_{jt}^{n} = W_{j,t-1}^{n} - X_{jt}^{n} + R_{jt}^{n} + \sum_{k} Y_{jt}^{kn} , \forall n, j, t \tag{2}$$

$$I_{jt}^{n} = I_{j,t-1}^{n} + X_{jt}^{n} - \sum_{k} Y_{jt}^{nk} - \sum_{p} PB_{jt}^{np} , \forall n, j, t \tag{3}$$

Constraints (2) define the inventory balance equation for the WIP inventory before the work centre. Constraints (3) define the inventory balance equation for the inventory after the work centre. Different from the ACF model, the $\sum_p PB_{jt}^{np}$ in constraints (3) represents parts transferred to the buffer p that will be assembled into product p.

$$pt_{jt}^{n} X_{jt}^{n} \le pt_{jt}^{n} \beta_{ns}^{P} \left(W_{jt}^{n} + X_{jt}^{n} \right) + \alpha_{ns}^{P} Z_{jt}^{n}, \forall n, j, t, s \tag{4}$$

$$\sum_{j=1}^{J} Z_{jt}^{n} = 1, \forall n, t \tag{5}$$

The main purpose of CFs is to model the maximum output of work centre n as a (non-linear) function of the workload over a given period (Jakob Asmundsson et al., 2009; Guhlich, Fleischmann, Moench, & Stolletz, 2018). Constraints (4) and (5) define the relationship between output and workload at work centre n by segmented linearised CF. The buffer p is used to store parts waiting to be assembled into product p. For the assembly centre, if the parts do not match the assembly BOM, the assembly operation cannot proceed. Constraints (6) define the WIP inventory balance relationship for the buffer between the processing stage and assembly stage and represent the transformation relationship between part j and product p.

$$BW_{jt}^{p} = BW_{j,t-1}^{p} + \sum_{n} PB_{jt}^{np} - \sum_{m} K_{pj} BA_{pt}^{m}, \forall p, j, t \tag{6}$$

Constraints (7) and (8) define the WIP balance relationship in assembly stage:

$$AW_{pt}^{m} = AW_{p,t-1}^{m} + BA_{pt}^{m} - AX_{pt}^{m}, \forall m, p, t \tag{7}$$

$$AI_{pt}^{m} = AI_{p,t-1}^{m} + AX_{pt}^{m} - AY_{pt}^{mf}, \forall m, p, t \tag{8}$$

where the AY_{pt}^{mf} denotes the product p transfer from assembly centre m to the finished goods inventory (FGI). Constraints (9) and (10) define the capacity constraints of assembly centres:

$$at_{jt}^{m} AX_{pt}^{m} \le at_{jt}^{m} \beta_{ms}^{A} \left(AW_{pt}^{m} + AX_{pt}^{m} \right) + \alpha_{ms}^{A} AZ_{pt}^{m}, \forall m, p, t, s \tag{9}$$

$$\sum_{p=1}^{P} AZ_{pt}^{m} = 1, \forall m, t \tag{10}$$

$$F_{pt} - B_{pt} = F_{p,t-1} - B_{p,t-1} + \sum_{m=1}^{M} AY_{pt}^{mf} - D_{pt}, \forall p, t \tag{11}$$

Constraints (11) define the inventory balance relationship for the finished goods. Constraints (12) guarantee that part j is only released to the work centre of its first operation. In addition, constraints (13) (14) and (15) ensure the correct processing route of part j and assembly route of product p. Finally, constraints (16) and (17) guarantee that the variables are non-negative.

$$R_{jt}^{n} = \begin{cases} \ge 0, if\ O_{j,1}^{n} = 1 \\ 0, else \end{cases}, \forall n, j, t \tag{12}$$

$$Y_{jt}^{nk} = \begin{cases} \ge 0, if\ T_{jnk}^{P} = 1 \\ 0, else \end{cases}, \forall n, k, j, t \tag{13}$$

$$PB_{jt}^{np} = \begin{cases} \ge 0, if\ T_{jnp}^{B} = 1 \\ 0, else \end{cases}, \forall n, p, j, t \tag{14}$$

$$BA_{pt}^{m} = \begin{cases} \ge 0, if\ T_{pm}^{A} = 1 \\ 0, else \end{cases}, \forall m, p, t \tag{15}$$

$$W_{jt}^{n}, X_{jt}^{n}, I_{jt}^{n}, R_{jt}^{n}, BW_{jt}^{p}, Z_{jt}^{n}, F_{pt}, B_{pt} \geq 0, \forall n, j, p, t \tag{16}$$

$$AW_{pt}^{m}, AX_{pt}^{m}, AI_{pt}^{m}, AY_{pt}^{mf}, AZ_{pt}^{m}\ AZ_{pt}^{m} \geq 0, \forall m, j, p, t \tag{17}$$

3.3 Estimation of CF parameters

The clearing function measures the nonlinear relationship between the workload and output of work centres in production systems (Gopalswamy & Uzsoy, 2019; Kacar & Uzsoy, 2014; Kacar & Uzsoy, 2015), which allows the load and output of a stochastic production system to be modelled and solved linearly. In this paper, the main steps in estimating the parameters of CF are as follows:

(1) Data collection

We obtain data on the loads and outputs of each work centre through a discrete-event simulation (DES) model and estimate CF according to the simulation data (Kacar & Uzsoy, 2015). The steps are as follows.

Step 1: Modelling discrete-event simulation of the studied production system, focusing on detailed modelling of work centres, processing (assembly) times, assembly relationships, process routes, random failure, etc.
Step 2: Assuming an exponential distribution of demand arrival intervals, orders are then released into the production system as soon as they arrive (as introduced in '4.3 Demand setting'). The utilisation of the bottleneck work unit in the simulation system can be controlled to be 50%, 60%, 70%, 80%, 85%, 90%, and 95% by setting the different average intervals of order arrivals.
Step 3: The IMM rule is used in each utilisation scenario, and each simulation runs for 100 periods to collect the workload data including $W_{jt}^{n}, X_{jt}^{n}, AW_{pt}^{m}, AX_{pt}^{m}$.
Step 4: Merge all data and graph the data for each work centre for visual inspection.

(2) Linear regression

The CFs are fitted to the above simulation data. As shown in (18), the CFs are fitted as a min-affine function with S affine segments.

$$f(\Lambda_t^n) = min\{\alpha_s^n + \beta_s^n \Lambda_t^n\}, s \in \{1, \dots, S\} \tag{18}$$

where Λ_t^n denotes the workload of the work centre n in period t, S is set to 2. The α_s^n, β_s^n denote the intercept and the slope of the sth linear segment of the CF of work (or assembly) centre. We use a heuristic algorithm to solve the large-scale CF fitting problem. The pseudocode for segmented linear regression is shown in Appendix A.

4. Computational experiments

This section describes the computational experiments, including the simulation model, experimental procedure and parameter settings.

4.1 Simulation model

The assembly shop processes 24 (2 × 3 × 4) kinds of parts in the processing stage, which can be assembled into 6 products. These 24 parts are divided into 6 groups, each containing 4 different part types. Each group can be assembled into 1 product, 1 of each kind of part is required. The processing stage is a hybrid flow shop with three stages, and each stage consists of nonidentical parallel work centres. The stages consist of 2, 3 and 4 work centres respectively. In stages 1, 2 and 3, the processing time of the part is 40, 60 and 50 time units, respectively. The number of assembly centres is set to 1. The assembly times will be set in the experiment. Assume that the system works every day without interruption. At each work centre, products are processed according to FCFS rule. For a product, the assembly process cannot begin until all the parts are matched. A planning period length is set at 1440 time units. In the experiments in this paper, parts for different products are not shared because we want to avoid the impact of part assignment decisions. For example, if part C is a common part of product A and product B. When part C is released, a decision must be made about which product (A or B) C should be assembled to. To avoid that decision, we can directly define part C as two kinds of parts: C-A and C-B, but they are the same in practice.

4.2 Experimental design

As shown in Table 2, the first column indicates the groups of different experiments, the second column denotes the experimental factor included in that experimental group, the third column denotes the value taken for that experimental factor, and the fourth denotes the number of values taken for the experimental factor. We conducted full factorial experiments on factors except utilisation and failure. Experimental group A enables to compare the optimisation of various approaches in different scenarios and to observe the influence of different factors. The assembly time is set to 60 time units. In the long failure scenarios, experiments are conducted only in the 80% utilisation condition. Experiment group B enables to observe the influence of different assembly centre loads on the optimisation effect of each approach. In Experiment group B, other factors are: exponential demand, 80% bottleneck utilisation, normally distributed processing time, and short failure.

Table 2
Experimental design

Group	Factor	Treatments	Quantity
A	Approach	IMM, BILA, LUMS COR, AACF	4
	Demand	Exponential, Seasonal	2
	Utilisation	80%, 90%	2
	Processing Time	Deterministic, Normally	2
	Failure	None, Short, Long	3
B	Approach	IMM, BILA, LUMS COR, AACF	4
	Assembly time	60,70, 80, 90	4

The IMM, BILA (Lu et al., 2011) and LUMS COR (Thürer et al., 2012; Thürer et al., 2013) approaches in the literature are used for comparison. A total of 96 scenarios are simulated. In order to overcome the effect of the randomization factor and to ensure that the results of the experiments are statistically significant, each scenario is run independently 20 times and the mean values are taken for comparison, with a total of 1920 experiments run.

