Chapter 2

Literature Review

An overview of relevant and recent literature focusing on the defined research field in Chapter 1 is presented in this chapter. Due to the huge amount of literature only the set of literature which forms the basis for the developed models is discussed.

2.1 JIT Goals

Section 2.1 of this review covers the literature on key performance indicator utilisation and JIT philosophy in a hierarchical production planning setting. The reviewed literature lacks of a differentiated investigation on utilisation for each hierarchical production planning level for a production system avoiding waste in terms of a JIT perspective.

2.1.1 Hierarchical Production Planning Models

In this section some basic formulations on how production planning decisions are split up into different hierarchical levels are presented. In 1975 Hax and Meal (1975) formalise a hierarchical production planning approach. Bitran and Hax (1977) suggest an iterative procedure to optimise the subproblems defined in Hax and Meal (1975). Wight (1984) develops the hierarchical production planning method MRPII (Manufacturing Resource Planning), which implements the three planning levels long, medium and short term (see Figure 2.1). The basis for MRPII is the closed loop MRP approach invented by Orlicky (1975). The MRP concept has some shortfalls, such as capacity infeasibility, long planned lead times and system nervousness. The integrated hierarchical planning system MRPII should eliminate these flaws (Hopp and Spearman, 2008; Jodlbauer, 2008b).

A. Hübl, *Stochastic Modelling in Production Planning*,
DOI 10.1007/978-3-658-19120-7_2

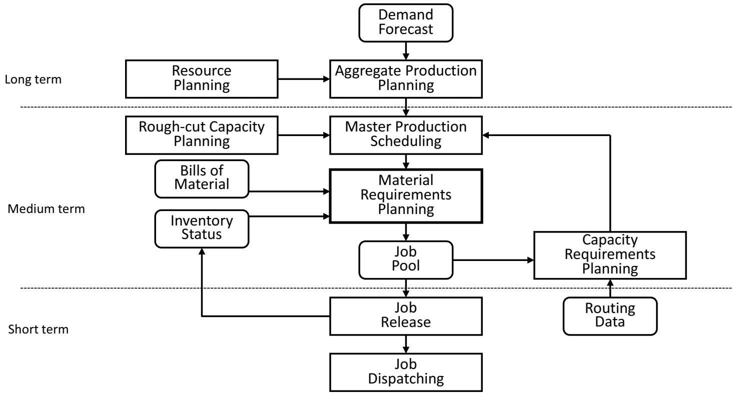

Figure 2.1: Manufacturing resource planning MRPII (Wight, 1984; Hopp and Spearman, 2008)

The long term level basically influences capacity investment decisions based on marketing parameters and results in rough prediction about future production mix and volume. Long term planning covers three activities: demand forecast (predicting future demand), resource planning (determines the required capacities over the long term horizon) and aggregate production planning (determines the level of production, staffing, inventory on part families).

The medium term level combines the information from the long term planning along with information about customer orders. The prediction about production mix and volume is translated into production orders with related due dates (Vollmann et al., 1997; Hopp and Spearman, 2008). The Master Production Schedule (MPS) converts the long-term aggregated forcasted demand into a detailed forecast while tracking individual customer orders. The rough-cut capacity planning provides capacity checks of a few critical resources to ensure the feasibility of the MPS. MRP (Orlicky, 1975) translates the demands resulting from the MPS into production orders with related due dates by the use of Bill-Of-Materials (BOM), routing data and the inventory status. Hopp and Spearman (1996, 2008) implement pull strategies such as Conwip (Constant Work-In-Process) (Spearman et al., 1990) or Drum Buffer Rope (DBR) (Goldratt, 1988) into the MRPII concept. Additionally, some methods are used that are not based on production plans, such as Kanban (Ohno, 1988) and Reorder Policies (Arrow et al., 1951). The result of medium term level are production plans (Job Pools) stating the work process that need to be carried out (Hopp and Spearman, 2008). These production plans are based on presumptions and expectations concerning a number of factors, such as capacity, production mix, equipment, inventory, lead times, disturbances, scrap and personnel (Wedel and Lumsden, 1995). Capacity requirements planning provides a more detail capacity check than the rough-cut capacity planning and generates loading profiles based on the production plans. The production orders can be based on customer orders (Make-To-Order, MTO) or on forecasts (Make-To-Stock, MTS) if the customer required lead time is likely to be shorter than the production lead time (Jodlbauer, 2008c; Hübl et al., 2010; Altendorfer et al., 2014).

