ΠΡΑΞΗ: «ΕΝΙΣΧΥΣΗ ΜΕΤΑΔΙΔΑΚΤΟΡΩΝ ΕΡΕΥΝΗΤΩΝ/ΕΡΕΥΝΗΤΡΙΩΝ» MIS 5001552

ΑΝΑΣΚΟΠΗΣΗ ΜΕΤΑΔΙΔΑΚΤΟΡΙΚΗΣ ΕΡΕΥΝΑΣ: «ΜΟΝΤΕΛΟΠΟΙΗΣΗ, ΕΛΕΓΧΟΣ ΚΑΙ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗ ΣΤΟΧΑΣΤΙΚΩΝ ΣΥΣΤΗΜΑΤΩΝ ΠΑΡΑΓΩΓΗΣ»

ΜΕΤΑΔΙΔΑΚΤΟΡΑΣ ΕΡΕΥΝΗΤΗΣ:
ΑΛΕΞΑΝΔΡΟΣ ΞΑΝΘΟΠΟΥΛΟΣ (ΤΟΥ ΣΩΤΗΡΙΟΥ)

ΕΠΙΣΤΗΜΟΝΙΚΟΣ ΤΟΜΕΑΣ:
ΦΥΣΙΚΕΣ ΕΠΙΣΤΗΜΕΣ ΚΑΙ ΕΠΙΣΤΗΜΕΣ ΜΗΧΑΝΙΚΟΥ

ΙΔΡΥΜΑ ΥΛΟΠΟΙΗΣΗΣ ΕΡΕΥΝΑΣ:
ΔΗΜΟΚΡΙΤΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΡΑΚΗΣ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΑΡΑΓΩΓΗΣ ΚΑΙ ΔΙΟΙΚΗΣΗΣ
ΕΡΓΑΣΤΗΡΙΟ ΒΙΟΜΗΧΑΝΙΚΗΣ ΠΑΡΑΓΩΓΗΣ

ΕΠΙΒΛΕΠΩΝ ΚΑΘΗΓΗΤΗΣ:
ΔΗΜΗΤΡΙΟΣ ΚΟΥΛΟΥΡΙΩΤΗΣ, ΚΑΘΗΓΗΤΗΣ ΔΠΘ

ΔΙΑΡΚΕΙΑ ΜΕΤΑΔΙΔΑΚΤΟΡΙΚΗΣ ΕΡΕΥΝΑΣ:
6/4/2017 – 5/4/2019

Η Μεταδιδακτορική Έρευνα υλοποιήθηκε με υποτροφία του ΙΚΥ η οποία χρηματοδοτήθηκε από την Πράξη «Ενίσχυση Μεταδιδακτόρων Ερευνητών/Ερευνητριών» από τους πόρους του ΕΠ «Ανάπτυξη Ανθρώπινου Δυναμικού, Εκπαίδευση και Διά Βίου Μάθηση» με άξονες προτεραιότητας 6, 8, 9 και συγχρηματοδοτείται από το Ευρωπαϊκό Κοινωνικό Ταμείο – ΕΚΤ και το ελληνικό δημόσιο

ACT: «SUPPORT OF POSTDOCTORAL RESEARCHERS» MIS 5001552

REVIEW OF POSTDOCTORAL RESEARCH: «MODELLING, CONTROL AND OPTIMIZATION OF STOCHASTIC PRODUCTION SYSTEMS »

POSTDOCTORAL RESEARCHER:
ALEXANDROS XANTHOPOULOS (OF SOTIRIOS)

SCIENTIFIC FIELD:
PHYSICAL AND ENGINEERING SCIENCES

HOSTING RESEARCH INSTITUTION:
DEMOCRITUS UNIVERSITY OF THRACE
DEPARTMENT OF PRODUCTION & MANAGEMENT ENGINEERING
LABORATORY OF INDUSTRIAL PRODUCTION

SUPERVISOR:
DIMITRIOS KOULOURIOTIS, DUTH PROFESSOR

DURATION OF POSTDOCTORAL RESEARCH:
6/4/2017 – 5/4/2019

This Postdoctoral Research was implemented under scholarship from IKY which was funded by Act “Support of Postdoctoral Researchers” from resources of EP “Development of Human Resources, Education and Life-Long Learning”, with priority axes 6,8,9 and it is co-funded by the European Society Fund-EKT and the Greek state.

1. INTRODUCTION

Real-life production systems are plagued by uncertainty, i.e. at least some degree of randomness is inherent in every system that manufactures goods. Many relevant examples can be cited, e.g.:

- The demand for finished products stems from customers that are positioned outside the boundaries of the production system. Consequently, the demand is exogenous and is beyond the control of the productions system's administration. Demand for finished goods is subjected to many influencing factors that cannot be predicted or controlled and thus, exhibits random fluctuations
- In realistic situations, production processes are rarely executed exactly as planned. This is especially true for production systems that are characterized by a relatively small degree of automation where many processes are foremost manual. Besides the manual labor factor, randomness in production process can also be generated by other unforeseen causes such as quality fluctuations of raw materials, power outages and so forth.
- Finally, an important source of production system randomness is the unreliability of equipment. Power-tools, workstations and other mechanical equipment that is used in production, is subjected to wear and tear and this causes unexpected malfunctions. In such events, production processes are disrupted with potentially long delays in delivery, unforeseen stockouts etc.

Because of the prevalence and importance of randomness in real-life manufacturing, this postdoctoral research focuses on *stochastic production systems*.

1.1 Modelling stochastic production systems

In order to *model* such stochastic production systems, we make use of two important methodologies: *discrete event simulation* and *Markov chains theory*.

Discrete event simulation is a methodological framework for modelling and simulating discrete-event systems. A discrete-event system is a dynamic system with the following characteristics:

- Time-invariant and non-linear
- Discrete-state and event-driven
- Discrete-time or continuous-time

Key point in the simulation of discrete-event systems is the fact that there are countable states (finite or infinite) and that the system state only changes at times where specific *events* occur. In a production system, as the ones studied in this research, such an event could be the arrival of a customer order, the completion of a new product etc.

A Markov chain is special type of a *stochastic process* with countable state space, countable or not countable parametric space where the *Markov property* applies. Informally, the Markov property states that "the next system state depends only on the current state and not on the complete history of state transitions up to the current time". The parametric space of a Markov chain usually denotes time (in this postdoctoral research, the parametric space is always associated with the concept of time). A Markov chain with a countable parametric space is called a *discrete-time* Markov chain. Otherwise, we have a *continuous-time* Markov chain. In this postdoctoral research, we model stochastic manufacturing systems as continuous-time Markov chains.

1.2 Control of stochastic production systems

Up to this point we have stressed the necessity of studying stochastic production systems and we have outlined the techniques that are used in this research to model such systems. Moving further, we highlight the *control* aspect regarding the stochastic production systems that were studied in this research. In broad terms, we examined *heuristic* and *optimal* control policies. Note that, a control policy is a mapping from system states to control actions. A heuristic control policy is a policy that does not comes with guarantees regarding its optimality. An optimal control policy is a policy that is (mathematically) proven to maximize or minimize some objective function that quantifies the performance of the underlying system.

In terms of heuristic policies, in this postdoctoral research, we studied production control policies that belong to the class of *pull-type* control policies. A pull-type control policy coordinates production processes based on actual demand occurrences and not on forecasts or advance-demand-information (ADI) such as MRP (Material Requirements Planning) systems. Notable pull-type production control policies are Kanban, Base Stock, CONWIP, Generic Kanban, CONWIP/Kanban Hybrid etc. Typically, in a pull-type policy, production is coordinated with the use of special signals called *production authorizations*, *kanbans*, *kanban cards* or simply *cards*. In earlier implementations, a kanban card was a physical, tangible card but in modern implementations kanbans are typically replaced by digital signals in the context of some production control software (these control systems are usually referred to as "e-kanban" systems where "e" stands for "electronic").

Standard pull-type control policies have a *fixed* number of cards and the card number has a pivotal role in their overall performance as it largely determines the throughput, the average number of backorders and the average inventory levels. Nonetheless, manufacturing systems often operate in turbulent environments, e.g. there is great uncertainty regarding the demand arrival process and/or the service times are subject to random fluctuations. Therefore, it is preferable to *dynamically adjust* the number of kanban cards in response to the current state of the system. As a result of the need to develop pull-type control policies that adapt to their environment, numerous heuristic approaches have been proposed in the literature over the years. These adaptive heuristics differ in many aspects, including the state representation of the system which is used to guide the control decisions, the mechanism for adding or retrieving cards from the manufacturing facility etc. Notable adaptive pull control polices are Extended Kanban, Generalized Kanban, Adaptive Kanban, among others.

For the purposes of this research, we have also studied heuristic periodic inspection and preventive maintenance policies. By means of an inspection we determine the current deterioration level of a production system. A periodic inspection policy establishes the time intervals between two successive inspections of the system. Preventive maintenance restores the system to a "better-than-before" state and thus, it decreases the frequency with which the manufacturing system is subjected to random breakdowns. A preventive maintenance policy determines the system states that signal the authorization of a maintenance epoch.

1.3 Optimization of stochastic production systems

By modelling the stochastic systems that are examined in this research as Markov chains or discrete-event models we can study their behavior and gain insight on their properties. Nonetheless, our ultimate goal is to *optimize* the manufacturing systems in question, i.e. derive optimal control policies or, informally, determine what is the best control action in any system state. To do so, we first need to define objective functions that quantify the performance of a such system. Performance metrics that are typically of interest in manufacturing are system throughput, average finished and work-in-process inventories, system availability and utilization, average idle and down time etc. In order to derive optimal production control policies we use Dynamic Programming and Reinforcement Learning.

Dynamic Programming is the classical approach in optimal control. It can be applied to situations where the underlying optimization problem can be formulated as a Markov Decision Process (MDP). Prominent Dynamic Programming algorithms for solving MDPs are policy iteration and value iteration. The Dynamic Programming approach has the significant advantage that it is guaranteed to converge to the optimal solution. Nonetheless, it also poses to significant drawbacks. In order for Dynamic Programming methods to apply, the complete model of the underlying system is needed, i.e. the transition probabilities for all states must be available. Furthermore, Dynamic Programming approaches are plagued by the so called *dimensionality curse*, i.e. they cannot be applied to large-scale optimization problems due to the resulting "computational explosion".

Reinforcement Learning is a methodology that belongs to the field of Machine Learning and it can be considered as some kind of approximate, stochastic, dynamic programming. Unlike Dynamic Programming though, Reinforcement does not require a complete model of the system in question and it can be used to solve optimization problems of any size. These are two important features that constitute the use of Reinforcement Learning very appealing.

In order to derive optimal production control policies, Reinforcement Learning (RL) is employed in this research. According to the RL paradigm, a decision-making agent is placed within an environment whose dynamics are initially unknown. The agent interacts at certain time points (decision epochs) with its environment. At a decision epoch, the agent receives a representation of the environment's current state and selects some action from a set of admissible controls. At the next decision epoch, the agent observes the result of its previous action selection. This cycle is repeated and after a sufficient number of decision epochs, the agent identifies through the process of trial-and-error the optimal control policy in respect to some performance metric.