4.3 Demand setting

Assuming that total demand is seasonal or exponentially distributed (Stefan Haeussler & Netzer, 2020; S. Haeussler et al., 2020). For the exponentially distributed demand, orders arrive at exponentially distributed intervals with an average interval of τ minutes. The product type is determined by the discrete uniform distribution. The order due date is calculated by Eq. (19). Note that the optimization model does not know the product type and due date until the orders arriving.

$$DD_i = AT_i + unif\{\eta_{min}, \eta_{max}\}, \forall i \tag{19}$$

where DD_i is the due date of order i, η_{min} is the minimum interval, η_{max} is the maximum interval (unit: period). If η_{min} is too small, the order will not be completed on time; if η_{max} is too large, information about all orders will not be available within planning horizons. After many simulation experiments, we found that the throughput time of products in steady state is generally concentrated in 10~15 periods. Therefore, the η_{min}, η_{max} is set to 10 and 15, respectively. For the seasonal demand, there are two steps, the first step determines the total demand for period t, through the equation (20).

$$D_t = k_1 + k_2 \sin\left(\frac{2\pi}{k_3} t + \pi\right), \forall t \tag{20}$$

Where D_t is the total demand for period t, k_1 is the constant part of total demand, k_2 is the seasonal variation of total demand, and k_3 is the seasonal variation cycle. In the second step, after obtaining the total demand for each period, we obtain the quantity of each product type by discrete uniform distribution, given by equation (21).

$$AT_i = DD_i - unif\{\eta_{min}, \eta_{max}\}, \forall i, t \tag{21}$$

The approximate ranges for the parameters can be given: $\tau \in [50,100]$, $k_1 \in [18,22]$, $k_2 \in [18,22]$, $k_3 \in [8,12]$. Note that in both of these demand scenarios, the parameters need to be adjusted according to the utilisation of the workshop in steady state. It can be seen from the simulation model that the bottlenecks will occur in stage 1 (work centre 1~2) and stage 2 (work centre 3~5). Assuming that U_r is the required utilisation rate and U_{max} is the maximum utilisation rate for all work centres, steps for order data generation: (S1) the parameters τ, m, α, c are constantly tuned and simulated, until the average utilisation of stage 1 and stage 2 has an error of less than $\pm 2\%$ with respect to UR; (S2) After determining the parameters τ, m, α, c, the simulation is run continuously until $|U_{max} - UR| \leq 2\%$, the demand will be adopted.

4.4 Random factors

In this paper, discrete event simulation software is used to simulate random factors such as time and failure.

For the normally distributed processing time, we define the mean and standard deviation of the normal distribution as pt_{jt}^n and $pt_{jt}^n/10$, respectively. For the normally distributed assembly time, we define the mean and standard deviation of the normal distribution as at_{pt}^m and $at_{pt}^m/10$, respectively.

For the failure of work centres and assembly centres, we define the mean time to failure (MTTF) and the mean time to repair (MTTR) as an exponential distribution. For short failure, the machine availability is defined as 0.95. MTTF is set to 1368 time units and MTTR is set to 72 time units. For long failure, the machine availability is defined as 0.90. MTTF is set to 1296 time units and MTTR is set to 144 time units.

4.5 Parameters of approaches

For all approaches, the parameters $\omega_{jt}^{n}, \pi_{jt}^{n}, \rho_{pt}^{m}, \chi_{pt}, \phi_{pt}$ in the objective function are set to 1, 1, 1, 10, and 100.

4.5.1 BILA

BILA rule assumes that the time required for production of a product is highly correlated with the number of operations, and is relatively fixed. Under the BILA rule, the release of parts is subject to periodic review. At the start of each cycle, parts scheduled for release within half a cycle before and after the review point are released immediately at the review point. The planned release date for each part of the product is determined by the formula (22).

$$R_{ij} = DD_i - PF_1 \times L_j - AT_i, \forall i, j \tag{22}$$

where R_{ij} represents the release time of part j of order i; DD_i represents the delivery time of order i; PF_1 is the planning factor, and a positive integer; AT_i indicates the assembly time of order i. The pool sequencing rule sorts all parts according to the PRDs (R_{ij}).

4.5.2 LUMS COR

Currently, the LUMS COR approach is considered to be widely used and effective, there are four parameters that influence the performance (Stefan Haeussler & Netzer, 2020) : (i) the release time of part; (ii) the time limit (TL), the rule selects those parts whose PRDs fall within TL; (iii) the upper bound of the load of the work centre; (iv) the release frequency. In this paper, the upper bound of the load of work centre n (or assembly centre m) is determined by the maximum intercept of the corresponding CF (different for each centre and each scenario), and the release frequency is set to a period length. The PRD of each part of the product is determined by formula (23).

$$R_{ij} = DD_i - PF_2 \times L_j - RT_j - AT_i \tag{23}$$

where RT_j denotes the total processing time of part j; PF_2 is the planning factor and a positive integer. The LUMS COR approach also contains a continuous workload trigger. Pool sequencing rule sorts all parts according to the PRDs (R_{ij}). Since TL is applied, only parts within the TL are included ($R_{ij} \leq current\ time\ +\ TL$).

4.5.3 Pre-experiments for parameters

For each experiment scenario, the full factorial simulation pre-experiments will be used to confirm the parameters of both rules to minimise the objective function (1). In pre-experiments for BILA, the range of PF_1 is 1 to 2000, the step is set to 10, and a total of 200 simulations are run for each scenario. In the pre-experiments for the LUMS COR approach, the range of PF_2 is 1 to 2000, the step is set to 10, the range of TL is 1 to 10 periods, the step is set to 1, and a total of 2000 (200×10) simulation are run for each scenario.

4.6. Implementation

The FactorySimulation software (www.factorysimulation.net) is used to simulate the production system, and the Gurobi software (www.gurobi.com) is used to solve the mathematical model. Figure 2 shows the flow of the simulation experiment, the specific steps are as follows:

Step 1: According to '4.3 Demand setting', orders for 30 periods are obtained through a simulation model. The simulation model is run for 15 periods to warm up, and the incomplete orders between periods 15 and 30 are taken as the demand for 15 periods.

Step 2: The IMM rule is used to make release decisions for the incomplete orders, if the utilisation of the bottleneck reaches the requirements, go to Step 3, otherwise, return to Step 1.

Step 3: The BILA, LUMS COR, and AACF approaches are used to make release decisions for the incomplete orders, respectively. Run the simulation models with the FCFS dispatching rule to obtain the results.

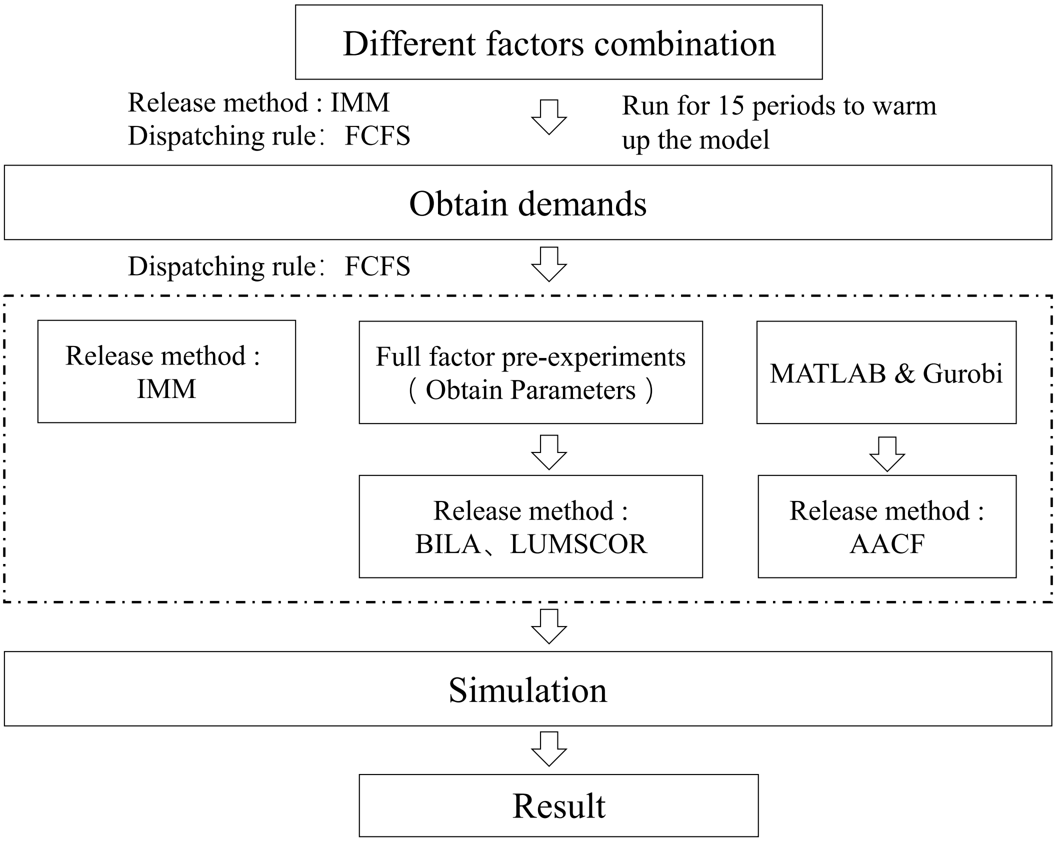

Fig. 2. Procedure for computational experiments

5. Experiment results

5.1 Experiment group A

We evaluated the performance of different order release methods in measures of cost, and timing. Cost measures include WIPC, FGIC, and BOC. Timing measures include service level (SL, the proportion of orders delivered in time), mean absolute due date deviation (MDD), the standard deviation of due date deviation (SDD), shop floor throughput time which is the average time an order takes from order release to completion (SFTT), and the standard deviation of the SFTT (σ_{SFTT}).