According to Hopp and Spearman (2008), the short term level controls the real-time flow of materials (Job Release) based on the work schedule developed on the medium term level. Dispatching rules are often used to decide when which production lot has to be produced on which machine (Job Dispatching) (Panwalkar and Iskander, 1977; Blackstone et al., 1982).

The hierarchical production planning model has been extended in the past decades a lot and has become an important method in production planning. Graves (1982) uses Lagrange multipliers to solve a mixed integer linear program based on Hax and Meal's (1975) hierarchical production planning approach. Kok (1990) discusses the mathematical logic of a hierarchical planning approach and proposes a procedure to compute the aggregate production volume and the allocated quantities in order to achieve a target service level.

Gfrerer and Zäpfel (1995) propose robust production plans for the aggregated planning level and a disaggregation method for the lower level for uncertain demand. Zäpfel (1996) develops a hierarchical model for uncertain demand, which can be incorporated into the MRPII concept.

Mula et al. (2006) provide a literature review for production planning models under uncertainty. Rafiei et al. (2013) investigate the medium term and short term production planning level in a hierarchical production planning approach. They develop an MTO/MTS hybrid production planning strategy, which is modelled as a meta-heuristic algorithm. Chakrabortty and Hasin (2013) explore the aggregated production planning with forecasted demand, related operating costs and capacity for a multi-product, multi-period model. Their approach minimises total costs including inventory levels, labor levels, overtime, subcontracting and back ordering levels, as well as labour, machine and warehouse capacity. Jansen et al. (2013) investigate the non-linear relationship between Work-In-Process (WIP) of a production unit and lead time in a hierarchical production planning setting. They present a two-step lead time anticipation procedure where a Linear Program (LP) is solved regardless of the production capacity available. A local smoothing heuristic is applied which tracks the stochastic workload during the planning horizon.

Yang and Fung (2014) use a hierarchical structure to solve an ATP decision model for a multi-site MTO production supply chain. Ponsignon and Mönch (2014) combine medium and short term levels in their simulation study to evaluate the performance of their master planning approach for the semiconductor industry. Moreover, the authors apply the methodology of reduced simulation models (Hung and Leachman, 1999) to speed up the run time of the experiment.

2.1.2 Utilisation in Production Planning

The classical long term objective is to minimise the capacity invested or equivalently to maximise the utilisation. In the medium term, capacity adjustments are made resulting from demand fluctuations. In the short term view, minimising the utilisation means to produce efficiently because the losses according to the "seven zeros" philosophy are avoided. In the following literature, opposing goals on each hierarchical planning level are reviewed.

According to Hopp and Spearman (2008) utilisation is controversial because on the one hand high utilisation leads to low costs per unit but on the other hand low utilisation allows high sales. Bradley and Glynn (2002) show that higher capacity invested (this means lower utilisation) allows less inventory. Mieghem and Rudi (2002) and Angelus and Porteus (2002) address the joint (inventory vs. capacity investment) decision problem. Jodlbauer and Altendorfer (2010) conclude that there is an impact of the utilisation on the cost for capacity available and inventory needed. Especially in the case of a small ratio of unit holding cost over unit capacity cost, the costs are considerably increased if the utilisation is only a little higher or smaller than optimal utilisation. Carrillo and Gaimon (2000) investigate the manufacturing performance through process changes and knowledge creation with an optimal control model. They argue that process changes may lead to long term increase in effective capacity but during implementation typically reduce short term capacity. After reviewing relevant literature it becomes clear that on each hierarchical level, utilisation is treated differently.