In this postdoctoral research, the agent environment is some production system. The system's dynamic behavior is obtained by means of simulation. In order to obtain the optimal production control policies, the decision-making/learning agent is interfaced with the production system simulation model. A wide range of alternative learning algorithms (e.g. Schwartz's R-learning, R-smart etc.) and exploration strategies (e-greedy, adaptive pursuit etc.) have been proposed in the literature over the years.

2. RESEARCH QUESTIONS AND SCIENTIFIC CONTRIBUTION OF POSTDOCTORAL RESEARCH

The research questions that are addressed by this postdoctoral research belong to the field of Industrial Engineering. In broad terms we study problems that pertain to production control in:

- Single-stage and multi-stage systems
- Single-product and multi-product systems
- Standard and adaptive systems

We also examine inspection and maintenance control policies and more specifically:

- Periodic inspection policies
- Threshold-type preventive maintenance policies

An important and innovative feature of this postdoctoral research is that it examines *joint or integrated* production and maintenance problems, i.e. control problems where there is an interaction between production control and inspection/maintenance decisions. Another significant theoretical contribution of this research is that it emphasizes adaptive control policies, a scientific field that is significantly under-represented in the existing bibliography. Finally, a salient feature of this research is that it uses state-of-the art, i.e. Machine Learning-based solution approaches that have not been applied before in the literature to solve hard optimization problems which have not been addressed up to now.

The relevant research questions, the novelty and the contribution to science of this postdoctoral research are elaborated below.

2.1 Research direction 1: Pull-type production control policies for multi-product manufacturing systems

According to a common definition of pull-type production control, a pull system is one in which production operations are coordinated based on actual demand occurrences and not on advance demand information or forecasts. An excellent review of pull control methods and critical comparisons with alternative production control paradigms is given in Liberopoulos (2013).

Numerous pull control strategies (or policies) have been proposed in the relevant literature and a considerable number of papers have been devoted to the modeling, evaluation and comparison of alternative pull systems. The reader is referred to Koulouriotis et al. (2010) and Xanthopoulos and Koulouriotis (2014) for some indicative examples.

Nonetheless, pull production control policies have been mostly studied in the context of single product type systems up to now. This is rather surprising because pull-type production control was initially proposed as a means for coordinating complex production processes. In recent years, this problem tends to be alleviated with the emergence of a new research direction that examines multi-product systems (Onyeocha et al. 2015; Renna, 2018).

This postdoctoral research addresses the following research questions:

- How can we expand known pull-type control policies in the literature such as CONWIP, Base Stock etc., to account for multi-product systems? More specifically, how can we develop *queueing network models* for the multi-stage, multi-product CONWIP, Base Stock and CONWIP/Kanban systems?
- What is the ranking of the alternative production control mechanisms for multi-product systems under the metrics of average number of backorders, average finished product inventories and average waiting time of backordered demand? How are they compared in a series of simulation experiments?
- What insights are gained on the behavior of the different pull production control methods and what are the related managerial implications?

2.2 Research direction 2: Adaptive pull-type control policies for single-stage manufacturing systems

As a result of the need to develop pull-type control approaches that adapt to their environment, numerous *heuristic* approaches have been proposed in the literature over the years. These adaptive heuristics differ in many aspects, including the state representation of the system which is used to guide the control decisions, the mechanism for adding or retrieving cards from the manufacturing facility etc. There are several comparative evaluations of adaptive pull-type policies that pertain to specific types of manufacturing systems and performance metrics. However, no conclusive results have been published which indicate that some adaptive heuristic, in general, outperforms other adaptive approaches. More importantly, none of the published adaptive pull-type policies has been proven to be *optimal* or deviate from the optimal up to some specific extent.

In this research we study adaptive pull-type production control in the context of a single-stage manufacturing system. More specifically, we address the following questions:

- How can we formulate the derivation of optimal adaptive pull-type production control policies as a Markov Decision Process?
- What are the optimal policies which are obtained by means of a Dynamic Programming approach?
- What are the properties and what is the structure of the optimal policy? Can this analysis lead to conclusions regarding existing, adaptive heuristics?
- How are existing, adaptive control policies compared to the optimal one, in terms of their observed performance? Are there any situations/settings where heuristic control policies approximate the optimal one adequately?

2.3 Research direction 3: Adaptive pull-type control policies for multi-stage manufacturing systems

This research direction is a straightforward extension of the former. Namely, we are interested in obtaining optimal adaptive control policies, comparing them to existing heuristics and drawing conclusions regarding their performance. They key differentiation here is that we study multi-stage and not single-stage manufacturing systems. This seemingly minor diversification dramatically perplexes the whole solution approach.

Recall from section 1 the "dimensionality curse" of Dynamic Programming, i.e. its inherent inability to deal with large-scale optimization problems. Turning a single-stage system to a multi-stage, rapidly increases the number of possible system states. This causes the Dynamic Programming approach to be no longer applicable. This research direction poses the following questions:

- How can we derive optimal adaptive pull-type policies for multi-stage manufacturing systems?
- What are the implementation details of a solution approach that is based on Reinforcement Learning?
- How does the optimal control policy compare to heuristic ones in an extended series of simulation experiments?

2.4 Research direction 4: Integrated production, inspection and maintenance control in stochastic manufacturing systems

Many real-world manufacturing systems have some defining characteristics which are cited hereafter. In high volume production and due to the size of the equipment, the transition of the manufacturing system from idle to working state requires substantial preparatory activities such as establishing the supply of raw materials, setting the cooling systems online etc. Consequently, each new production batch entails significant costs. In standardized production, finished goods and easy to store whereas the pressure to meet customer demand is high. As a result, holding costs are low in relation to lost sales costs. The adverse consequences of equipment failures are considerable because of the time/costs needed for conducting repairs and the related productivity drop. Preventive maintenance activities can be undertaken to prevent hard failures but they do not come without the associated costs. Because of the complexity that characterizes many manufacturing systems, inspections are needed so as to determine the current deterioration state accurately. Inspections and maintenance actions are preferable to be scheduled when the system is idling in order to increase productivity.

Motivated from the above we examine joint production, inspection and maintenance control problems. More specifically, we try to answer research questions such as:

- how do changes in, e.g. production rate and holding cost factors, affect the system's performance?
- which is the best inspection/maintenance/production control policy for a given manufacturing system?
- in what situations do the costs of inspection/preventive maintenance action counterbalance the related benefits?
- why does some control policy outperforms another policy in certain production environments?
- how should the batch size be set in respect to production cost factors?
- which cost components are pivotal in minimizing total cost?

3. STRUCTURE OF REVIEW

In this section we elaborate on the structure of the remainder of this review.

- In chapter 4, we examine pull-type production control policies in multi-product manufacturing environments (refer to research direction 1, section 2.1)
- In chapter 5, we derive optimal adaptive control policies for single-stage systems using Dynamic Programming (refer to research direction 2, section 2.2)
- In chapter 6, we calculate optimal pull-type control policies for multi-stage manufacturing systems by means of a Reinforcement Learning-based approach (refer to research direction 3, section 2.3)
- In chapter 7, we develop the Markovian model of a stochastic manufacturing system that is subjected to deterioration with usage. We study several heuristic production, inspection and maintenance control policies. We define metrics that quantify the system's performance and find optimal or near-optimal control parameters for the system. The behavior of the system in respect to alternative configurations is studied extensively (refer to research direction 4, section 4.1)
- Chapter 8 contains the concluding remarks of this postdoctoral research and lays down several plausible extensions of it.

4. PULL-TYPE PRODUCTION CONTROL OF MULTI-PRODUCT MANUFACTURING SYSTEMS

In this chapter, we examine the CONWIP, Base Stock, and CONWIP/Kanban Hybrid pull strategies (Liberopoulos, 2013; Koulouriotis et al., 2010; Xanthopoulos & Koulouriotis, 2014) in multi-product manufacturing systems. In such a manufacturing system, several product types are manufactured by utilizing the same resources, i.e. machines, conveyors, workstations etc. A setup is required when switching from one product type to another, however, more than one types can be interspersed in the manufacturing system at the same time.

Pull production control policies have been mostly studied in the context of single-product systems up to now. In recent years, a new research direction has emerged that examines mixed-model or multi-product systems (Renna, 2018; Onyeocha et al., 2015; Onyeocha, Khoury and Geraghty, 2015a-b). We advance the research in this field by developing queueing network models of multi-stage, multi-product manufacturing systems operating under the three aforementioned pull control strategies. Discrete event simulation models of the alternative production systems are implemented in the simulation software JaamSim (King & Harrison, 2013). A comparative evaluation of CONWIP, Base Stock and CONWIP/Kanban Hybrid in multi-product manufacturing is carried out in a series of simulation experiments with varying demand arrival rates, setup times and control parameters. The control strategies are compared based on the metrics of average wait time of backordered demand, average finished products inventories, and average length of backorders queues.

4.1 System description – Production control policies

The system under investigation is comprised of several stages in tandem and manufactures a number of product types. In the remainder of this section, *i* and *j* will be used to denote an arbitrary production stage and product type, respectively.

Raw materials enter the system and are processed in all production stages starting from the first one and moving to the downstream stage. Finished products of type *j* are outputted by the last stage and stored in the respective finished goods inventory. Raw materials are assumed to be continuously available, i.e. the raw materials buffers are never empty. Demands for finished products arrive dynamically to the system and the times between successive demand arrivals are stochastic. Upon a demand arrival, one unit of type *j* product is requested instantly. If there are available finished products of type *j*, then the demand is satisfied immediately. If not, then the demand enters the, type *j*, backorders queue and waits until inventory is made available.

All production stages have a manufacturing facility that is composed of a single machine (with stochastic service times) and the associated input queue. A machine can process all product types; it process products one-by-one and undergoes a setup when switching from one type to another. A type *j* product that completes its processing in the *i*-the stage is stored in output buffer *i*,*j*.

The flow of materials from one production stage to the next is coordinated by a pull-type production control policy. In broad terms, a production control policy determines when stage *i* should pull a type *j* part from the upstream output buffer in order to process it. In this chapter, we examine the CONWIP, Base Stock and CONWIP/Kanban Hybrid policies for mixed-model manufacturing. We develop the queueing network models of the respective systems in the following three sections.

4.2 Multi-product CONWIP, Base Stock and CONWIP/Kanban Hybrid system

Figures 4.1, 4.2 and 4.3 show a two-stage, two-product type CONWIP, Base Stock and CONWIP/Kanban Hybrid system, respectively. Note that the properties of the CONWIP/Base Stock/Hynrid system presented here hold for any number of product types and production stages. All queues shown in Figures 4.1, 4.2 and 4.3 operate according to the First-Come-First-Served rule.

In Figure 4.1, M_i is the i-th manufacturing facility and $P_{0,j}$ is the raw materials buffer for product type j. The output buffer of stage i and product j is denoted as queue $P_{i,j}$ and queue $D_{3,j}$ contains demands for type j finished products. Finally, queue $D_{1,j}$ contains demands for stage – 1 products of type j.

Initially, i.e. at time 0, all machines are idling and all queues are empty except $P_{0,j}$ (by definition) and $P_{i,j}$, for all i, j. At time 0, queue $P_{i,j}$ contains $S_{i,j}$ parts, where $S_{i,j}$ is the base stock (initial inventory) of stage-i and part-j products. The integers $S_{i,j}$, $\forall i, j$, are the control parameters that characterize a multi-product CONWIP system. The sum of the $S_{i,j}$ parameters equals the constant number of parts that "circulate" in the manufacturing system.