5.1.1 Exponential demand

This section analyses the results under different experimental conditions for the exponential demand case. Table 3 indicates the results for the no failure scenarios. Table 4 indicates the results for the short failure scenarios. The tables are divided into two sections, denoting deterministic processing time in the upper part and normally distributed processing time in the lower part. The upper part of Table 3 shows the result for exponential demand, deterministic processing time and no failure scenarios. The IMM rule generates the highest total cost, with FGIC and WIPC accounting for the largest proportion and BOC accounting for a lower proportion. This suggests that the IMM rule leads to early order release and completion, resulting in more WIP and FGI accumulation. The AACF model consistently generates the lowest total cost in all scenarios, followed by the LUMS COR approach. In terms of cost structure, the LUMS COR approach yields the lowest WIPC but generates more FGIC, suggesting that the LUMS COR approach tends to facilitate early order completion. The lower part of Table 3 shows the result for exponential demand, normally distributed processing time and no failure scenarios. At 80% utilisation rate, the FGIC and BOC of the AACF model are significantly lower, highlighting a significant cost advantage. The BOC of all approaches rises sharply as utilisation increases. The AACF model continues to keep the lowest FGIC, resulting in the lowest total cost. However, the cost advantage of the AACF model gradually diminishes as the utilisation rate increases. The SL, MDD, and SDD of the AACF model outperform those of the LUMS COR approach and the BILA rule, suggesting superior performance in on-time delivery, due date deviation and stability. The LUMS COR approach generates the lowest values for SFTT and σ_{SFTT}, suggesting that it can control the order lead time more efficiently and consistently, leading to a smoother production process. For all approaches, an increase in the utilisation rate corresponds to a significant decrease in service level and a reduction in due date deviation. The SFTT and σ_{SFTT}, values of the AACF model also decrease as utilisation increases. The effect of stochastic processing time on the various measures can be analysed by comparing the upper and lower parts of Table 3. In cost measures, the stochastic time contributes to a slight rise in BOC and a slight increase in total cost. In timing measures, the stochastic time leads to a minor increase in due date deviation and a slight decrease in due date deviation stability, which is a small impact overall.

Table 3
Results for exponential demand and no failure

Exponential demand, deterministic processing time and no failure										
Utilization	Method	WIPC	FGIC	BOC	SUM	SL	MDD	SDD	SFTT	σ_{SFTT}
80%	AACF	958.0	1607.4	115.1	2680.4	0.97	0.88	0.45	2.44	1.70
	IMM	4423.9	4587.2	201.0	9212.1	0.97	2.55	1.52	6.06	3.16
	BILA	1188.4	2335.4	671.5	4195.4	0.90	1.31	0.74	1.65	0.82
	LUMS	578.8	2511.4	558.1	3648.3	0.92	1.39	0.74	0.88	0.39
90%	AACF	965.5	1435.2	3222.2	5622.9	0.81	0.82	0.50	1.86	1.33
	IMM	5911.4	2610.0	3940.6	12462.1	0.76	1.41	0.95	6.92	3.61
	BILA	1971.2	1961.4	3661.1	7593.6	0.75	1.09	0.69	2.33	1.04
	LUMS	804.5	2265.3	2969.2	6039.0	0.80	1.20	0.78	1.04	0.40
Exponential demand, normally distributed processing time and no failure										
Utilization	Method	WIPC	FGIC	BOC	SUM	SL	MDD	SDD	SFTT	σ_{SFTT}
80%	AACF	909.9	1549.1	228.9	2687.9	0.97	0.88	0.45	2.36	1.72
	IMM	4237.5	4795.4	357.6	9390.4	0.97	2.75	1.55	5.98	3.14
	BILA	1193.2	2406.2	986.4	4585.8	0.89	1.40	0.76	1.71	0.89
	LUMS	574.8	2604.6	839.3	4018.8	0.91	1.51	0.78	0.89	0.42
90%	AACF	965.6	1482.7	3777.8	6226.1	0.82	0.88	0.56	1.90	1.33
	IMM	5801.9	3018.4	4112.9	12933.2	0.77	1.64	1.10	6.87	3.59
	BILA	1960.6	2165.7	3703.8	7830.1	0.76	1.21	0.76	2.34	1.09
	LUMS	768.1	2478.0	3273.0	6519.1	0.78	1.33	0.84	1.01	0.40

The upper part of Table 4 shows the results for exponential demand, deterministic processing time, and short failure scenarios. The AACF model consistently produces the lowest total cost in all scenarios, demonstrating superior performance in on-time delivery, due date deviation, and stability measures. At 80% utilisation, the IMM rule performs better on service level but worse on other timing measures. Meanwhile, the IMM rule generates more WIPC and FGIC, suggesting that the IMM rule leads to early release and early completion of orders, sacrificing cost for good on-time delivery performance. The lower part of Table 4 presents the results for exponential demand, normally distributed processing time and short failure scenarios. In cost terms, short failures lead to a decrease in FGIC, a significant increase in BOC and a minimal impact on WIPC, resulting in a significant increase in total cost. The failure increases the randomness of the process and the risk of order delays, resulting in: (i) more FGIC converted to BOC, and (ii) a significant decrease in service level and a slight increase in the throughput cycle of orders.

Table 4
Results for exponential demand and short failure

Exponential demand, deterministic processing time and short failure										
Utilization	Method	WIPC	FGIC	BOC	SUM	SL	MDD	SDD	SFTT	σ_{SFTT}
80%	AACF	1064.4	1345.9	904.9	3315.1	0.89	0.76	0.46	2.47	1.55
	IMM	4883.7	3775.3	905.1	9564.1	0.92	2.07	1.26	6.51	3.36
	BIL	1471.9	2279.1	1224.1	4975.1	0.88	1.27	0.76	1.98	0.93
	LUMS	679.5	2411.9	1412.9	4504.2	0.86	1.35	0.78	1.00	0.43
90%	AACF	1012.8	1200.8	5792.4	8005.9	0.76	0.85	0.60	1.94	1.40
	IMM	6046.7	2363.1	7788.7	16198.5	0.66	1.51	1.06	7.22	3.83
	BIL	2023.7	1684.0	5644.8	9352.5	0.69	1.07	0.70	2.44	1.14
	LUMS	854.5	1939.9	5846.1	8640.5	0.69	1.19	0.80	1.13	0.43
Exponential demand, normally distributed processing time and failure										
Utilization	Method	WIPC	FGIC	BOC	SUM	SL	MDD	SDD	SFTT	σ_{SFTT}
80%	AACF	997.7	1379.3	1761.7	4138.7	0.87	0.82	0.52	2.28	1.51
	IMM	5026.7	3346.2	1728.3	10101.1	0.88	1.87	1.10	6.61	3.37
	BIL	1508.4	2128.3	2038.0	5674.7	0.82	1.22	0.72	2.01	0.95
	LUMS	739.9	2289.7	2067.5	5097.0	0.83	1.30	0.76	1.07	0.45
90%	AACF	1026.9	1168.4	7577.9	9773.3	0.72	0.90	0.71	1.88	1.38
	IMM	6160.9	2208.8	8068.3	16438.0	0.64	1.42	0.98	7.26	3.77
	BIL	2230.1	1426.7	8861.1	12517.9	0.60	1.08	0.75	2.65	1.24
	LUMS	892.7	2026.8	7306.8	10226.3	0.68	1.29	0.90	1.17	0.42

Table 5 shows the results for exponential demand, deterministic processing time, 80% utilisation and long failure scenarios. The tables are divided into two sections, denoting deterministic processing time in the upper part and normally distributed processing time in the lower part. As failures increase, the trend is generally the same for all approaches: (i) a continued decrease in FGIC, a further increase in BOC, and an increase in total cost; and (ii) a continued decrease in service level and due date deviations, and a continued increase in throughput cycle of orders (except the AACF model).

Table 5
Results for exponential demand and long failure

Exponential demand, deterministic processing time, 80% utilization and long failure										
Proc-Time	Method	WIPC	FGIC	BOC	SUM	SL	MDD	SDD	SFTT	σ_{SFTT}
Deterministic	AACF	1056.3	1010.1	4338.4	6404.7	0.73	0.75	0.58	2.18	1.42
	IMM	5318.6	2435.7	5018.9	12773.2	0.72	1.56	1.05	6.97	3.63
	BIL	1836.4	1751.2	4359.5	7947.1	0.73	1.15	0.73	2.45	1.11
	LUMS	878.4	2029.7	4838.0	7746.1	0.72	1.32	0.90	1.27	0.46
Normally distributed	AACF	1033.7	890.7	5667.6	7592.1	0.67	0.75	0.57	2.01	1.24
	IMM	5435.2	2236.7	5355.0	13026.9	0.71	1.46	1.04	7.07	3.67
	BIL	1813.1	1455.0	5412.2	8680.3	0.67	1.03	0.66	2.38	1.07
	LUMS	875.1	1708.8	6387.1	8970.9	0.67	1.22	0.82	1.27	0.47

In summary, in exponential demand scenarios, the AACF model outperforms other methods in terms of total cost, service level, and due date deviation. The LUMS COR approach excels in WIPC and the throughput cycle of orders. The randomness of processing time has a minor impact on various metrics, while the randomness of failures increases the risk of order delays, significantly affecting both cost and service level.