2.1.3 Impact of JIT Activities on Performance

Many industries are facing strong global competition because product life cycles are shortened, time-to-market decreases and customers require fast deliveries of a variety of products of an appropriate quality. Therefore, it is absolutely necessary that a company ensures that the right product of the right quality is available to the customer in the right quantity at the right time. Companies applying lean practices such as Just-In-Time (JIT), Total Productive Maintenance (TPM) or Total Quality Management

(TQM) contribute substantially to the operating performance (Ahuja and Khamba, 2008; Emde et al., 2012; Shah and Ward, 2003; Aspinwall and Elgharib, 2013).

Various empirical studies examine the influence of JIT activities on a company's performance. The results of Sim and Killough (1998) provide empirical evidence that performance gains from synergies of JIT activities resulting from combining JIT activities with performances goals. In a survey of US manufacturing companies White et al. (1999) identify ten JIT activities that are appropriate for implementation. Changes in performance depend on the degree of JIT implementation and the company size. White and Pearson (2001) propose how JIT activities can be implemented into the decision making process of manufacturing companies.

According to their empirical study covering the Canadian automotive parts manufacturing industry, Callen et al. (2005) show that JIT intensive plants have more capacity waste than other plants, but they generate more profit. Inman et al. (2011) observe the influence of JIT activities on a firm's manufacturing agility with their structural equation model. They argue that if JIT activities in the manufacturing processes are already implemented in a company, then an increased supplier/ customer integration could show a greater impact on agility than JIT activities alone. Obermaier and Donhauser (2012) analyse the financial performance of companies as a function of inventory holding in their empirical study. They identify a positive relation between inventory holding and financial performance. According to them, those companies with highest inventory show the best financial performance and vice versa. Their findings indicate that the core principle of JIT, avoiding inventory, has its limits.

JIT is a philosophy with the primary goal of continuously reducing all forms of waste (Sugimori et al., 1977; Ohno, 1988; Golhar and Stamm, 1991; Daugherty et al., 1994). Suzaki (1987) identifies waiting time, transportation, processing, inventory and motion as possible forms of waste. According to Brown and Mitchell (1991) there are two major forms of waste: high inventories and unnecessary delays. Daugherty et al. (1994) define six prerequisites for continuous improvements: maximum equipment availability, TQM, minimum changeover, optimum physical layout, multi-trained workforce and standardised operations.

TPM literature provides different concepts for classifying waste. Nakajima (1988), the founder of TPM, describes six waste types (equipment failure; set-ups and adjustments; idling and minor stoppages; reduced speed; quality defect and rework; reduced yield). Golhar and Stamm (1991) give an extensive literature review about JIT and identify eleven critical variables for eliminating waste. Ljungberg (1998) identifies in his study that many sources of possible loss are not tackled by companies and he suggests the use of a comprehensive model of loss management. Muchiri and Pintelon (2008) give an overview of the performance measurements of TPM.

Already in 1983 Edwards (1983) has introduced the ”seven zeros” – zero defects, zero lot size, zero set-ups, zero breakdowns, zero handling, zero lead time, zero surging – as JIT goals, which have to be achieved to eliminate all forms of waste, especially inventories. The term zero should express that these figures should be continuously minimised.

2.2 Capacity Setting Methods for Medium Term Capacity Planning

Section 2.2 is dedicated to literature covering capacity setting problems. In the beginning, capacity investment problems are reviewed. This is followed by a discussion about decision problems with capacity expansion and/or reduction. Finally recent queuing state depended models are introduced.

On the one hand the due dates of the customers can be negotiated to create a smoother capacity demand (Hopp and Roof Sturgis, 2000; Hegedus and Hopp, 2001; Keskinocak and Tayur, 2004; Corti et al., 2006). On the other hand the capacity can be adjusted to the fluctuations of the customer demand (Kok, 2000; Bradley and Glynn, 2002; Mieghem and Rudi, 2002; Defregger and Kuhn, 2006; Li et al., 2009; Mincsovics and Dellaert, 2009; Buyukkaramikli et al., 2013). The methods discussed in this section are based on capacity adjustment literature whereby a flexible capacity with upper and lower bounds is assumed. Therefore, this literature review focuses on the capacity adjustment literature stream. Finally, the reviewed literature lacks of a periodical decision support for short and medium term capacity setting for improving service level and tardiness whereby stochastic customer behaviour and stochastic production process is assumed.