The control logic of the CONWIP policy is the following. All stages except the first one are constantly authorized to produce. Consequently, it can be argued that production stages 2, 3, … operate according to a push strategy. The first stage receives an authorization to process a new type-j part at the moment when a type-j finished product exits output buffer $D_{3,j}$ (transmission of information is assumed to be instantaneous).

In Figure 4.2, M_i denotes the i-th manufacturing facility and $P_{0,j}$ symbolizes the raw materials inventory for product type j. Queue $P_{i,j}$ contains stage-i completed parts of type j. $D_{i,j}$ contains demands for type-j parts; e.g. an element of queue $D_{3,j}$ is a demand for a finished product of type j and an element of queue $D_{2,j}$ authorizes the production of a new stage-2 part of type j.

At time 0, all machines are idle and all queues are empty with the exception of $P_{0,j}$ and $P_{i,j}$, $\forall i, j$. It is reiterated that an infinite supply of raw materials is assumed. Initially, queue $P_{i,j}$ contains $S_{i,j}$ parts. Similarly to CONWIP, the base stocks $S_{i,j}$, for all i, j, are the only control parameters of a multi-product Base Stock system. However, in a Base Stock system, there are no limits on the Work-In-Process and finished goods inventory levels.

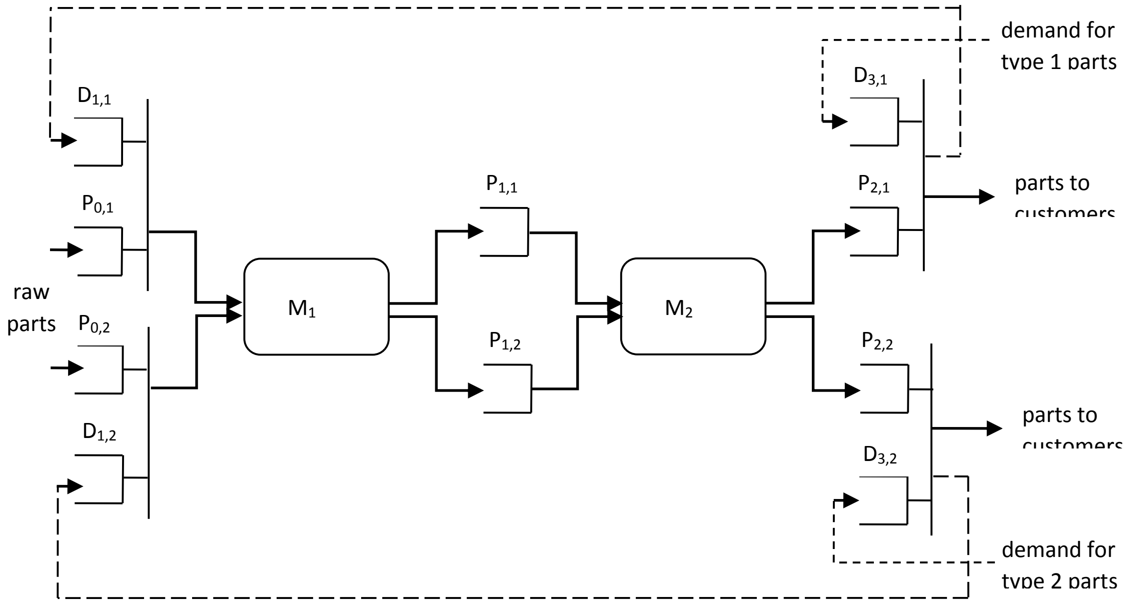

Fig. 4.1. A CONWIP system with two stages and two product types.

The Base Stock system operates as follows. At the time when a demand for a type-j finished product arrives to the system, an analogous demand is transmitted to all queues $D_{i,j}$ authorizing the production of a new type-j, stage-i part, for all i. This way, the production of a new part can commence even if no finished goods inventory has been consumed, allowing for increased flexibility in following demand fluctuations.

M_i symbolizes the manufacturing facility i and $P_{0,j}$ is the raw parts buffer for type j, in Figure 4.3. Queue $PA_{i,j}$ contains stage-i, type-j completed parts with kanbans (production authorizations) attached on them. Queue $P_{2,j}$ has finished products of type j and queue $D_{3,j}$ contains demands for such products. CONWIP-type demands are held in queues D_1 and D_2. Finally, queues $DA_{1,j}$ contain kanban/demand pairs for stage-1 parts of type j.

Initially, all machines are idling and all queues are empty except for the raw parts buffers, which are always non-empty by definition, and queues $PA_{1,j}$ and $P_{2,j}$, for all j. The latter contain $K_{1,j}$ and $S_{2,j}$ parts, respectively. The number of stage-1, type-j kanbans $K_{1,j}$ and the base stocks $S_{2,j}$ are the control parameters of the system shown in Figure 4.3.

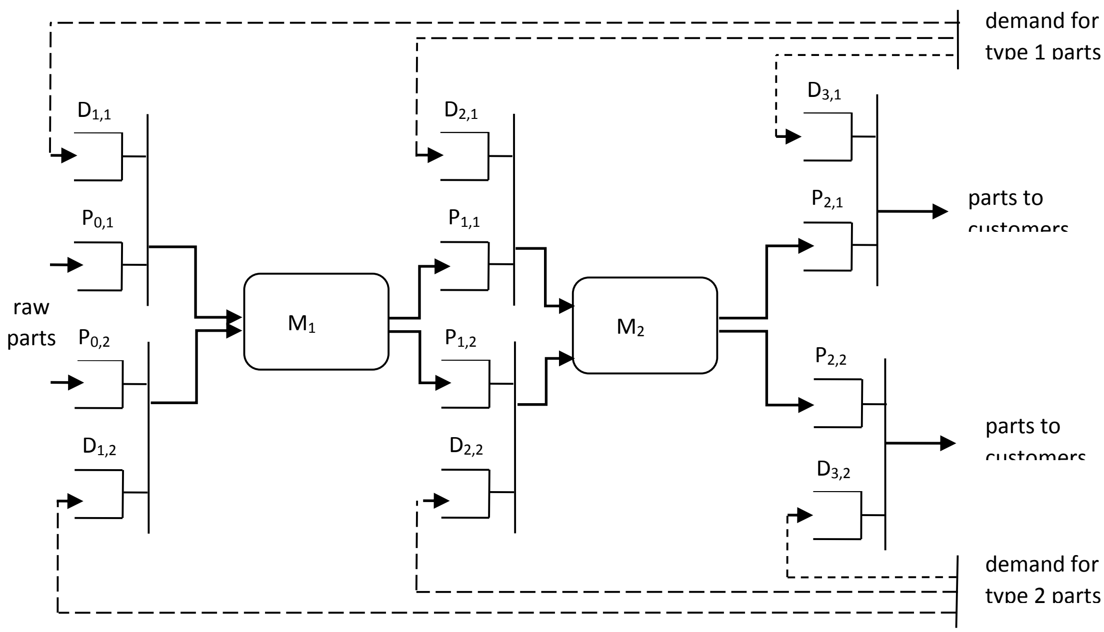

Fig. 4.2. A Base Stock system with two stages and two product types.

The CONWIP/Kanban Hybrid system operates as follows. The last stage has a perpetual authorization to produce, similarly to the pure CONWIP policy. At the time when a type-*j* finished product is delivered to a customer, a relevant demand is sent to queue D_j at the beginning of manufacturing line. The first stage is authorized to produce a new type-*j* part if there is at least one element in each of the D_j and $DA_{1,j}$ queues. All other stages operate under a Kanban control policy (for additional details refer to Koulouriotis, Xanthopoulos and Tourasis, 2010). The rationale behind the philosophy of the CONWIP/Kanban Hybrid policy is to combine the swift turnaround of the CONWIP system with the tight coordination between production stages offered by Kanban.

4.3. Experimental results

The investigated control policies were compared in a series of simulation experiments that pertained to a manufacturing system with five stages and two product types. We defined a base simulation case as the starting point of our analysis and then varied i) the average time between arrivals, ii) the setup time for switching from one product type to another and iii) the policies' control parameters, in order to study the behavior of the alternative control mechanisms.

The base simulation case is defined as follows: times between arrivals are exponentially distributed with mean 1.26 time units. Upon a demand arrival, a type 1 (or type 2) finished product is requested with probability 0.5. The service times of all machines are exponential with mean 0.8 and 1.2 for type 1 and type 2 products, respectively. When a machine switches from one part type to another, a setup with duration 0.25 time units is incurred.

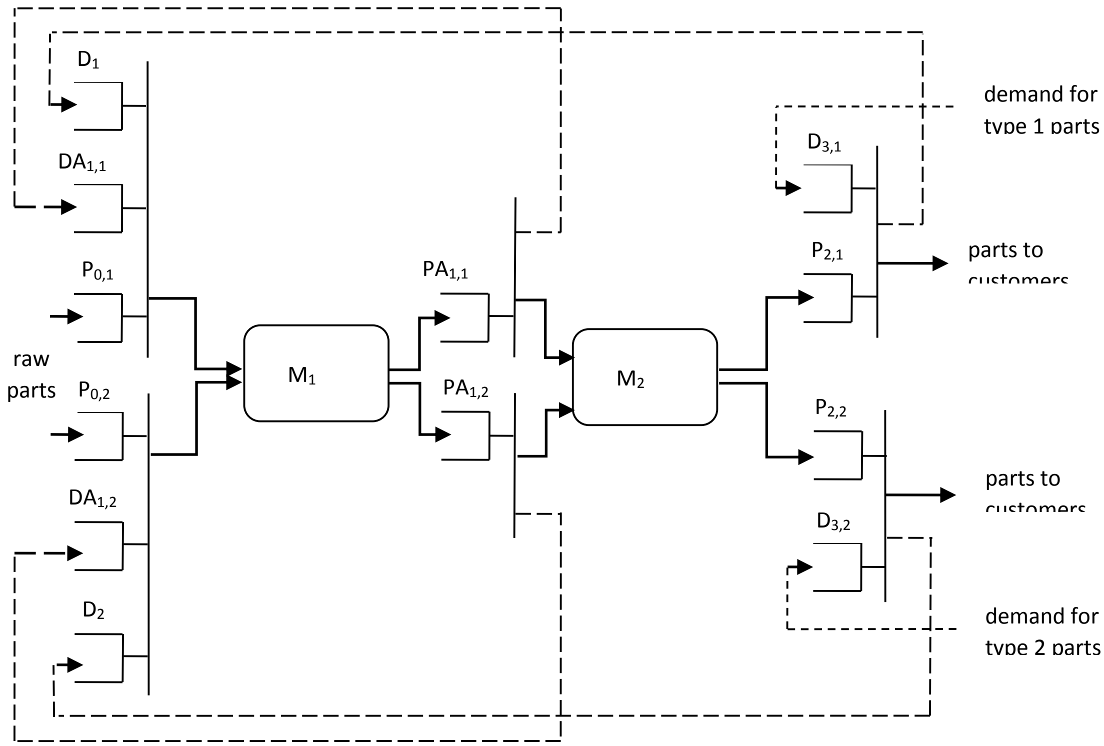

Fig. 4.3. A CONWIP/Kanban Hybrid system with two stages and two product types.