5.1.2 Seasonal demand

In this section, the results for different experimental conditions in the seasonal demand scenarios are analysed. Table 6 shows the results for no failure scenarios. Table 7 shows the results for short failure scenarios. The tables are divided into two sections, denoting deterministic processing time in the upper part and normally distributed processing time in the lower part. As shown in Table 6, in terms of cost measures, the AACF model consistently generates the lowest total cost in all scenarios, followed by the LUMS COR approach. The cost advantage of the AACF model lies in consistently avoiding being released too early. Similar to exponential demand, the LUMS COR approach tends to promote early order completion, resulting in the lowest WIPC and higher FGIC. Different from exponential demand, the LUMS COR approach generates a higher SL. The AACF model generates the lowest SL, suggesting that it effectively controls order delays. In addition, the AACF model generates the lowest MDD and SDD, suggesting excellent stability of due date deviation. The SFTT and σ_{SFTT} of the LUMS COR approach are the lowest, suggesting effective and stable control over order lead times. The effect of stochastic processing time on the various measures can be analysed by comparing the upper and lower parts of Table 6. In cost measures, the stochastic time has less impact, mainly resulting in a slight increase in BOC (excluding the IMM rule) and a slight increase in total costs. In timing measures, the stochastic time leads to a slight decrease in the SL, which is not more than 0.02.

Table 6
Results for seasonal demand and no failure

Seasonal demand, deterministic processing time and no failure										
Utilization	Method	WIPC	FGIC	BOC	SUM	SL	MDD	SDD	SFTT	σ_{SFTT}
80%	AACF	1030.8	1345.6	468.9	2845.3	0.86	0.75	0.72	2.55	1.45
	IMM	4578.1	4529.9	477.4	9585.4	0.95	2.47	1.58	6.20	3.22
	BIL	1663.9	2808.7	827.8	5300.4	0.90	1.56	0.97	2.28	1.11
	LUMS	651.2	2913.4	648.2	4212.7	0.91	1.60	0.96	0.97	0.41
90%	AACF	1130.7	1209.6	4453.7	6793.9	0.71	0.76	0.68	2.29	1.47
	IMM	6148.9	2623.1	4987.3	13759.4	0.73	1.43	1.03	7.07	3.69
	BIL	2550.6	2118.5	4963.3	9632.4	0.70	1.20	0.83	2.96	1.31
	LUMS	876.5	2534.5	3799.8	7210.8	0.75	1.34	0.97	1.11	0.40
Seasonal demand, normally distributed processing time and no failure										
Utilization	Method	WIPC	FGIC	BOC	SUM	SL	MDD	SDD	SFTT	σ_{SFTT}
80%	AACF	1044.2	1338.5	563.6	2946.2	0.85	0.76	0.70	2.61	1.56
	IMM	4530.7	4673.8	420.5	9625.0	0.96	2.56	1.41	6.18	3.19
	BIL	1626.2	3041.3	788.4	5455.9	0.91	1.70	0.99	2.25	1.13
	LUMS	650.5	2992.4	810.3	4453.2	0.90	1.67	0.97	0.98	0.43
90%	AACF	1092.3	1222.4	4229.1	6543.8	0.71	0.76	0.68	2.11	1.31
	IMM	6061.0	2635.2	4919.3	13615.5	0.72	1.44	1.04	7.00	3.70
	BIL	2468.2	2069.4	4743.2	9280.8	0.70	1.17	0.82	2.87	1.27
	LUMS	822.3	2443.1	3696.0	6961.4	0.75	1.29	0.95	1.05	0.39

As shown in Table 7, similar to the no failure scenario, the AACF model generates the lowest total cost, followed by the LUMS COR approach. The AACF model generates lower FGIC, higher BOC and lower SL. The SFTT of the LUMS COR approach outperforms others, demonstrating better balance.

Table 7
Results for seasonal demand and short failures

Seasonal demand, deterministic processing time and short failures										
Utilization	Method	WIPC	FGIC	BOC	SUM	SL	MDD	SDD	SFTT	σ_{SFTT}
80%	AACF	1049.6	1011.7	1453.6	3515.0	0.76	0.65	0.63	2.64	1.60
	IMM	4440.1	4189.3	1245.4	9874.8	0.90	2.45	1.63	6.32	3.33
	BILA	1639.1	2583.4	1399.7	5622.2	0.86	1.55	0.95	2.36	1.18
	LUMS	675.5	2594.1	1421.5	4691.0	0.86	1.55	0.96	1.06	0.46
90%	AACF	1182.7	861.0	6312.2	8355.9	0.61	0.70	0.66	2.29	1.45
	IMM	6190.9	2063.8	6434.2	14688.9	0.68	1.27	0.86	7.28	3.85
	BILA	2459.4	1744.3	7096.9	11300.5	0.65	1.15	0.76	2.91	1.31
	LUMS	939.9	2252.8	5589.5	8782.2	0.71	1.32	0.91	1.22	0.43
Seasonal demand, normally distributed processing time and short failures										
Utilization	Method	WIPC	FGIC	BOC	SUM	SL	MDD	SDD	SFTT	σ_{SFTT}
80%	AACF	1073.9	963.1	1890.8	3927.8	0.74	0.65	0.59	2.64	1.59
	IMM	4506.4	4147.6	889.8	9543.8	0.91	2.39	1.56	6.36	3.35
	BILA	1628.6	2605.5	1497.9	5732.0	0.85	1.55	0.93	2.33	1.17
	LUMS	681.6	2609.7	1694.8	4986.1	0.85	1.56	0.95	1.06	0.46
90%	AACF	1154.5	870.1	7005.0	9029.5	0.58	0.74	0.69	2.25	1.39
	IMM	6142.3	2181.3	7484.2	15807.8	0.64	1.38	0.97	7.27	3.86
	BILA	2448.6	1669.8	7555.9	11674.3	0.62	1.14	0.78	2.93	1.32
	LUMS	901.0	2313.7	6190.0	9404.8	0.68	1.38	1.01	1.18	0.43

Table 8
Results for seasonal demand and long failures

Seasonal demand, deterministic processing time, 80% utilization and long failures										
Proc-Time	Method	WIPC	FGIC	BOC	SUM	SL	MDD	SDD	SFTT	σ_{SFTT}
Deterministic	AACF	1033.9	960.8	3167.6	5162.3	0.72	0.65	0.54	2.16	1.28
	IMM	5460.2	2799.9	3710.0	11970.1	0.75	1.60	1.09	6.95	3.75
	BILA	2251.1	2310.5	2898.8	7460.3	0.76	1.31	0.93	2.89	1.31
	LUMS	845.7	2239.8	3287.8	6373.3	0.73	1.30	0.93	1.19	0.39
Normally distributed	AACF	870.1	867.4	4479.9	6217.3	0.63	0.68	0.57	1.77	1.02
	IMM	5335.5	1897.1	5453.5	12686.1	0.68	1.27	0.88	6.97	3.64
	BILA	1861.1	1505.8	4924.7	8291.6	0.68	1.04	0.66	2.46	1.13
	LUMS	875.7	1956.6	4421.5	7253.7	0.69	1.24	0.88	1.29	0.48

Table 8 shows the results for the seasonal demand and long failure scenarios. The tables are divided into two sections, denoting deterministic processing time in the upper part and normally distributed processing time in the lower part. By comparing Table 6, Table 7 and Table 8, it can be observed that failures lead to a decrease in FGIC and a significant increase in BOC for all approaches, no change in WIPC and a corresponding increase in total cost. However, combining the analysis of SL, MDD, and SFTT, it is evident that: (i) failures do not cause a substantial increase in MDD and SFTT, and (ii) failures only result in more delayed deliveries, and the delays aren't significant. The increase in utilisation, randomness of processing time and failure all lead to an increase in costs. Among these, utilisation and failures have a greater impact on costs, mainly leading to an increase in BOC for all approaches. The stochasticity leads to more conservatism in estimating the parameters of CF of the work centre, resulting in a lower maximum capacity of the work centre in the mathematical model (Jakob Asmundsson et al., 2009). As the utilisation increases and the input load reaches the limit of the work centre, the effect of stochastic factors will be more pronounced for the AACF model. In seasonal demand scenarios, the AACF model consistently generates the lowest total cost and performs in due date deviation. Meanwhile, the LUMS COR approach performs best in terms of SFFT. The cost structures of the various approaches are similar, with the cost advantage of the AACF model lying in FGIC. The AACF model shows excellent on-time delivery in exponential demand scenarios, while the LUMS COR approach shows the optimal on-time delivery performance in seasonal demand scenarios.

5.2 Experiment group B

Table 9 shows the results for different assembly centre loads, the first column indicates the different assembly times. As the assembly centre load increases, the probability of the assembly centre being a bottleneck also rises, leading to: (i) a continued increase in the total cost of each approach, (ii) an increase in FGIC and BOC for the AACF model, and a decrease in WIPC, and (iii) a decrease in FGIC for LUMS, and an increase in WIPC and BOC. In addition, the SFTT and σ_{SFTT} of the AACF model show a continuous decrease, while those of the LUMS COR approach continue to increase, and the BILA rule remains stable in terms of due date deviation. The left and right parts of Fig. 3 respectively depict the distribution of due date deviations for assembly times of 80 and 90. The x-axis represents the number of delivery lead time deviations, while the y-axis represents the number of orders.