Capacity expansion problems are first studied in capacity investment literature (Chenery, 1952; Manne, 1961; Luss, 1982; Kok, 2000; Pibernik and Yadav, 2009). Chenery (1952) and Manne (1961) assume a deterministic increasing demand and whenever demand exceeds the capacity available, the capacity is expanded. Manne (1961) includes probabilities instead of a constant rate of growth in demand and backlogs to the model of Chenery (1952). Luss (1982) conducts an extensive literature review about capacity expansion problems and classifies capacity investment problems in several categories emphasising modelling approaches and algorithmic solutions. Segerstedt (1996) develops a capacity constrained multi-stage inventory and production control problem. He minimises the inventory costs and shortage costs, whereby the cumulated capacity concept is applied as constraint. The cumulated demanded capacity is not allowed to exceed the cumulated provided capacity. Kok (2000) compares two capacity allocation strategies.

A fixed capacity is assumed and if the demand exceeds the fixed capacity the orders are delayed. Moreover, an additional capacity is introduced by hiring additional personnel.

Decision problems with capacity expansion and/or reduction are modelled in most cases as dynamic programs (Mieghem and Rudi, 2002; Bradley and Glynn, 2002; Li et al., 2009). Bradley and Glynn (2002) develop an analytic model for a single-machine and single product system which describes the optimal long term balance between capacity and inventory. They show that optimal inventory policy varies with capacity investment and that higher capacity invested allows less inventory. Moreover, the authors describe how inventory should be optimally substituted for capacity in order to minimise costs when the capacity level varies. Mieghem and Rudi (2002) and Angelus and Porteus (2002) deal with this issue in a more general situation and gain similar results. Obviously there is a trade off between capital invested in capacity and costs of the capital employed in inventories. Defregger and Kuhn (2006) develop an MTO order acceptance model with limited inventory capacity, where it is possible to adjust the inventory levels. In Li et al. (2009), capacity allocation methods with mixed integer programming methods are compared for supply chain optimisation. The authors identify that an integrated planning approach achieves better results than an approach where each subproblem is treated separately.

The MTO ability of production systems is evaluated in Jodlbauer (2008c) depending on the capacity provided, the customer required lead time distribution, and the demand fluctuation. The result of this evaluation shows that applying a capacity adjustment method which enables the reaction to short term peaks can decrease the capacity provided for MTO environment. Jodlbauer and Altendorfer (2010) present a concept for optimising the overall capacity provided, for which they use the customer required lead time distribution too. The result of this paper indicates that flexible capacity on a short and medium term basis can lead to a cost decrease.

The papers of Balakrishnan et al. (1996) and Balakrishnan et al. (1999) discuss the capacity rationing problem for a two-product production system whereby one product class leads to higher profits per unit. An order rejection policy for the lower profit products to maximise company profit is described. In Kok (2000), capacity allocation is discussed where capacity has to be allocated to different product groups while minimising a total cost function. In this model the production is triggered by an order-up-to policy for each product group.

Yang and Fung (2014) present an ATP decision support for order acceptance/selection, due date assignment and order scheduling in a multi-site MTO production supply chain system. They propose a two-stage hierarchical structure.

In recent research queuing state dependent capacity adjustment models focusing on the transient behaviour of the queuing system when switching between different capacity levels are studied (Mincsovics and Dellaert, 2009; Buyukkaramikli et al., 2013). In Mincsovics and Dellaert (2009) a continuous setting is discussed in which an up-switching-point and a down-switching-point are identified and each switch incurs costs. A periodic setting with two possible capacity levels has been studied extensively in Buyukkaramikli et al. (2013).

2.3 Conwip

Section 2.3 summarises literature about the production planning method Conwip, whereby the literature is distinguished between analytical models and simulation. The production planning and control method Conwip is firstly introduced by Spearman et al. (1990) and Spearman and Zazanis (1992). The basic parameters Wipcap, work-ahead-window and capacity trigger are already introduced in these papers. The method has since then been further developed and compared to the most dominant production planning and control methods MRP and Kanban either analytically or with simulation. For a detailed review on Conwip and its applications see Framinan et al. (2003). Prakash and Chin (2014) review 15 modified Conwip systems and classify them according to: (I) the feedback signal from demand source triggers production, (II) the products are pushed between workstations and (III) the WIP level (Wipcap) is limited within workstations.