The control parameters (i.e. base stocks and/or kanbans) for all policies, stages and products are set to the value of 5. For all simulation models, the length of a replication is set to 10000 time units and the number of independent replications for all models is 20. All simulation models are built using the JaamSim software (King and Harrison, 2013).

Figure 4.4 shows the average waiting time of backordered demands of each control policy for average time between arrivals that varies in the range [1.2, 1.38]. It is observed that, average waiting time of demand is an increasing function of the arrival rate. For relatively low arrival rates, it is observed that the performance of the alternative control policies is practically the same. However, for relatively high arrival rates the Base Stock and the CONWIP/Kanban Hybrid policies are clearly the best and worst performing mechanisms, respectively.

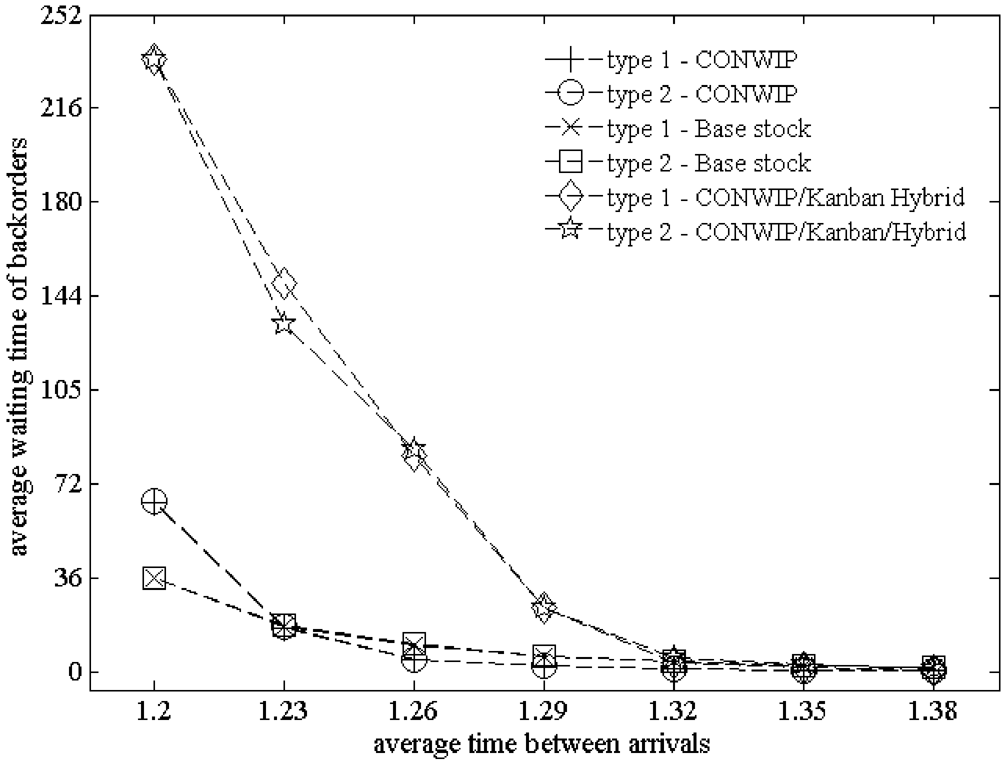

Fig. 4.4. Average waiting time of backordered demands for alternative production control policies and varying arrival rates.

This can be attributed to the following qualitative characteristics of these two control policies: the CONWIP/Kanban Hybrid system has a very tight coordination between the various production stages whereas the Base Stock does not coordinate at all production operations at different stages. Consequently, the Base Stock system responds rapidly to demand fluctuations, compared to CONWIP/Kanban Hybrid.

Figure 4.5 shows the average number of backorders of each control policy for varying setup times. It is seen that the average number of backorders is an increasing function of the setup time. Again, for relatively small setup times, the differences between the alternative control policies are rather negligible. For relatively large setup times the ranking of the policies is Base Stock, CONWIP, CONWIP/Kanban Hybrid. Increasing the setup time has a similar effect to the system as increasing the arrival rate or decreasing the service rate, i.e. the workload imposed on the manufacturing system increases. This explains the observed performance of the various control mechanisms. Overall, we can argue that CONWIP/Kanban Hybrid is significantly affected by the magnitude of the workload that is imposed on the system whereas the Base Stock system is relatively insensitive to changes in the workload. The Base Stock policy appears to be a good choice when the manufacturing system operates close to its capacity.

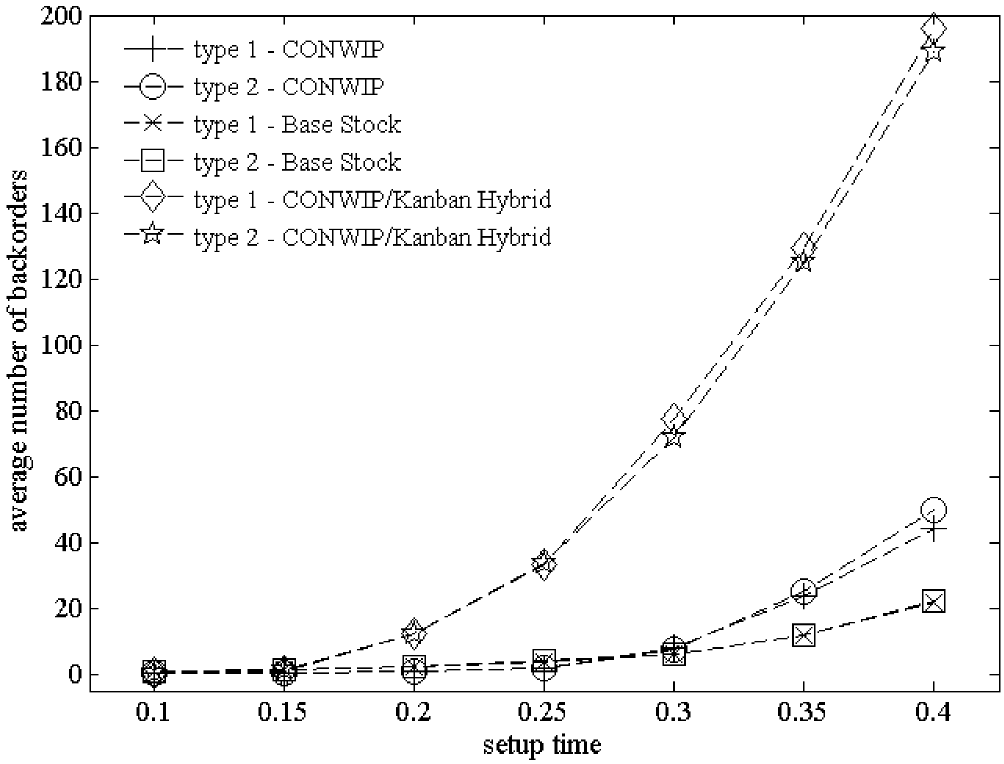

Fig. 4.5. Average number of backorders for alternative production control policies and varying setup times.

In order to examine the effect of the control parameters to the performance of the investigated control mechanisms we used a fractional factorial design. The factors (parameters) of the experimental design are the base stocks or number kanbans for each production stage and product type. The low and high level for all factors was set to 2 and 5 respectively. A 2^{4-1} fractional factorial design of resolution IV was generated (Xanthopoulos and Koulouriotis, 2014) and it is presented in Table 4.1.

Figure 4.6 shows the average finished product inventories of all investigated manufacturing systems, for alternative control parameter sets. It is observed that the CONWIP policy is by far the most sensitive control strategy in respect to changes of parameters $S_{i,j}$.

Exactly the opposite holds for CONWIP/Kanban Hybrid, whereas the Base Stock policy defines a somewhat intermediate situation. From Figure 4.6, it is seen that the average finished goods inventories in a Base Stock system depend primarily on the base stocks of the last stage. However, in a CONWIP system the average inventories of finished products depend mostly to the sum of base stocks in all production stages. This clearly can be attributed to the fact that in a CONWIP system, Work-In-Process constantly flows without interruption to the last stage.

parameter set	$S_{1,j}$ ($K_{1,j}$)	$S_{2,j}$ ($K_{2,j}$)	$S_{3,j}$ ($K_{3,j}$)	$S_{4,j}$ ($K_{4,j}$)	$S_{5,j}$
No 1	2	2	2	5	5
No 2	5	2	2	2	2
No 3	2	5	2	2	5
No 4	5	5	2	5	2
No 5	2	2	5	5	2
No 6	5	2	5	2	5
No 7	2	5	5	2	2
No 8	5	5	5	5	5

Table 4.1. The investigated sets of control parameters, where j = 1, 2. Parameters $K_{i,j}$ apply only to the CONWIP/Kanban Hybrid system.

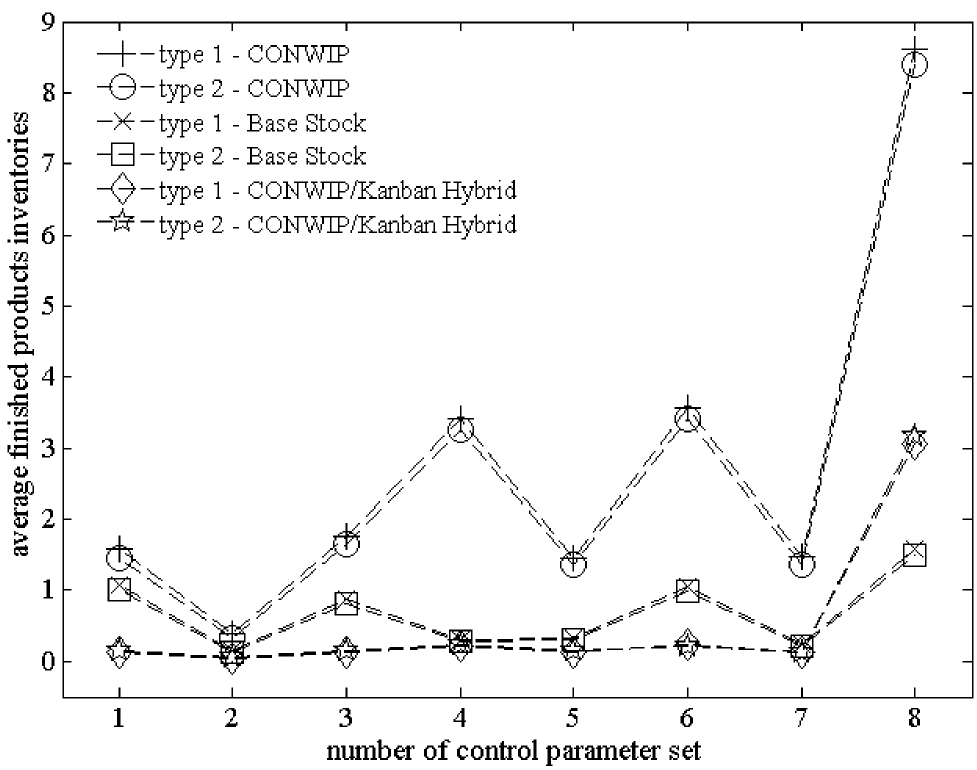

Fig. 4.6. Average inventories of finished products for alternative production control policies and varying arrival rates.