Table 9
Results for experiment group B

Time	Method	WIPC	FGIC	BOC	SUM	SL	MDD	SDD	SFTT	σ_{SFTT}
60	AACF	1064.4	1345.9	904.9	3315.1	0.89	0.76	0.46	2.47	1.55
	IMM	4883.7	3775.3	905.1	9564.1	0.92	2.07	1.26	6.51	3.36
	BILA	1471.9	2279.1	1224.1	4975.1	0.88	1.27	0.76	1.98	0.93
	LUMS	679.5	2411.9	1412.9	4504.2	0.86	1.35	0.78	1.00	0.43
70	AACF	917.6	1342.9	1609.0	3869.4	0.86	0.80	0.51	2.03	1.35
	IMM	4884.7	3373.8	1725.4	9984.0	0.87	1.91	1.24	6.54	3.38
	BILA	1444.2	1990.5	2176.8	5611.6	0.82	1.18	0.71	1.97	0.96
	LUMS	714.3	2335.3	2088.0	5137.6	0.85	1.36	0.77	1.06	0.43
80	AACF	708.7	1938.1	1551.4	4198.3	0.87	1.09	0.82	1.25	0.59
	IMM	4975.6	2995.7	2043.0	10014.4	0.85	1.72	1.13	6.63	3.44
	BILA	1577.9	1790.6	2244.5	5613.0	0.80	1.06	0.69	2.14	0.92
	LUMS	1038.2	1810.7	2413.7	5262.6	0.79	1.08	0.69	1.48	0.72
90	AACF	680.5	1782.1	3000.1	5462.7	0.78	1.10	0.94	1.12	0.45
	IMM	5089.7	2249.3	5126.5	12465.4	0.71	1.47	1.05	6.98	3.59
	BILA	1557.0	1426.4	4348.0	7331.5	0.69	0.98	0.66	2.31	1.02
	LUMS	1095.4	1405.8	5000.8	7501.9	0.66	1.01	0.66	1.71	0.81

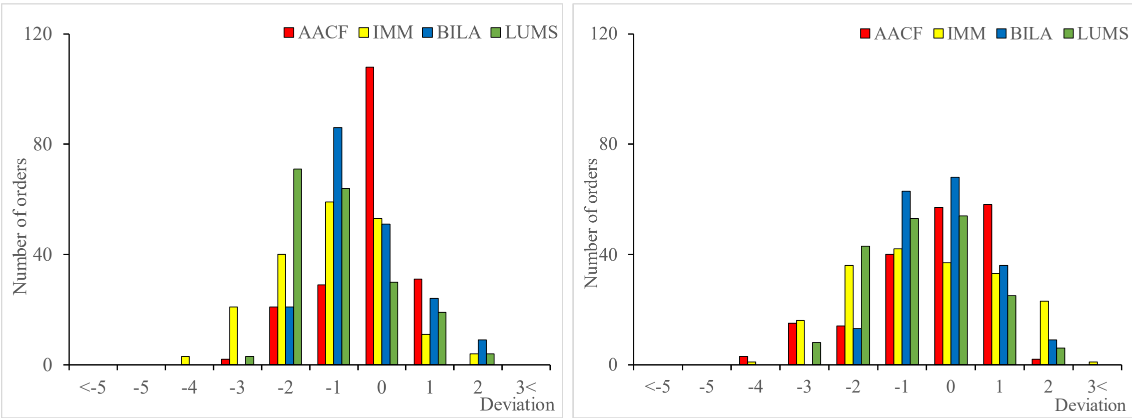

Fig. 3. Order delivery deviation of different approaches

It can be observed that when the assembly time is 80, the delivery of orders is more concentrated for the AACF model. When the assembly time is 90, the number of on-time deliveries for the AACF model decreases, and there is an increase in the number of orders with delivery deviations in the interval [-2,2]. We posit that when the assembly centre serves as a bottleneck, the "order delivery misalignment" phenomenon occurs, where some orders are delivered early while others are delayed. Since the order release approaches only control the decision for releasing parts, it cannot control the sequence in the production system. The "order delivery misalignment" may be caused by the sequence at the assembly centre, but the scheduling issue is beyond the scope of this paper and will not be discussed further. If the scheduling policy is considered, another circular problem arises, i.e., the clearing function depends on the scheduling policy and therefore on the outcome of the scheduling algorithm. On the other hand, the schedules are highly dependent on the release schedule and, thus, the planning algorithm (Asmundsson et al., 2006).

5.3 Sensitivity analysis

In this section, we conduct computational experiments with different cost structures to analyse whether the advantages of the AACF model remain valid. We classify the cost parameters into 3 categories: unit WIPC ($\omega_{jt}^{n}, \pi_{jt}^{n}, \rho_{pt}^{m}$), unit FGIC ($\chi_{pt}$), and unit BOC ($\phi_{pt}$), and perform sensitivity analyses on different approaches. Other experimental conditions: exponential demand, 80% utilisation, deterministic processing time, assembly time of 60 time units, and no failures. To avoid the influence of random factors, we will avoid random events as much as possible and run different parameters using the same demand. We adjusted the values of the above three types of parameters and conducted full factorial experiments, parameter ranges: $\omega_{jt}^{n}, \pi_{jt}^{n}, \rho_{pt}^{m} \in \{1,2,3\}, \chi_{pt} \in \{1,2,3\}, \phi_{pt} \in \{1,2,3\}$. For pithy, only partial results are shown.

Table 10 shows the results for unit WIPC of 2 and 3. The results for unit WIPC of 1 can be found in the previous experiment. It is intuitive and normal for the WIPC and SUM of all approaches to increase significantly when the unit WIPC increases. However, for the AACF model, when $\omega_{jt}^{n}, \pi_{jt}^{n}, \rho_{pt}^{m} = 3$, the FGIC and the due date deviation increases, and the throughput cycle of orders decreases. We believe that as unit WIPC increases, the AACF model reduces WIPC growth by driving parts to product earlier, which also explains the change in the throughput cycle of orders.

Table 10
Result for different unit WIP cost

Method	WIPC	FGIC	BOC	SUM	SL	MDD	SDD	SFTT	σ_{SFTT}
Cost structure: $\omega_{jt}^n, \pi_{jt}^n, \rho_{pt}^m = 2, \chi_{pt} = 10, \phi_{pt} = 100$									
AACF	1786.3	1627.0	461.6	3874.9	0.95	0.89	0.46	2.22	1.57
LUMS	1108.8	3082.1	715.0	4905.9	0.93	1.67	0.88	0.84	0.44
BILA	2770.1	3097.5	328.5	6196.1	0.95	1.66	0.88	1.87	0.95
Cost structure: $\omega_{jt}^n, \pi_{jt}^n, \rho_{pt}^m = 3, \chi_{pt} = 10, \phi_{pt} = 100$									
AACF	1777.1	2405.9	499.9	4682.8	0.93	1.30	0.84	0.97	0.41
LUMS	1663.1	3079.5	727.0	5469.7	0.92	1.67	0.88	0.84	0.44
BILA	4150.2	3097.0	316.0	7563.3	0.95	1.66	0.88	1.87	0.95

Table 11
Result for different unit finished goods inventory cost

Method	WIPC	FGIC	BOC	SUM	SL	MDD	SDD	SFTT	σ_{SFTT}
Cost structure: $\omega_{jt}^n, \pi_{jt}^n, \rho_{pt}^m = 1, \chi_{pt} = 20, \phi_{pt} = 100$									
AACF	873.2	3321.4	306.6	4501.3	0.96	0.89	0.52	2.12	1.48
LUMS	647.3	3882.4	2058.3	6588.0	0.77	1.14	0.71	0.94	0.41
BILA	1354.7	3205.6	3100.6	7660.9	0.75	1.01	0.66	1.83	0.99
Cost structure: $\omega_{jt}^n, \pi_{jt}^n, \rho_{pt}^m = 1, \chi_{pt} = 30, \phi_{pt} = 100$									
AACF	868.1	4984.6	308.7	6161.3	0.95	0.90	0.53	2.05	1.40
LUMS	649.3	5810.0	2099.1	8558.5	0.76	1.14	0.71	0.94	0.42
BILA	1356.1	4796.6	3114.7	9267.4	0.74	1.01	0.66	1.84	0.99

Table 11 shows the results for unit FGIC of 20 and 30. The results for unit FGIC of 10 can be found in the previous experiment. We find that changes in unit FGIC have a smaller effect on the AACF model (increasing FGIC only) and a larger effect on the BILA rule and the LUMS COR approach. As unit FGIC increases, the FGIC and BOC of the BILA rule and the LUMS COR approach also increase significantly, and service levels decrease significantly. We believe that the BILA rule and the LUMS COR approach tend to release parts later when unit FGIC increases, which could lead to a much higher risk of order backorders. Table 12 shows the results for unit BOC of 50 and 150. The results for unit BOC of 100 can be found in the previous experiment. We find that changes in unit BOC have little effect on the AACF model. For the BILA rule and the LUMS COR approach, the increase in unit BOC improves service levels and reduces the risk of delayed delivery of orders. With the above sensitivity analyses, we find that the AACF model exhibits good robustness.

Table 12
Result for different unit backorder cost

Method	WIPC	FGIC	BOC	SUM	SL	MDD	SDD	SFTT	σ_{SFTT}
Cost structure: $\omega_{jt}^n, \pi_{jt}^n, \rho_{pt}^m = 1, \chi_{pt} = 10, \phi_{pt} = 50$									
AACF	882.8	1638.2	201.4	2722.3	0.95	0.89	0.48	2.06	1.42
LUMS	645.8	1940.9	1029.5	3616.2	0.75	1.14	0.71	0.95	0.42
BILA	1357.6	1598.9	1555.1	4511.7	0.75	1.01	0.66	1.84	0.99
Cost structure: $\omega_{jt}^n, \pi_{jt}^n, \rho_{pt}^m = 1, \chi_{pt} = 10, \phi_{pt} = 150$									
AACF	903.5	1585.2	680.2	3168.9	0.95	0.86	0.44	2.17	1.45
LUMS	678.1	3496.2	139.9	4314.3	0.98	1.85	0.91	0.99	0.41
BILA	1384.0	3097.2	485.5	4966.7	0.95	1.66	0.88	1.87	0.95

6. Engineering case

From the experiments in the previous section, it is clear that the AACF model and the LUMS COR approach perform better. Therefore, in this section, the AACF model and the LUMS COR approach are applied in a practical engineering case in a storage tank assembly shop.