Looking at the literature discussed in the review below shows that Conwip performs well in comparison to other traditional production planning and control methods. However, the implementation of Conwip in an MTO system is yet rarely discussed as mostly average inventory and throughput are compared. Especially research on the question of improving service level and tardiness measures by changing the Wipcap definition itself and the evaluation of a safety stock implementation in Conwip is still not available.

Analytical models usually describe the behaviour and performance of Conwip in comparison to Kanban (or other production planning and control methods) or they deliver approximations for basic logistical key performance indicators, such as throughput, WIP and production lead time when Conwip is applied to a more complex production system.

The first paper introducing Conwip, Spearman et al. (1990) model a multi-stage serial production system as closed loop queuing network. Moreover, the authors compare its performance to an open loop queuing network representing the push production control method.

Similarly the superior performance of Conwip in comparison to Kanban and push methods as MRP is proven in Spearman and Zazanis (1992).

Hence, Gstettner and Kuhn (1996) demonstrate that Kanban outperforms Conwip due to the higher number of optimiseable parameters. Geraghty and Heavey (2004) show that the optimal hybrid push/pull method proposed in Hodgson and Dingwei Wang (1991a) and Hodgson and Dingwei Wang (1991b) modelled as a Markov decision process is a Conwip system.

Herer and Masin (1997) develop an optimisation problem formulation where the right order in which the parts should be produced in a Conwip system is optimised.

An accurate approximation of the average production rate in a closed loop system for more complex production system structures than discussed in the papers mentioned above is developed in Gershwin and Werner (2007) which can also be applied to a Conwip system. Furthermore, Heragu et al. (2011) develop an approach to identify the waiting time for semi-open queuing networks, also taking customer waiting time in front of the system into consideration. Such semi-open queuing networks can be used to model Conwip.

An analytical method for determining Wipcap and Work-Ahead-Window (WAW) of Conwip based on the customer required lead time distribution is presented in Jodlbauer (2008a).

Helber et al. (2011) optimise the production rate and/or short term profit of a Conwip system by the use of mixed integer linear program in combination with a stochastic simulation. The accuracy of their method proposed depends on the variability of the processing times and the WIP.

Satyam and Krishnamurthy (2013) model the batch size constraints for a Conwip system as a multi-class closed queuing network with synchronisation stations. Based on a routing matrix for each station a random variable for set-up time and processing time is modelled. The authors use a decomposition approach as approximation because exact methods are hard to solve. They identify that batch size decisions have an impact on average production lead time. Practitioners can decide whether they adjust batch sizes or Wipcap. Park and Lee (2013) study a multi-product Conwip production system with a Poisson arrival process with a fixed probability that an order requires a particular item made of different subcomponents. The authors also apply a decomposition approach for modelling and solving.

Lagershausen et al. (2013) propose an approximation for the throughput of a closed loop queuing network with generally distributed processing times.

Gong et al. (2014) quantify and measure information to compare the information amounts in MRP, Kanban and Conwip to study how information amount affects the decision-making delay. They argue that MRP requires the largest amount of information since the WIP in an MRP system is higher compared to Kanban and Conwip.

Simulation studies either discuss extensions of Conwip and their performance, which is also the objective of this section since the integration of FGI into the Wipcap for MTO production systems is an extension, or they compare the performance of Conwip to the performance of other production planning and control methods.

Simulation studies which compare the performance of Conwip to that of MRP, Kanban and/or DBR are Huang et al. (1998b), Huang et al. (1998a) and Jodlbauer and Huber (2008), all stating that Conwip outperforms the other methods. Furthermore, Gilland (2002) and Grosfeld-Nir and Magazine (2002) compare the performance of Conwip to their developed methods by simulation. They mainly find that Conwip outperforms MRP and DBR, although their methods outperform Conwip in the respective environments. Sepehri and Nahavandi (2007) also compare Conwip to other WIP constraining work release methods and find that it is partly outperformed by them. Gstettner and Kuhn (1996) identify that Kanban leads to a lower average WIP than Conwip for a given production rate. The WIP levels in the buffers between the working stations of a Kanban system show the "WIP bowl phenomenon" while in the Conwip system the WIP has equal values in all stations except at the bottleneck station where the WIP accumulates.