5. OPTIMAL ADAPTIVE PULL-TYPE PRODUCTION CONTROL FOR SINGLE-STAGE MANUFACTURING SYSTEMS

In this chapter, we derive optimal adaptive pull production control policies for single-stage systems. In order to achieve this goal, we formulate the respective problem as a Markov Decision Process and apply a standard Dynamic Programming algorithm to solve it, that is, value iteration. We study the properties of the optimal policy through a set of numerical experiments. Moving further, we compare existing, heuristic policies to the optimal one. Based on this analysis, we draw conclusions regarding the expected performance of heuristic pull-type policies in different system settings. In the next section 5.1, we offer a brief overview of relevant works.

5.1 Theoretical background

Adaptive pull-type approaches that have been proposed in the literature often differ in numerous aspects and are not directly comparable to others. Some relevant examples are cited hereafter with one of the earliest works being the STC policy of Hopp & Roof (1998) where both the number of cards and the capacity of the manufacturing system was adjusted.

The approaches of Korugan & Gupta (2014) and Takahashi et al. (2014) pertained to remanufacturing systems. The approaches of Takahashi (2003) and Takahashi & Nakamura (2002) focus on statistical methods for detecting changes in the manufacturing system's environment. The method of Gupta et al. (1999) makes a strong assumption that future demand is known beforehand. The approach of Liu & Huang (2009) is only applicable to pure flowshops where each stage is treated as an M/G/1 queue. Belisario & Pierreval (2015) use genetic programming to obtain adaptive control policies but their approach is only tested in two simulation cases and so, a complete comparative evaluation of their results with alternative methods is not possible. Renna (2015) applies fuzzy control to dynamically adapt the number of cards but the proposed method is characterized by numerous parameters and, in the absence of guidelines for setting them, numerical result reproduction is very difficult.

The aforementioned adaptive approaches are not considered in this study because of their singularities that prohibit their extensive comparison with other approaches. In Xanthopoulos et al. (2017) several adaptive Kanban-type systems are shortlisted which are heuristics that make very few assumptions regarding the characteristics of the controlled manufacturing system. Consequently, they can be easily applied, fine-tuned, and tested to manufacturing systems of arbitrary size and structure, such as flowshops, parallel machine systems etc. This is a very favorable characteristic because it facilitates the complete comparative evaluation of alternative control schemes.

From the set of heuristics examined in Xanthopoulos et al. (2017), three approaches are also tested in this research: the Extended Kanban policy (Dallery & Liberopoulos, 2000), the Adaptive Kanban policy (Tardif & Maaseidvaag, 2001) and the approach of Framinan et al. (2006) which will be referred to as FSL hereafter, for short. The remaining adaptive policies that were studied in Xanthopoulos et al. (2017), i.e. Generalized Kanban, AEK, RCK, RC, FTH, are excluded from this paper on the following grounds. In the single-stage system which is examined in this research, the Extended Kanban policy operates exactly as the Generalized Kanban (Buzacott & Shanthikumar, 1993) and AEK. The AEK and the RCK (Renna et al., 2013) policies are flowshop-oriented control schemes. The FTH and RC policies were not found to yield encouraging results in numerous cases (Xanthopoulos et al. (2017)).

5.2 Model development for the single-stage manufacturing system

The investigated system consists of a *single manufacturing facility*. The manufacturing facility is comprised of N identical, independent and parallel machines. Machines can process parts one by one, i.e. there is no batch production. No preemption is allowed, meaning that the processing of a part cannot be interrupted before its completion. The machines' service times are exponentially distributed with mean $1/\mu$. A single type of end-items is manufactured. At the time when a machine completes the processing of a part, the finished product is stored in the *finished goods buffer*.

The demand for finished goods is exogenous, i.e. generated by customers, and thus, beyond the control of the production system manager. Demand for finished goods arrives randomly at the system and the times between successive demand arrivals are exponentially distributed. A *seasonal* demand pattern is assumed with *M demand levels*, where the mean inter-arrival time at level i is $1/\lambda_i$. The system transits to level $i + 1$ and 1 with rate θ when the current demand level is $i < M$ and $i = M$, respectively. The demand level transition times are considered to be exponential. An example of this cyclical demand pattern is shown in figure 5.1. A single demand level is also examined in our model, as a special case.

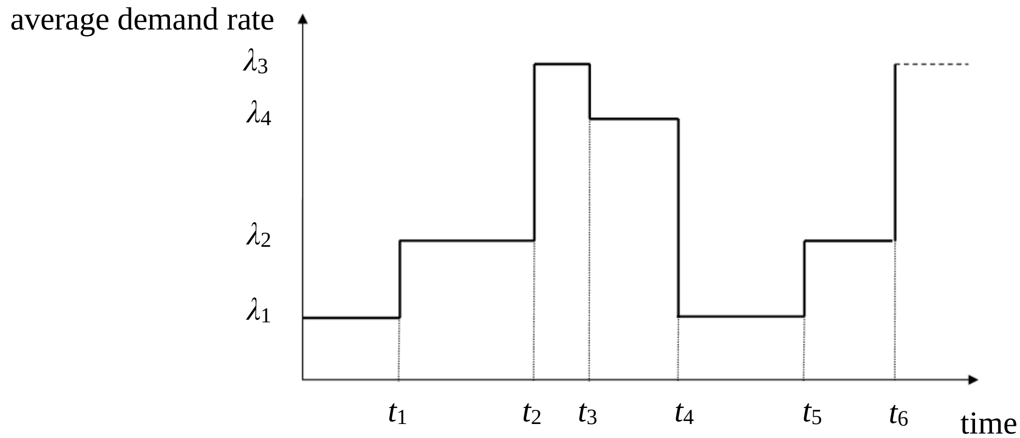

Fig 5.1. An example of a cyclical demand pattern with four levels and $\lambda_1 < \lambda_2 < \lambda_4 < \lambda_3$. The average length of the intervals $(t_2 - t_1)$, $(t_3 - t_2)$, …. is $1/\theta$.

At the time of a demand arrival, one end-item is requested from the finished goods buffer. If there is available inventory at that time, the demand is satisfied instantaneously. If there is no inventory available, the demand is placed in the *backorders queue*. The discipline of the backorders queue is First-Come-First-Served. At the time when inventory is made available, the first pending demand in the backorders queue is satisfied instantaneously. The maximum allowed length of the backorders queue is B_{max}. If a demand arrives while the backorders queue length is B_{max}, then the newly arrived demand is lost to the system.

An infinite supply of raw materials is assumed, i.e. raw materials are perpetually available when needed. The system operates under an adaptive Kanban control policy. Initially, all machines are idling and all system queues are empty with the exception of the finished goods buffer which has a number of end-items. Each end-item has a *kanban card* attached to it. As soon as an end-item exits the finished goods buffer, the card which was attached to it is sent to the manufacturing facility in order to authorize the production of a new part. If there is at least one machine available at that time, the card is attached to a raw part and the raw part is loaded to an available machine for processing. If all machines are busy at that time, the raw part with the kanban card attached to it is placed in the *input queue* of the manufacturing facility. The raw part which has been authorized to receive processing remains in the input queue until a machine becomes available. Production authorizations cannot be cancelled, i.e. raw parts that are placed in the input queue cannot be scrapped or returned to the raw parts inventory.

Symbol	Description
N	number of machines
M	number of demand levels
λ_i	average demand arrival rate at level *i*
θ	rate of switching from one demand level to the next
μ	service rate
K_{max}, K_{min}	maximum and minimum number of kanban cards
B_{max}	maximum length of backorders queue
h_1	cost per time unit per Work-In-Process part
h_2	holding cost per time unit per stocked end-item
b	cost per time unit per backordered demand
WIP	average Work-In-Process level
I	average finished goods inventory
B	average length of backorders queue
J	average cost

Table 5.1. Definition of symbols pertaining to system description.

The number of cards that "circulate" within the system can vary in the range $[K_{min}, K_{max}]$ so as to adjust the system's capacity in response to random fluctuations of the demand and the production process. Recall that in a Kanban system there must be at least one kanban card or the manufacturing facility will never be given the authorization to produce, and thus $K_{min} = 1$.

Holding and backorder costs are considered in our analysis. A *Work-In-Process part* incurs a cost of h_1 monetary units per time unit. Work-In-Process parts are parts which are held in the input queue or being processed by the manufacturing facility. The cost of storing an end-item in the finished goods buffer per time unit is h_2. Finally, the cost of one backordered demand per time unit is b. Our goal is to derive the optimal adaptive Kanban policy in respect to the minimization of the following cost function:

$$J = h_1 WIP + h_2 I + bB \tag{5.1}$$

where WIP is the average Work-In-Process level, $I$is the average finished goods inventory, and B is the average length of the backorders queue.

5.3 Derivation and analysis of the optimal adaptive control policy

In this section, we outline the Markov Decision Process formulation of the underlying optimization problem. The objective is to find adaptive Kanban-type policies that minimize the long-run average cost defined in (5.1). The evolution of the examined production/inventory system is driven by three types of events: i) demand arrival, ii) production completion of an end-item, and iii) demand level change. The system's state changes at the occurrence of these events.

The state of the system is described by the integer variables x, y, and z where x is the number of Work-In-Process items, y is the *inventory position* of the system, i.e. y = (finished goods buffer level) – (length of backorders queue), and z is the current demand level. Recall from section 5.2 that there can be no more than B_{max} pending demands in the backorders queue. Furthermore, note that in a Kanban system the Work-In-Process level plus the finished goods inventory must equal the number of circulating kanban cards. Consequently, the state space of the system can be written as:

$$S = \{(x, y, z) \mid 0 \le x \le K_{\max}, -B_{\max} \le y \le K_{\max}, 1 \le z \le M, K_{\min} \le x + \max\{0, y\} \le K_{\max}\}.$$

In the overwhelming majority of heuristic adaptive Kanban policies which have been proposed in the relevant literature, a single kanban card is added to or removed from the production system in each decision epoch. We adopt this convention in our analysis and so the admissible action set is formulated as: A = {"do nothing", "release card", "capture card"}. When decision "do nothing" is made, the system operates exactly as a standard Kanban system. The decisions "release card" and "capture card" result in the number of circulating cards to be increased and decreased by one, respectively.

Following the standard stochastic Dynamic Programming approach (see e.g. Bertsekas, 1995 and Puterman, 1994) the problem can be formulated as Markov Decision Process. We define as $V_k(x, y, z)$ the optimal average cost over the first k (decision epochs) when the initial state is (x, y, z). Next, we derive the *Bellman equations* for the optimal total average cost over the first $k + 1$ events, where $V_0(x, y, z) = 0$ for every state. The optimal long-run average cost J^* is given by:

$$J^* = v \lim_{k \to \infty} \left(V_k(x, y, z) - V_{k-1}(x, y, z)\right), \quad (x, y, z) \in S \tag{5.2}$$

and can be approximated numerically by the *value iteration algorithm*:

1: Set $V_0(x, y, z) = 0$ for all x, y, z and $k = 0$. Initialize error parameter e. Go to 2.

2: Estimate all $V_{k+1}(x, y, z)$ from Bellman equations. Go to 3.

3: Set $\Delta_{\max} = \max_{x,y,z}\{V_{k+1}(x, y, z) - V_k(x, y, z)\}$ and $\Delta_{\min} = \min_{x,y,z}\{V_{k+1}(x, y, z) - V_k(x, y, z)\}$. Go to 4.