6.1 Introduction

The production system is a large storage tank welding assembly job shop, with a long order cycle, multi-species, small batch mixed-flow production, etc. The production system processes 9 types of parts in the processing stage and assembles 3 types of products in the assembly stage. The assembly bill of materials, processing times, the process routes, and the division of work centres are shown in Fig. 4. For example, the processing routes of parts 1, 4, and 7 are work centres 4, 3, 5, and 6, and the processing times of the work centres are 24, 8, 24, and 8 hours, respectively, and 3 parts 1, 1 part 2, 2 parts 3 are assembled into a product 1, which can only be assembled on assembly work centre 1. The original plan for 26 orders is shown in Appendix B, where tank A, tank B and tank C orders are 13, 8 and 5 respectively and the due date (DD) is known. There is a stochastic processing time and failure in this production system.

The original plan derives the release time of each order is obtained by backward scheduling (delivery time production cycle). Due to poor planning, the following problems are caused: (i) Workload imbalance, i.e., equipment is overloaded beyond its normal capacity at certain times and extra shifts are required to complete the overtime work, but the average utilisation of centres is low; (ii) Long waiting times due to uneven loads; (iii) WIP backlogs, and early/late delivery of orders. The workshop uses 2 forklifts for logistics transport and 29 workers to operate its equipment. The simulation experiment verifies that the above resources have little influence on the production planning effect of the production system, so they are ignored.

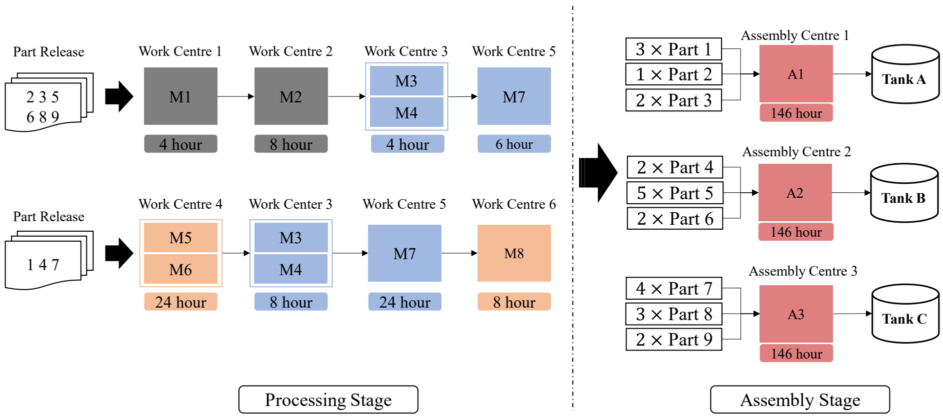

Fig. 4. Process routes for parts and products

6.2 Parameters

The assembly job shop works 5 days a week, 8 hours a day. A planning period is set to 28 days (4 weeks), so there are 32 periods in total. The load and output of each work centre are collected by simulation. Both processing time and assembly time follow normal distributions with means pt_{jt}^{n} and at_{pt}^{m} and variances $pt_{jt}^{n}/10$ and $at_{pt}^{m}/10$. MTTF follows a normal distribution with mean of 28 days and standard deviation of 1 day. MTTR follows a normal distribution with a mean of 1 day and a standard deviation of 2 hours.

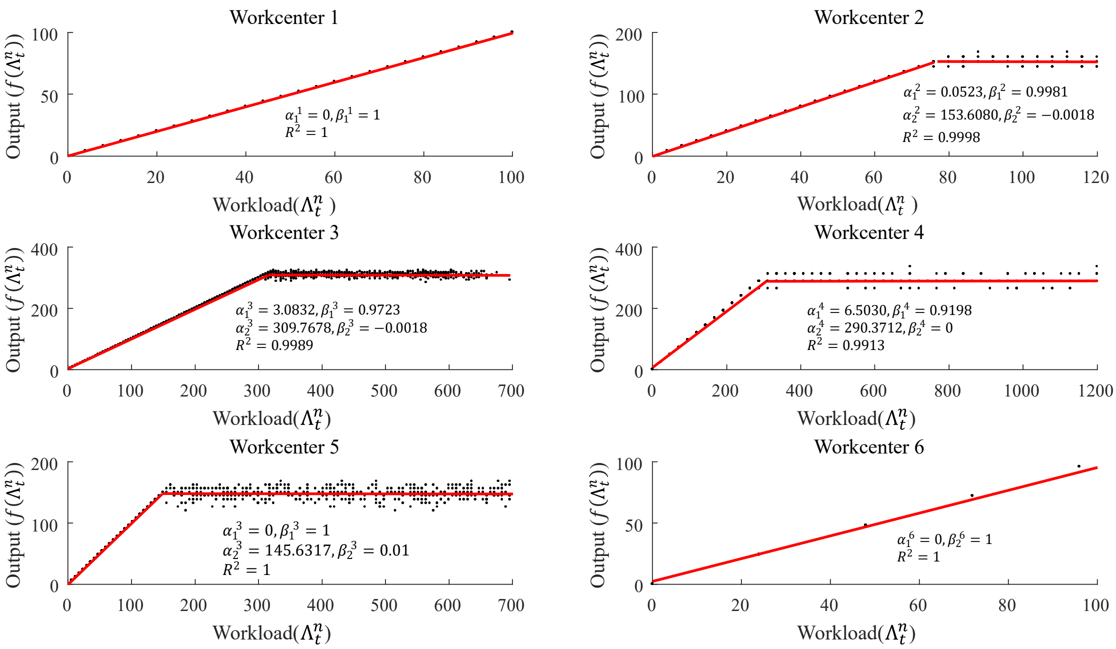

Fig. 5. The segmented linear fitting results for CFs of work centres

The slopes and intercepts of CF are obtained by segmented linear regression. The data of the work centres and the fit of CFs are shown in Fig. 5. The goodness of fit (R^2) of the linear fits are above 0.9, indicating an excellent fit. The parameters of the

LUMS COR approach refer to '4.5 Parameters of approaches'. Theoretically, the slope of the last line segment of the work centre should be 0. However, there are positive or negative values in the graphs, which are related to the error of the sampled data and the small sample size in this paper. However, the slope is generally close to 0, which has little effect on the results.

6.3 Result

The results of the original plan and the optimised plan are shown in Table 13. To avoid the effect of random factors, the simulation is run 20 times, and the average value is taken as the result. All the costs of the AACF model are lower than the others, yielding the lowest total cost. The LUMS COR approach also performs well, while the original plan performs the worst. The AACF model causes virtually no delivery delays and the BOC is virtually zero. The due date deviation of the AACF model also exhibits optimal performance. Consistent with the experimental results, the LUMS COR approach performs well in terms of the throughput cycle of orders.

Table 13
Results for the storage tank welding assembly job shop case

Method	WIPC	FGIC	BOC	SUM	SL	MDD	SDD	SFTT	σ_{SFTT}
Origin	1360.60	2003.06	165.37	3529.03	0.92	7.77	5.76	9.88	5.77
AACF	249.11	225.39	0.44	474.94	1	0.87	0.40	2.71	1.33
LUMS	325.99	308.56	215.58	850.14	0.81	1.27	0.78	2.61	0.92

Fig. 6 shows the variation of WIP over time for different centres with different approaches. The orange, blue and red lines represent the change curves for the original plan, LUMS COR and AACF approaches, respectively. The horizontal coordinate represents the time and the vertical coordinate represents the number of WIP at the centre. The WIP of the original plan exhibits more significant fluctuations, suggesting that the workload in the original plan is concentrated and unbalanced, and the lead time for orders is inaccurately estimated. Consequently, the original plan resulted in an increased WIP and premature order fulfilment. The WIP of the LUMS COR approach and the AACF model are more balanced in their distributions over time. The curves for bottleneck (work centre 5) and assembly centre indicate that the AACF model is superior to the LUMS COR approach.

Fig. 7 shows the Gantt charts of the original plan and the plan optimised by the AACF model, respectively. The blue bar indicates the processing time of the order within the planned completion time; if the order is completed ahead of schedule, the lead time is indicated by the green bar; if the order is delayed, the lead time is indicated by the red bar. The original plan has a total early order completion of approximately 5,500 days, an order delay of approximately 50 days, and an average order cycle time of approximately 280 days. For the LUMS COR approach, the total early order completion is approximately 800 days, order delays are approximately 80 days, and the average order cycle time is approximately 280 days. The AACF model performs better, with a total of approximately 600 days of early order completion, no order extensions and an average order cycle time of approximately 77 days.