Duenyas et al. (1993) model a Conwip system with deterministic processing times, exponential failure and repair times as closed queuing network. Moreover the authors derive mean and variance of the output. By empirical tests, they show that their approach leads to robust solutions which can be the basis for selecting production quota and Wipcap for a Conwip line.

Framinan et al. (2000) develop a new Conwip card setting method, which leads to card counts per product, and discuss the performance of this card setting method with the application of different dispatching rules. A simulation study shows that this card setting method leads to decreased inventories and increased service level for an MTS production system.

A dynamic card setting method for the Wipcap is presented in Hopp and Roof (1998) whereby the number of cards is identified by statistical process control applied to the throughput. They test the performance of this method by simulation and find that it is robust against environmental changes.

Bahaji and Kuhl (2008) conduct a simulation study comparing the influence of different dispatching rules on logistical performance in a wafer fabrication facility. For work release they compare Conwip to a push method which is a simplified MRP setting. They find that their simplified MRP method outperforms Conwip when applying some multi-objective composite dispatching rules developed in their paper. However, they also find that these rules do not perform significantly better than First-Come-First-Serve (FCFS) in a Conwip setting which supports the recommendation of

Spearman et al. (1990) and Hopp and Spearman (2008) to apply First-In-System-First-Served (FISFS) within a Conwip production system.

Hübl et al. (2011) develop a simulation model where Conwip is embedded in a hierarchical production planning structure as proposed by Hopp and Spearman (2008).

Chong et al. (2013) investigate an integrated Kanban and Conwip production system by the use of discrete event simulation. They use total output, average WIP, average production lead time and average bottleneck utilisation as performance metrics.

Huang et al. (2015) consider a Conwip assembly production system for mass production with multi-loop, multi-products, low volume and one-of-a-kind production environments. The authors develop eight Conwip loop control policies based on five basic design patterns. The result of the simulation study shows that the multi-loop design performs better than a single-loop. Moreover, they have also proposed a WIP upper bound heuristic algorithm for searching the optimal Wipcap.

2.4 Dispatching and Production Lead Time

In Section 2.4 some literature is reviewed where dispatching rules influence average production lead time. The goal of scheduling in production planning is to generate a profitable balance between conflicting objectives in production planning. Most real-world scheduling applications are NP-hard (Non-deterministic Polynomial-time hard) problems and therefore in practice dispatching rules are applied. Dispatching rules consider only local and current states of the production system and neglect further jobs. However, dispatching rules effect the production lead time (Hopp and Spearman, 2008). The reviewed literature lacks an analytical relationship between dispatching rules and production lead time.

Short production lead times offer several advantages. According to Little's Law (Little, 1961), shorter production lead times also result in lower WIP levels at the same utilisation and therefore less capital is employed.

Schonberger (1986) uses production lead time reduction to classify companies into different levels of world class manufacturing. According to Neely et al. (1995) and Wacker (1996) manufacturing performance is defined in terms of quality, delivery speed, delivery reliability, price (cost) and flexibility, whereby delivery speed includes the production lead time. Merschmann and Thonemann (2011), Yang et al. (2011) and Inman et al. (2011) include the metrics "reduction on production lead time" for measuring manufacturing flexibility in their structural equation model.

Production lead times for assembly processes can vary because of queuing or transportation delays. Moreover, variable set-up times and variable processing times result in a distribution of the production lead time (Yano, 1987).

Wedel and Lumsden (1995) show in their case studies that the production planning process is affected by reductions in the total manufacturing lead time because the planning department has a lack of confidence in the manufacturing department or factory.