4: IF $\Delta_{\max}-\Delta_{\min} < e$

set $J^{*}= v\ \Delta_{\max}$ and terminate.

ELSE

set $k \leftarrow k + 1$ and go to 2.

In the remainder of this section we investigate numerically the properties of the optimal adaptive Kanban-type policy for the examined manufacturing system. For the purposes of this investigation we use a case that pertains to a system with the following parameters: $N = 10$, $M = 2$, $\lambda_1 = 1/6$, $\lambda_2 = 1/4$, $\mu = 1/15$, $\theta = 1/100$, $h_1 = 1$, $h_2 = 2$, $b = 100$. This example case has two demand levels, nevertheless, the observations made regarding the structure of the optimal policy can be straightforwardly extended for any value of M. We solve the related Dynamic Programming model for $K_{max} = 50$, $K_{min} = 1$, $B_{max} = 50$. The optimal policy in terms of the next event being "demand arrival" and "production completion" is depicted graphically in figure 5.2 and figure 5.3, respectively. Recall that the "demand level change" event does not alter the Work-In-Process level or the inventory position, and therefore, no decision is made at the occurrence of that event.

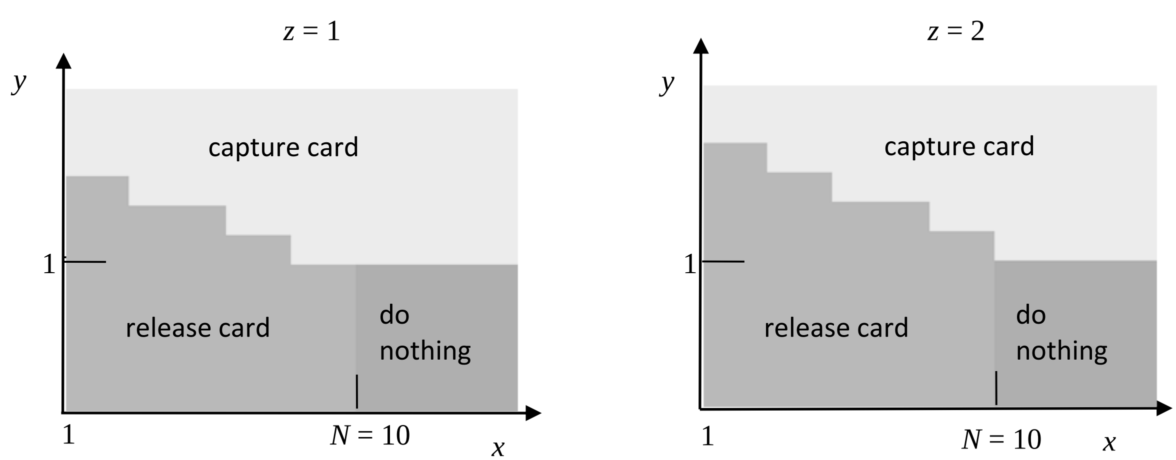

Fig 5.2. Optimal adaptive Kanban-type policy at the occurrence of arrival events for $\lambda_1 = 1/6$ (on the left) and $\lambda_2 = 1/4$ (on the right).

The optimal decisions when the current state is (x, y, z) and demand arrival occurs are interpreted as follows. If the finished goods inventory is relatively high, then the demand is satisfied but no production authorization is issued to replenish stock. The same decision is made when the finished goods inventory is at relatively low levels and the number of working machines is close to N. On the other hand, if the finished goods inventory is relatively low and few machines are currently working, then a standard (Kanban-type) production authorization is issued to replenish the item that exited the finished goods buffer plus an additional one (card release). If $y < 1$ at the time of a demand arrival and the number of working machines is less than N, an extra kanban card is added to the system. However, if all machines are working then the optimal decision is to keep the number of kanban cards fixed.

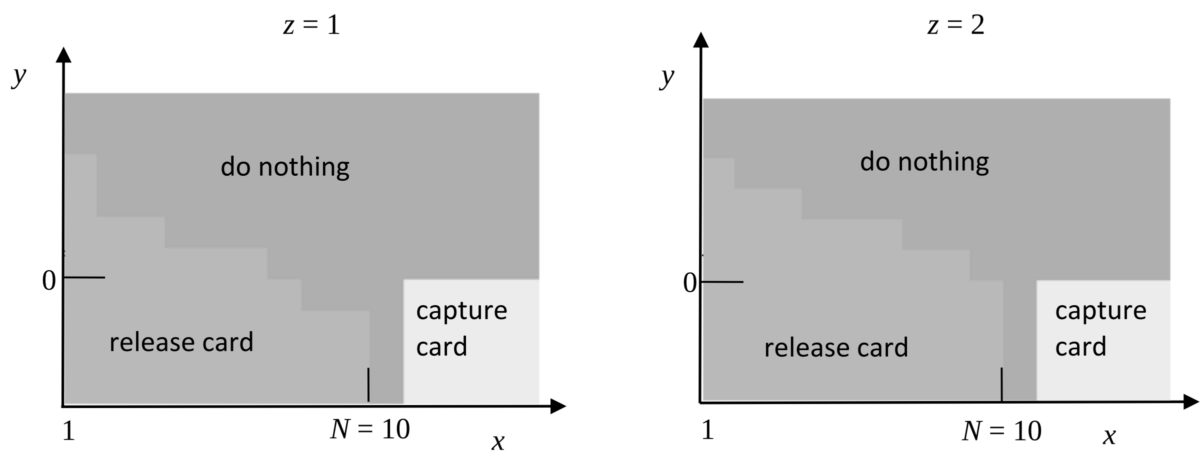

Fig 5.3. Optimal adaptive Kanban-type policy at the occurrence of production completion events for $\lambda_1 = 1/6$ (on the left) and $\lambda_2 = 1/4$ (on the right).

The interpretation of the optimal decisions when the current state is (x, y, z) and production completion occurs is given hereafter. If there is backordered demand, N Work-In-Process items and production occurs, the optimal decision is "do nothing", i.e. the produced end-item exits the system immediately to satisfy demand and the related kanban card is sent back to the manufacturing facility to authorize the production of a new part. However, under the same conditions, if the Work-In-Process is greater than N, the kanban card that is detached from the finished end-item is captured. In general, if the inventory position y is relatively low and few machines are working, the "release card" decision is the optimal at the event of production completion. In the case of a negative y this means that two kanban cards (a standard Kanban-type and an extra card) are sent to the manufacturing facility whereas in the case of $y \geq 0$ a single card is issued to authorize production of a new part. The "do nothing" decision when $y \geq 0$ ($y<0$) means that no (one) card is sent to the manufacturing facility.

It is important to stress the fact that, in the analysis of the optimal policy's structure, the meaning of the terms "relatively high/low inventory" depends on the current demand level that is imposed on the manufacturing system. This is made evident by comparing the left-hand/right-hand plots in figures 5.2 and 5.3. For example, in figure 5.2, it is seen that the higher the arrival rate is, the higher the switching curve of decision "capture card" and "release card" in the x-y plane is positioned. On the other hand, it is never optimal for x to be greater than N because this would lead to an unnecessary build-up of Work-In-Process, since the manufacturing system consists of a single stage.

5.4 Comparison between Kanban, Extended Kanban, Adaptive Kanban, FSL and the optimal adaptive control policy

In this section, we examine the behavior of the alternative adaptive heuristics in relation to the optimal policy. Two series of experiments are carried out: one that pertains to a single demand level and one that pertains to multiple demand levels. A comparative evaluation of the investigated control mechanisms is carried out for various levels of production system workload and cost factors.

The optimal cost for all optimization problem instances is calculated by solving the Dynamic Programming model of section 5.3 for the cases with $M > 1$ and $M = 1$, respectively. The performance of the heuristic production control policies is assessed by means of discrete-event simulation (Xanthopoulos et al., 2016). 20 independent replications of each simulation model were executed and each replication lasted 1000000 time units in order for the system to reach steady state and for statistically significant results to be obtained. The simulation models were also verified by comparing their output against published results. To obtain the best parameters for each control policy, exhaustive search over the space of reasonable solutions was carried out. For the continuous parameter T_{SL} of the FSL policy, the search space was discretized using a step equal to 0.01.

We use the base case of Tardif & Maaseidvaag (2001) as the starting point of our analysis because it is often used as benchmark in relevant publications ($N = 10$, $M = 1$, $\lambda = 1/5$, $\mu = 1/6$, $h_1 = 1$, $h_2 = 1$, $b = 1000$). We vary parameters μ, h_1, h_2, b one at a time and observe the effect on the production system output. Figures 5.4 – 5.5 show the % relative difference between the best average cost attained by each heuristic control policy and the optimal one. In all cases with $M = 1$, the % relative difference between the optimal cost and the cost attained by Kanban, Extended Kanban, Adaptive Kanban and FSL is found to be in the range [4.25, 22.76], [3.72, 18.42], [0.69, 13.31] and [4.14, 22.49], respectively.

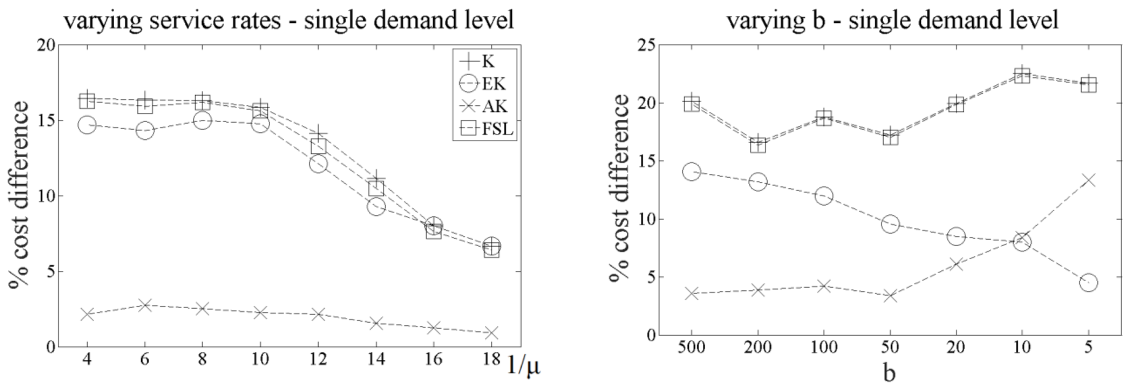

Fig 5.4. Comparison between the examined control policies and the optimal for cases with a single demand level in respect to varying $1/\mu$ (on the left) and b (on the right). K, EK, and AK stand for Kanban, Extended Kanban and Adaptive Kanban, respectively.

The Adaptive Kanban policy is found to be a very good approximation of the optimal policy in most cases that pertain to a single demand level. It outperforms all other heuristics in all cases except two. Extended Kanban is superior to the Adaptive Kanban policy in cases where parameter b is relatively low. For M = 1, the Extended Kanban policy is found to be a reasonably good approximation of the optimal adaptive policy. It ranks second in this group of adaptive Kanban-type policies since it is outperformed in all but two cases by Adaptive Kanban and outperforms FSL in all cases except those with $1/\mu$ = 16 and $1/\mu$ = 18. With the exception of these two aforementioned cases, the FSL policy is outperformed by both Extended Kanban and Adaptive Kanban in all cases. The only cases where the cost attained by FSL approximates the optimal one is when the service rates μ is relatively low or, the Work-In-Process costs are high. The performance of the FSL policy is only marginally superior to that of the standard Kanban policy which ranks last for all cases with a single demand level.