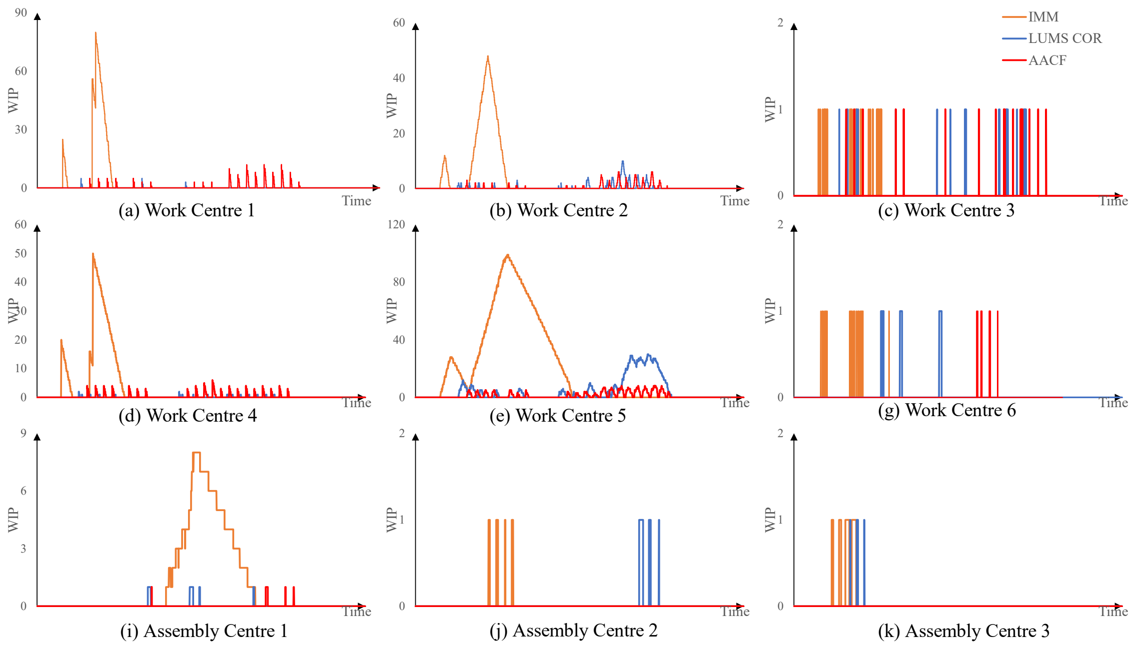

Fig. 6. WIP of different work centres over time

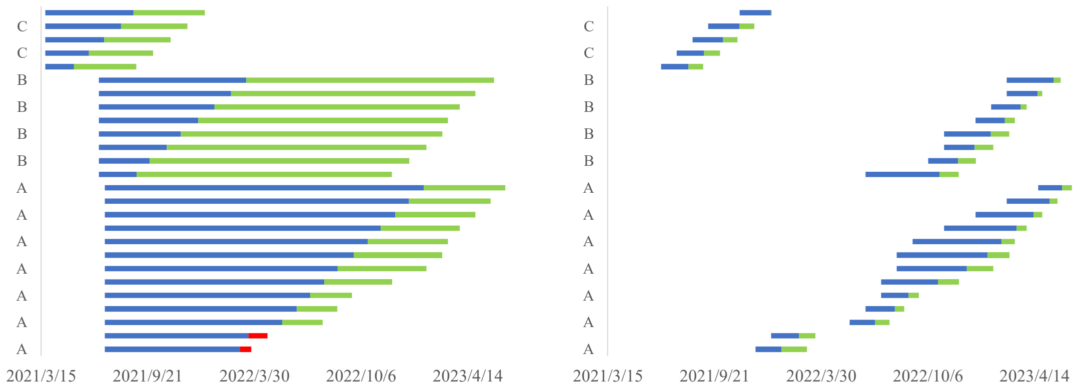

Fig. 7. The Gantt charts of the original plan and the AACF model

In conclusion, the results show that the optimised plan is more reasonable in terms of the distribution of release times and part quantities, and can balance the workload of the centres. Optimised production planning by the AACF model leads to better order lead times and fewer early and late completions. It also highlights the validity of the AACF model proposed in this paper in guiding practical production.

7. Conclusions and implications

A higher level of coordination is required for components in assembly shops compared to general job shops. Prior to this, only a limited number of rule-based order release approaches have been investigated in assembly shops. This paper focuses on an optimisation-based approach to order release in assembly job shop. The AACF model is proposed based on the ACF model by introducing the material flow constraints and the complex assembly structure constraints in the assembly stage, and the clearing functions of centres are fitted segmentally and linearly by simulation data.

In experiments, we designed two groups of experiments (a total of 96 scenarios, 1920 simulations) with factors including different demands, processing time randomness, utilisation, failures and assembly centre load. The experiments compare the proposed AACF model with the LUMS COR approach, the BILA rule, and the IMM rule in different scenarios, perform a sensitivity analysis of the parameters of the objective function, and analyse the optimisation effects under different assembly centre loads. We compared and analysed the impact of different experimental factors on order release approaches in terms of cost (WIPC, FGIC, and BOC) and time (SL, MDD, and SFTT). The results suggest that: (i) the AACF model consistently generates the optimal total cost in all scenarios, primarily due to its advantage in FGIC; (ii) the AACF model exhibits superior service level and due date deviation compared to others, while the LUMS COR approach outperforms in WIPC and SFTT; (iii) the randomness of processing time has a minor impact on various measures, while failure increases the risk of order delays and have a significant impact on costs, service levels and order throughput lead times; (iv) After sensitivity analyses, we find that the cost parameters have an impact on the cost structure, service level and order throughput cycle time of each approach, but overall, the AACF model still performs well and has good robustness; (v) As assembly centre loads increase, the cost structure and order throughput cycle time of the AACF model and the LUMS COR approach are significantly impacted. Additionally, when the assembly centre serves as a bottleneck, the "order delivery misalignment" phenomenon occurs that may be caused by the sequence at the assembly centre, which can also be used as a reference research direction in the future.

Then, the AACF model and the LUMS COR approach are applied in an actual case study of an assembly workshop. The results indicate that the AACF model outperforms in terms of cost and due date deviation, while the LUMS COR approach excels in SFTT. The plan optimised by the AACF model exhibits a more reasonable distribution of release time and part quantities, effectively balancing the workload across work centres. It also highlights the validity of the AACF model proposed in this paper in guiding practical production.

This study provides a new approach to the assembly job shop order release problem, but its limitations should also be noted. (i) Since less research has been done on order release methods in assembly shops, we can only compare the AACF model to a few rule-based mechanisms; (ii) in this paper, only two-stage assembly shop scenarios have been investigated, whereas complex BOM structures and multi-stage assemblies are important features of assembly shops that can be further investigated; (iii) in addition, the presented approach can be verified in more complex production systems with more order types and process routes.

Funding

This work is supported in part by the National Key Research and Development Program of China under Grant No.

2023YFB3307900.

Disclosure statement

The authors report there are no competing interests to declare.

Data Availability Statement

Data is available on request from the authors. The data that support the findings of this study are available from the corresponding author, Haiping Zhu, upon reasonable request.

References

Asmundsson, J., Rardin, R. L., Turkseven, C. H., & Uzsoy, R. (2009). Production planning with resources subject to congestion. *Naval Research Logistics (NRL), 56*(2), 142-157.

Asmundsson, J., Rardin, R. L., & Uzsoy, R. (2006). Tractable nonlinear production planning models for semiconductor wafer fabrication facilities. *IEEE Transactions on Semiconductor Manufacturing, 19*(1), 95-111.

Bechte, W. (1988). Theory and practice of load-oriented manufacturing control. *International Journal of Production Research, 26*(3), 375-395.

Gopalswamy, K., & Uzsoy, R. (2019). A data-driven iterative refinement approach for estimating clearing functions from simulation models of production systems. *International Journal of Production Research, 57*(19), 6013-6030.

Graves, S. (1986). A Tactical Planning Model for a Job Shop. *Operations Research, 34*(4), 522-533.

Guhlich, H., Fleischmann, M., Moench, L., & Stolletz, R. (2018). A clearing function based bid-price approach to integrated order acceptance and release decisions. *European Journal of Operational Research, 268*(1), 243-254.

Haeussler, S., & Missbauer, H. (2014). Empirical validation of meta-models of work centres in order release planning. *International Journal of Production Economics, 149*, 102-116.

Haeussler, S., & Netzer, P. (2020). Comparison between rule- and optimization-based workload control concepts: a simulation optimization approach. *International Journal of Production Research, 58*(12), 3724-3743.

Haeussler, S., Stampfer, C., & Missbauer, H. (2020). Comparison of two optimization based order release models with fixed and variable lead times. *International Journal of Production Economics, 227*, 107682.

Hendry, L., & Kingsman, B. (1991). A Decision Support System for Job Release in Make‐to‐order Companies. *International Journal of Operations & Production Management, 11*(6), 6-16.

Hendry, L. C., & Kingsman, B. G. (1991). Job Release: Part of a Hierarchical System to Manage Manufacturing Lead Times in Make-to-order Companies. *Journal of the Operational Research Society, 42*(10), 871-883.

Hutter, T., Haeussler, S., & Missbauer, H. (2018). Successful implementation of an order release mechanism based on workload control: a case study of a make-to-stock manufacturer. *International Journal of Production Research, 56*(4), 1565-1580.

Kacar, N. B., Irdem, D. F., & Uzsoy, R. (2012). An Experimental Comparison of Production Planning Using Clearing Functions and Iterative Linear Programming-Simulation Algorithms. *IEEE Transactions on Semiconductor Manufacturing, 25*(1), 104-117.

Kacar, N. B., Mönch, L., & Uzsoy, R. (2013). Planning Wafer Starts Using Nonlinear Clearing Functions: A Large-Scale Experiment. *IEEE Transactions on Semiconductor Manufacturing, 26*(4), 602-612.

Kacar, N. B., Mönch, L., & Uzsoy, R. (2016). Modeling Cycle Times in Production Planning Models for Wafer Fabrication. *IEEE Transactions on Semiconductor Manufacturing, 29*(2), 153-167.

Kacar, N. B., & Uzsoy, R. (2014). A comparison of multiple linear regression approaches for fitting clearing functions to empirical data. *International Journal of Production Research, 52*(11), 3164-3184.

Kacar, N. B., & Uzsoy, R. (2015). Estimating Clearing Functions for Production Resources Using Simulation Optimization. *IEEE Transactions on Automation Science and Engineering, 12*(2), 539-552.