Altendorfer and Minner (2011) investigate a two-stage MTO production system with random demands, processing times and distributed customer due dates. They minimise total inventory holding and customer order tardiness costs to identify the optimal manufacturing capacities and planned lead times for each stage. As a result they have proven that the distribution of the customer required lead time has no influence on the optimal planned lead times whenever capacity is predefined.

Hayya et al. (2011) present a procedure for reducing the mean and variance for exponentially distributed lead times.

Furthermore, shorter production lead times allow quicker responses to market demand changes (Altendorfer and Jodlbauer, 2011). Average production lead time is influenced by the decisions of production planning and control. For short term scheduling, especially dispatching rules are often applied and so their influence on average production lead time is a subject of research (see the reviews of Blackstone et al. (1982); Waikar et al. (1995); Rajendran and Holthaus (1999) as well as the book of Pinedo (2008)). Dispatching rules are used to select the next order to be processed from waiting orders in front of the processing station.

Bednowitz et al. (2014) have tested different dispatching and loitering policies for unmanned aerial vehicle systems. This military application examines a system with multiple unmanned aerial vehicle systems assigned to respond to fixed-location,
multiple priority targets. The aim is to ensure rapid service for high priority targets and effective management of medium priority levels. The authors identify that dispatching rules with a think-ahead[1] and distance-based focus lead to best results.

Various simulation studies examine the influence of dispatching rules on average production lead time. Barrett and Kadipasaoglu (1990) show for a dynamic flow shop that the Shortest Processing Time (SPT) rule performs best in terms of average production lead time. Waikar et al. (1995) compare ten different dispatching rules under different shop loads in their simulation study. Evaluating the average production lead time of jobs, again the SPT rule performs best and First-In-First-Out (FIFO) is ranked fourth. Hung and Chen (1998) show that Shortest Remaining Processing Time (SRPT)

[1] According to Mantel and Landeweerd (1995) think-ahead implies that known tasks are combined to routes such that idle (empty) travel time is minimised.

and Earliest Due Date (EDD) are good dispatching rules to reduce average production lead time in semiconductor wafer fabrication. Land et al. (2014) focus on overcoming the conflict between order release and dispatching. The authors find out that operating due date dispatching rules are beneficial.

Jayamohan and Rajendran (2000), El-Bouri et al. (2008) and Chen and Matis (2013) compare new dispatching rules with standard rules with regard to different performance measures such as average, maximum and variance of production lead time and tardiness. The results of Jayamohan and Rajendran's (2000) simulation study show that SPT is the rule with the shortest average production lead time for flow shops.

Whenever simulation is applied to discuss the effect of dispatching rules on the average production lead time, the focus is mainly on comparing and evaluating the performance of dispatching rules and developing new dispatching rules. Very little literature is available on the development of analytical models based on simulation and empirical studies. One such stream is based on the application of the funnel model for single-stage models (Wiendahl et al., 1994; Wiendahl and Breithaupt, 1999) where approximation functions for production lead time and inventory are identified applying empirical data for their parametrisation. Based on this funnel model, Nyhuis and Wiendahl (2009) derive approximations for average production lead time using SPT and Longest Processing Time (LPT) as dispatching rules.

From the research work reviewed above, it is obvious that most literature dealing with the interaction between dispatching rules and production lead time either applies simulation or empirical data and approximation equations. Nevertheless, queuing theory can also be applied.

An analytical model based on queuing theory shows that the SPT rule minimises the expected production lead time in a static job shop with deterministic processing times (Buzacott and Shanthikumar, 1993; Hopp and Spearman, 2008).

The work by Nyhuis and Wiendahl (1999) and Nyhuis and Wiendahl (2009) introduce the range, which is the "processing time weighted average production lead time" in comparison to the arithmetic average production lead time. They show that the range is independent of the dispatching rule applied for a single-stage production system. The papers by Jodlbauer (2005) and Jodlbauer and Stöcher (2006) extend the framework for continuous input and output functions in a single-stage production system.

http://www.springer.com/978-3-658-19119-1

Stochastic Modelling in Production Planning
Methods for Improvement and Investigations on
Production System Behaviour
Hübl, A.

2018, XV, 139 p. 19 illus., 12 illus. in color., Softcover
ISBN: 978-3-658-19119-1