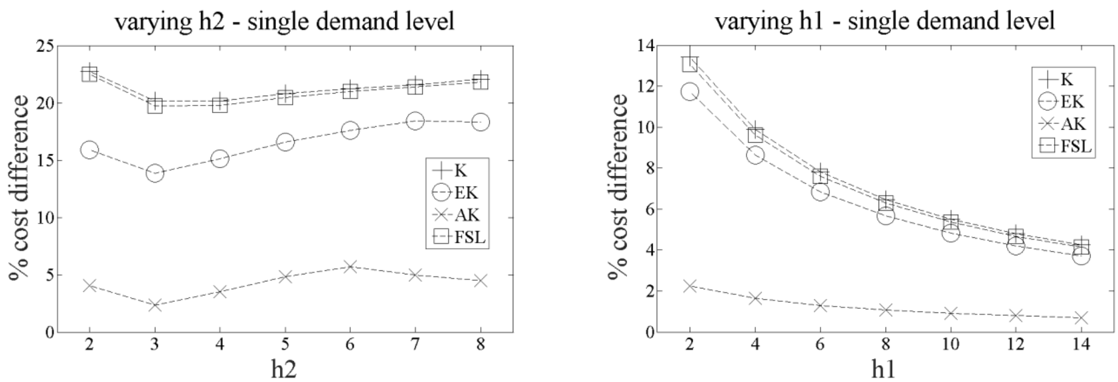

Fig 5.5. Comparison between the examined control policies and the optimal for cases with a single demand level in respect to varying h_2 (on the left) and h_1 (on the right). K, EK, and AK stand for Kanban, Extended Kanban and Adaptive Kanban, respectively.

In the next series of experiments, we examine a system with five demand levels. The base here is described by the following parameter set: N = 10, M = 5, λ_1 = 1/4, λ_2 = 1/4.5, λ_3 = 1/5, λ_4 = 1/5.5, λ_5 = 1/6, μ = 1/15, θ = 1/100, h_1 = 1, h_2 = 2, b = 100. We examine the optimal cost and the performance of the four heuristic policies in relation to varying parameters μ, b, h_1, h_2, one at a time.

Figures 5.6 – 5.7 show the % relative difference between the optimal cost and the best cost attained by each heuristic. Over all cases with M = 5, the % relative difference between the optimal cost and the best cost of Kanban, Extended Kanban, Adaptive Kanban and FSL is 10.45 – 44.08%, 7.99 – 21.58%, 2.1 – 12.44%, 10.31 – 44.24% respectively.

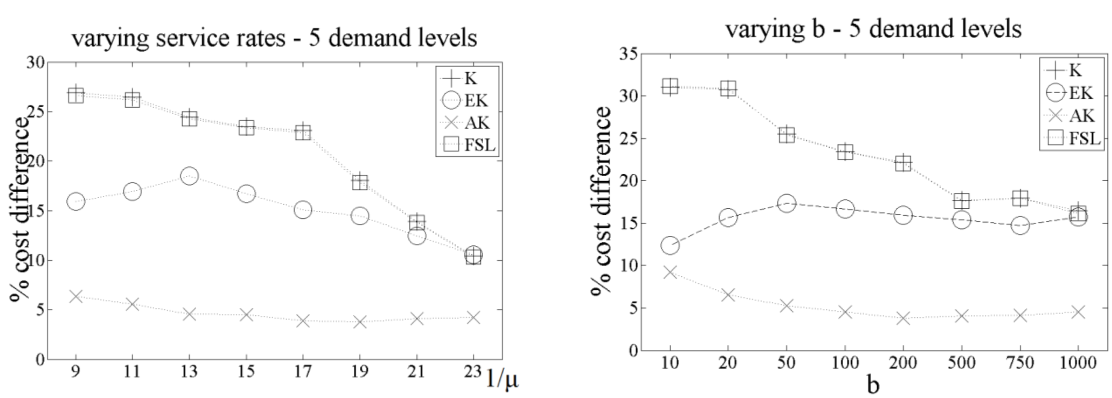

Fig 5.6. Comparison between the examined control policies and the optimal for cases with five demand levels in respect to varying $1/\mu$ (on the left) and b (on the right).

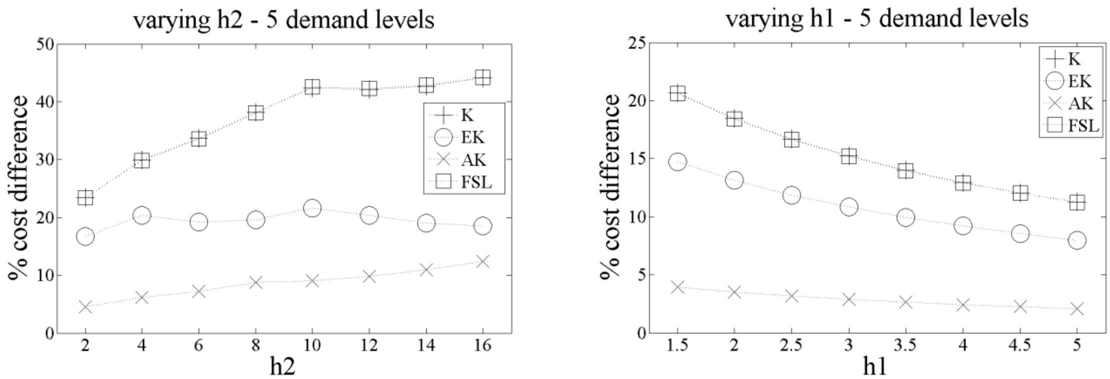

Fig 5.7. Comparison between the examined control policies and the optimal for cases with five demand levels in respect to varying h_2 (on the left) and h_1 (on the right). K, EK, and AK stand for Kanban, Extended Kanban and Adaptive Kanban, respectively.

In general, for more than one demand levels, the performance of Adaptive Kanban deteriorates in relation to the cases with $M = 1$. Nonetheless, figures 5.6 and 5.7 show that this adaptive mechanism is still a good approximation of the optimal policy and outperforms all other heuristics in all cases. The Extended Kanban policy ranks second in this series of experiments also, as it outperforms both Kanban and FSL in all cases except one. FSL manages to outperform Extended Kanban only in the case with $1/\mu = 23$. In all other cases with M = 5, the FSL policy is, at best, marginally superior to the standard Kanban mechanism. Furthermore, FSL together with Kanban are the control policies that their performance is severely affected by the increase of demand variability, compared to Extended Kanban and Adaptive Kanban.

5.5 Synopsis of experimental findings

In this section we highlight and comment on i) the similarities and ii) the differences between behavior of the examined control policies for the cases with a single and multiple demand levels.

Regardless of the number of demand levels, the average total cost is shown to be a decreasing function of the mean service rate μ and, an increasing function of the cost parameters h_1, h_2, b. The performance of all heuristic control policies is adversely affected by the increase of the demand variability. For example, the average relative difference between the optimal cost and the cost attained by Extended Kanban for the cases with $M = 1$ and $M = 5$ is 9.95% and 15.18%, respectively. This is explained by the properties of the optimal policy that were discussed in a previous section of chapter 5 and by the fact that none of the heuristic policies is *self-adaptive*, i.e.

no heuristic has a mechanism for adapting its *own* control parameters. All existing adaptive control policies can achieve near-optimal results only in cases where the variability of the inter-arrival times is rather limited.

6. OPTIMAL ADAPTIVE PULL-TYPE CONTROL POLICIES FOR MULTI-STAGE MANUFACTURING SYSTEMS

In this chapter we extend the research that was reported in the previous chapter 5 in two ways. First, we study manufacturing systems that are more complex than the single-stage system of chapter 5, i.e. serial production lines.

The dynamic behavior of the manufacturing system is obtained by means of discrete-event simulation (refer to Ch. 1 of this review). The second way in which this chapter is differentiated from the former is the solution approach. More specifically, in this chapter we derive optimal or near-optimal adaptive pull-type control policies using a Machine Learning technique known as Reinforcement Learning (RL).

6.1. Formal definition of a multi-stage manufacturing system

For the purposes of this research we consider a system that manufactures a single type of products. Raw materials are used to manufacture a product and the supply of raw materials is assumed to be infinite. The manufacturing system is comprised of several stages in series. In order for products to be manufactured, raw parts enter the first stage and are processed in all stages starting from the first one and finishing to the last stage. A part that completes its processing in some stage is always forwarded to the downstream stage. Processing times are assumed to be stochastic, so as to model unexpected disruptions in the production process.

Finished products are placed in the finished products buffer. Demands for finished products arrive dynamically to the system and at random time intervals. At the time of a demand arrival, one product is requested from the finished goods inventory. The demand is satisfied instantly if there is available inventory at that time. Otherwise, the demand is backordered until a new finished product is completed. It follows that the backorders queue and the finished products buffer cannot be non-empty at the same time. The backorders queue is assumed to be of infinite capacity and the backorders queue discipline is FCFS.

Each stage of the examined manufacturing system is comprised of a manufacturing facility and an output buffer; refer to figure 6.1 for a graphical example. A stage-*i* finished part is placed in the stage-*i* buffer at the time when it is completed. Clearly, the finished products buffer is the output buffer of the last stage. A stage-*i* completed part remains in the respective buffer up to the point when it is authorized by a *CONWIP-type production control policy* to be forwarded to the downstream stage. A manufacturing facility can process parts one-by-one; if it is currently busy then pending parts are held in a wait queue. All wait queues are assumed to have infinite capacity and abide by the FCFS discipline.

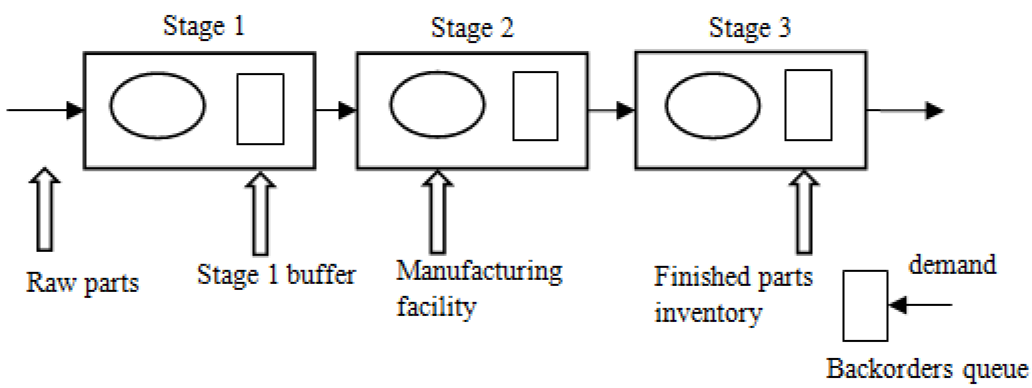

Fig 6.1. A production system with three stages in series.

6.2 CONWIP-type control policies

In this section we outline the CONWIP production control policy and two existing, adaptive CONWIP-type policies.

A CONWIP system is completely characterized by parameter *K* which equals the total number of Work-In-Process parts plus finished products that are allowed in the system at any time. The CONWIP system operates as follows: i) all stages except the first are constantly authorized to produce and ii) the first stage is granted the authorization to start manufacturing a new part at the time when a finished product exits the finished goods buffer to satisfy a demand.