Komaki, G. M., & Kayvanfar, V. (2015). Grey Wolf Optimizer algorithm for the two-stage assembly flow shop scheduling problem with release time. *Journal of Computational Science, 8*, 109-120.

Land, M. (2006). Parameters and sensitivity in workload control. *International Journal of Production Economics, 104*(2), 625-638.

Land, M., & Gaalman, G. (1996). Workload control concepts in job shops A critical assessment. *International Journal of Production Economics, 46-47*, 535-548.

Land, M. J., & Gaalman, G. J. C. (1998). The performance of workload control concepts in job shops: Improving the release method. *International Journal of Production Economics, 56-57*, 347-364.

Liu, J., Land, M. J., Bokhorst, J. A., & Chen, Q. (2023). Improving coordination in assembly job shops: redesigning order release and dispatching. *Flexible Services and Manufacturing Journal, 35*(3), 669-697.

Lu, H. L., Huang, G. Q., & Yang, H. D. (2011). Integrating order review/release and dispatching rules for assembly job shop scheduling using a simulation approach. *International Journal of Production Research, 49*(3), 647-669.

Melnyk, S. A., & Ragatz, G. L. (1989). Order review/release: research issues and perspectives. *International Journal of Production Research, 27*(7), 1081-1096.

Missbauer, H., & Uzsoy, R. (2022). Order release in production planning and control systems: challenges and opportunities. *International Journal of Production Research, 60*(1), 256-276.

Oosterman, B., Land, M., & Gaalman, G. (2000). The influence of shop characteristics on workload control. *International Journal of Production Economics, 68*(1), 107-119.

Paul, M., Sridharan, R., & Radha Ramanan, T. (2015). An Investigation of Order Review/Release Policies and Dispatching Rules for Assembly Job Shops with Multi Objective Criteria. *Procedia - Social and Behavioral Sciences, 189*, 376-384.

Portioli-Staudacher, A., & Tantardini, M. (2012). A lean-based ORR system for non-repetitive manufacturing. *International Journal of Production Research, 50*(12), 3257-3273.

Pürgstaller, P., & Missbauer, H. (2012). Rule-based vs. optimisation-based order release in workload control: A simulation study of a MTO manufacturer. *International Journal of Production Economics, 140*(2), 670-680.

Schneckenreither, M., Haeussler, S., & Gerhold, C. (2021). Order release planning with predictive lead times: a machine learning approach. *International Journal of Production Research, 59*(11), 3285-3303.

Silva, C., Stevenson, M., & Thurer, M. (2015). A case study of the successful implementation of workload control. *Journal of Manufacturing Technology Management, 26*(2), 280-296.

Stevenson *, M., Hendry, L. C., & Kingsman †, B. G. (2005). A review of production planning and control: the applicability of key concepts to the make-to-order industry. *International Journal of Production Research, 43*(5), 869-898.

Stevenson, M., Huang, Y., Hendry, L. C., & Soepenberg, E. (2011). The theory and practice of workload control: A research agenda and implementation strategy. *International Journal of Production Economics, 131*(2), 689-700.

Stevenson, M., & Silva, C. (2008). Theoretical development of a workload control methodology: evidence from two case studies. *International Journal of Production Research, 46*(11), 3107-3131.

Thürer, M., Land, M. J., Stevenson, M., Fredendall, L. D., & Godinho Filho, M. (2015). Concerning Workload Control and Order Release: The Pre-Shop Pool Sequencing Decision. *Production and Operations Management, 24*(7), 1179-1192.

Thürer, M., Stevenson, M., & Land, M. J. (2016). On the integration of input and output control: Workload Control order release. *International Journal of Production Economics, 174*, 43-53.

Thürer, M., Stevenson, M., Land, M. J., & Fredendall, L. D. (2019). On the combined effect of due date setting, order release, and output control: an assessment by simulation. *International Journal of Production Research, 57*(6), 1741-1755.

Thürer, M., Stevenson, M., & Silva, C. (2011). Three decades of workload control research: a systematic review of the literature. *International Journal of Production Research, 49*(23), 6905-6935.

Thürer, M., Stevenson, M., Silva, C., & Huang, G. (2012). The application of workload control in assembly job shops: an assessment by simulation. *International Journal of Production Research, 50*(18), 5048-5062.

Thürer, M., Stevenson, M., Silva, C., Land, M., & Filho, M. G. (2013). Workload control and order release in two-level multi-stage job shops: an assessment by simulation. *International Journal of Production Research, 51*(3), 869-882.

Yi-Feng, H., & Leachman, R. C. (1996). A production planning methodology for semiconductor manufacturing based on iterative simulation and linear programming calculations. *IEEE Transactions on Semiconductor Manufacturing, 9*(2), 257-269.

Zhang, Z., Guan, Z., & Yue, L. (2023). Multi-fidelity simulation optimization for production releasing in re-entrant mixed-flow shops. *International Journal of Industrial Engineering Computations, 14*(1), 99-114.

Appendix A

The segmented linear fitting algorithm according to the parameter the number of breakpoints (np) value range of loop traversal, according to the breakpoint algorithm breakpoint set, takes the first np breakpoints for segmented linear fitting. This fitting process is carried out several times to judge the goodness of fit R^2. If $R^2 < RC$, the number of breakpoints will be adjusted, and enter into the next loop and fit again. If $R^2 \geq RC$, the loop is skipped, and the fitting results are output. At the end of the traversal, if there is no qualified R^2, the fitting results corresponding to the largest R^2 in the traversal process are outputted. The parameter RC indicates the requirement for the degree of excellence for the fitted curves, which is generally above 0.9.

Table A1
Algorithm 1: segmented linear fitting algorithm

Segmented linear fitting algorithm

Input: $X0, Y0, np, gs, ge, gap, RC$
$D0 \leftarrow (X0, Y0)$ // D is a point set, $X0, Y0$ are the collected input and output load data
$X \leftarrow merge(X0)$// remove duplicate values
// take the average value to replace the point with the same horizontal coordinates
for $i = 1 : length(X)$ **then**
   // find points with the same horizontal coordinates
   $location = find(D0(:,1) == X(i))$
   // the mean of the vertical coordinates
   $Y \leftarrow mean(D0(location, 2))$
   $P \leftarrow (X(i), Y)$

```
end for
// sort the set of points P in ascending order of horizontal coordinates
// horizontal and vertical coordinates correspond to Λ_t^n and f(Λ_t^n) separately
P ← sort(P)
j ← 0
for np = 1: np do
  for gap = gs: gp: ge do
    bp ← lookpoint(gap, P, np) //invoking algorithm lookpoint
    for i = 1: np do
      // segmented linear regression
      L ← regress(Y(bp(i): bp(i + 1)), X(bp(i): bp(i + 1)))
    end for
    R(j) ← SSR/SST // calculate the goodness of fit
    j ← j + 1
    if R(j) > RC then
      Best_R ← R(j)
      Best_P ← P(j)
    else if (g + gp) > ge && np = np_num then
      Best_R ← max(R)
      Best_P ← P(location(Best_R))
    end if
  end for
end for
```

The breakpoint function algorithm solves for the difference in slope of the set of neighbouring points, where the difference is greater than a certain set value is considered a breakpoint, and solves for and stores all the breakpoint locations to get the set of breakpoints.

Table A2
Algorithm 2: breakpoint function algorithm

```
Breakpoint function algorithm: breakpoint
Input: gap, P, np
j ← 0
for ps = 1: gap: length(P) do // step size is gap
  // get each group points, find the mean value and store it
  Avg_{j,1} ← mean(P(ps : ps + gap, 1))
  Avg_{j,2} ← mean(P(ps : ps + gap, 2))
  j ← j + 1
  end for
for i = 2 : length(Avg) - 1 do
  // for 3 neighboring points, calculate the slope of 2 of the points
  s_1 ← (Avg_{i,2} - Avg_{i-1,2})/(Avg_{i,1} - Avg_{i-1,1})
  s_2 ← (Avg_{i+1,2} - Avg_{i,2})/(Avg_{i+1,1} - Avg_{i,1})
  S_i ← |s_1| - |s_2|
end for
// sort the set of points S in descending order
S1 ← sort(S)
// take the first np points with the largest values as breakpoints
l ← location(S1(1: np)) // l is an indexed set
breakpoints ← P(index)
Output breakpoints
```

Appendix B

Table B1 Original production plan of the assembly shop

No	Type	PRD	DD	No	Type	PRD	DD
1	A	2021\7\7	2022\3\5	14	B	2021\6\26	2022\11\30
2	A	2021\7\7	2022\3\20	15	B	2021\6\26	2022\12\31
3	A	2021\7\7	2022\7\30	16	B	2021\6\26	2023\1\31
4	A	2021\7\7	2022\8\25	17	B	2021\6\26	2023\2\28
5	A	2021\7\7	2022\9\20	18	B	2021\6\26	2023\3\10
6	A	2021\7\7	2022\11\30	19	B	2021\6\26	2023\3\31
7	A	2021\7\7	2023\1\31	20	B	2021\6\26	2023\4\28
8	A	2021\7\7	2023\2\28	21	B	2021\6\26	2023\5\31
9	A	2021\7\7	2023\3\10	22	C	2021\3\23	2021\9\1
10	A	2021\7\7	2023\3\31	23	C	2021\3\23	2021\10\1
11	A	2021\7\7	2023\4\28	24	C	2021\3\23	2021\11\1
12	A	2021\7\7	2023\5\25	25	C	2021\3\23	2021\12\1
13	A	2021\7\7	2023\6\20	26	C	2021\3\23	2022\01\01