In the work of Framinan et al. (2006), a heuristic approach is proposed for adjusting the maximum allowed finished goods inventory dynamically and in response to the current state of the system. In the remainder of this chapter, we will refer to this approach as Dynamic CONWIP for brevity. The control parameters of Dynamic CONWIP are the non-negative integers *K*, *E* and the non-negative real *T*. Let P_n^{max} denote the current, maximum allowed inventory of finished products and *TH* denote the current throughput of the manufacturing system. At time *t*, the throughput is evaluated as: *TH* = (number of finished products completed up to *t*)/*t*. In a Dynamic CONWIP system, the initial finished goods inventory is *K* (similarly to a standard CONWIP system) but

P_n^{max} can vary in the range $\{K, K+1, \ldots, K+E\}$. The adjustments of P_n^{max} are guided by parameter T which is the *throughput target*. Dynamic CONWIP works as a standard CONWIP system with one major modification: the throughput TH is evaluated and then compared to the target T, at the completion of a new finished product; a decision to increase/decrease/keep fixed the maximum allowed finished goods inventory P_n^{max} is made based on this comparison.

The single-loop, Generalized Kanban system bares resemblance to Dynamic CONWIP in the sense that it operates similarly to a standard CONWIP system but it also has a mechanism for increasing/decreasing the maximum allowed Wok-In-Process parts. Parameters K and E are the only control parameters of the single-loop Generalized Kanban system. Their physical meaning is similar to those of the Dynamic CONWIP; K is the initial finished goods inventory and E is the maximum number of *additional* parts that are allowed to "circulate" in the system.

6.3 RL-based adaptive CONWIP system

In this research, we propose an intelligent approach for deriving adaptive CONWIP-type control policies. The derived control policies have a fairly simple structure. The manufacturing system operates as if under a standard CONWIP policy but at certain system states production authorizations can be either "released to" or "captured from" the system. Releasing a production authorization to the system means that the first stage is authorized to start working on a new part while no part exits the finished goods inventory. Capturing a production authorization means that no authorization is sent to the first stage when a part exits the finished goods inventory.

The major novelty of the proposed approach is that decisions on capturing/releasing production authorizations are not ad-hoc/intuitive, as is the case in the Dynamic CONWIP and the Generalized Kanban systems. On the contrary, relevant decisions are the outcome of an optimization process widely known as Reinforcement Learning (Kaelbling et al., 1996; Sutton & Barto, 1998). The proposed approach is represented graphically in figure 6.2.

At some time point, a decision-making agent (RL agent) receives a representation of the manufacturing system's current state. The agent decides to capture a production authorization, or release a production authorization, or do nothing. Then, the manufacturing system transits to a new state and the associated *reward* (or cost) of the agent's decision is conveyed to it. This cycle is repeated "sufficiently" many times. Eventually, the decision-making agent converges to the optimal production control decisions in terms of some objective function.

The manufacturing system's state transitions and the reward signal for the agent's decisions are generated by discrete-event simulation models.

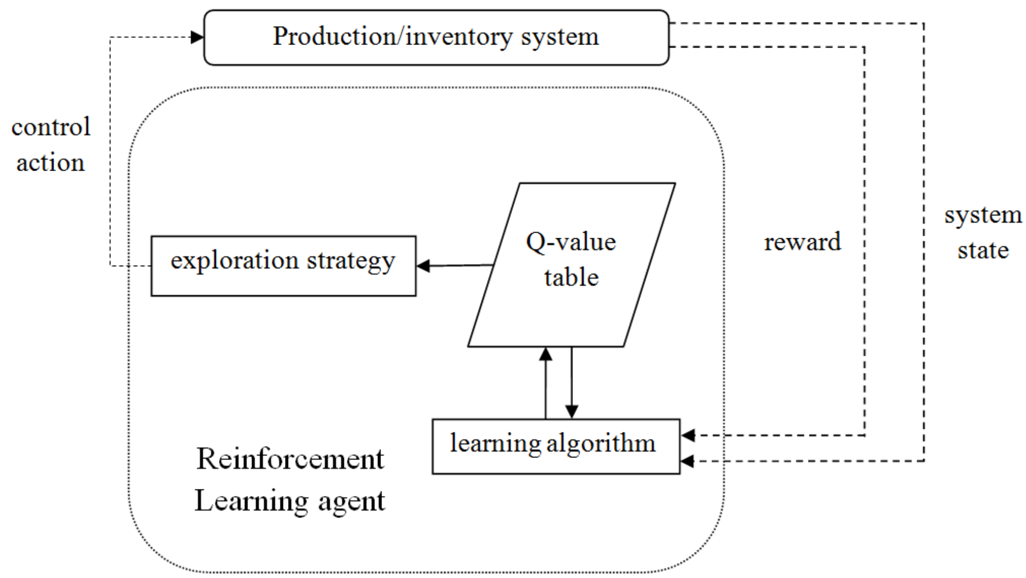

Fig 6.2. The solution approach for obtaining optimal adaptive control policies of multi-stage production systems.

In order to fully describe the RL-simulation interface that is depicted in figure 6.2 one needs to define:

- The representation of the system state, as it is perceived by the RL module
- The decision epochs (the time points where the RL agent interacts with the controlled system) and the admissible actions
- The objective function that is to be optimized and the mechanism according to which, numerical rewards are generated for the RL agent
- The learning algorithm and the exploration strategy of the RL module

6.4 Summary of numerical results

A series of simulation experiments was conducted in order to assess the performance of the various adaptive, CONWIP-type control policies. All simulation experiments pertained to manufacturing systems with three stages in series where each stage has a single machine. The service times of all machines are assumed to be exponential with mean 1.0. The times between demand arrivals are also exponentially distributed. Four simulation cases were examined that differ in terms of the volume of demand for finished products. The mean time between arrivals (MTBA) in the four simulation cases was 1.2, 1.4, 1.6 and 1.8, respectively. The CONWIP, Dynamic CONWIP, single-loop Generalized Kanban, and the proposed RL-based approach were compared in terms of the total average cost function, i.e. the long-run average (Work-In-Process inventory + finished goods inventory + number of backorders).

In this set of experiments, the cost function is seen to be decreasing in respect to the mean time between arrivals. This can be attributed to the fact that the lower the arrival rate is, the more easily the manufacturing system can cope with incoming demand. This results in fewer backordered demands on the average. In cases where the manufacturing system can easily satisfy demand, relatively low stock is also needed to hedge against fluctuations of the demand and production process. Consequently, the holding cost component of the objective function is also kept at relatively low levels.

In this series of experiments, the RL-based approach outperforms all other control policies in all cases. This is an indication that Reinforcement Learning is a good option for obtaining optimal or near-optimal production control policies for stochastic manufacturing systems. The Generalized Kanban system ranks second as it outperforms both the Dynamic and the standard CONWIP methods in all simulation cases. The CONWIP policy ranks last and this shows that the adaptive counterparts of standard pull-type control policies are highly likely to exhibit superior performance. Finally, the Dynamic CONWIP policy ranks third in this experimental as it is marginally superior to the standard CONWIP policy.

It is observed that the differences between the alternative control policies become more apparent in cases where the manufacturing system is under a relatively heavy workload, i.e. in cases with MTBA = 1.2 and 1.4. The RL-based approach is a stochastic optimization procedure and therefore, it is bound to yield favorable results. However, it has a relatively large number of parameters that need to be fine-tuned and it also has non-negligible computational requirements. This experimental trial shows that the single-loop Generalized Kanban system is a reasonable approximation of the optimal or near-optimal policy found by the RL agent in cases with MTBA = 1.6 and 1.8. Consequently, a practical rule of thumb could be suggested, i.e. to use the Generalized Kanban method as an approximation of the optimized, adaptive control policy in cases where the volume of finished products is relatively low.

7. INTEGRATED PRODUCTION, INSPECTION AND MAINTENANCE CONTROL IN STOCHASTIC MANUFACTURING SYSTEMS

This chapter studies an integrated production/inspection/maintenance control problem. Integrated or joint optimization problems pertaining to maintenance have been studied extensively over the years (e.g. refer to Zhou et al. 2019; Li and Ma 2017; Xu and Xu 2016). Relevant literature surveys can be found in the works of Ruschel et al. (2017) and Prajapati et al, (2012). The production/inventory system and the underlying control problem which is studied in this research has not been addressed in the literature up to now. Nonetheless, some of its individual features can be found in existing works. In this section, we categorize publications that are most relevant. Our aim is to show how this chapter is "positioned" in relation to existing works, demonstrate the gaps in the literature which are addressed by this research and highlight its contribution.

We examine a *single machine* system that produces a *single part type*, similarly to the works of Koutras et al. (2017) and Iravani and Duenyas (2002). To model the manufacturing system in question, the commonly adopted formalism of Markov chains is used (indicatively refer to Cekyay and Ozekici 2012; Rao and Naikan 2009; Liang and Parlikad 2015). We employ the machine deterioration scheme, according to which, the system's condition is discretized using a number of *deterioration stages* (Pavitsos and Kyriakidis 2009; Kazaz and Sloan 2013; Xanthopoulos et al 2015). Apart from the stage where the machine has experienced a hard failure, i.e. is broken down, its current deterioration is only made known by means of *periodic inspections* (Mousavi et al. 2017; Golmakani and Fattahipour 2011; Soemadi et al. 2014; Lee 2009).

Preventive maintenance activities are authorized on the basis of the current deterioration level of the machine, a scheme widely known as *condition-based maintenance* (Van and Berenguer 2012; Rausch and Liao 2010; Prajapati and Ganesan 2013; Jiang et al. 2018). When the machine is broken down, repair, i.e. corrective maintenance, activities are undertaken (Zhang and Sun 2018). Regarding the production control component of the joint optimization problem which is addressed in this research, we examine (s, S) – type (Gosavi et al. 2004) and *base stock* – type (Axsater 2015; Cheng et al. 2011) policies. Production (Axsater 2015; Cheng et al. 2011; Peng and van Houtum 2016), inventory Lee (2009), lost sales (Cheng et al. 2011), inspections, maintenance (Wolter and Helber 2016; Najid et al. 2011) and repair costs are considered in this research.

7.1 System description

The manufacturing system in question is comprised of a single machine and a finished goods buffer. It manufactures a single type of products. The supply of raw materials is assumed to be infinite, i.e. the machine never starves. The machine manufactures products one-by-one, with no preemption. The processing times are exponentially distributed with mean $1/\lambda_p$. Manufactured products are stored in the finished goods buffer. The cost of storing one product in the buffer for one time unit is C_h.

Demand for finished goods is stochastic and the demand inter-arrival time is exponential with mean $1/\lambda_a$. Upon a demand arrival, one unit of the product is requested. If there is available inventory, the demand is satisfied instantly, otherwise the demand is rejected. Rejecting a demand for a finished product costs C_a monetary units.

The machine deteriorates with usage, i.e. when it is working to manufacture products. If the machine is "good-as-new", then it is considered to be in *deterioration stage* 0. If the machine is broken-down and under repair, then it is in deterioration stage $d + 1$. In the case where the current deterioration stage is $0 < i < d + 1$, then the machine is deteriorated yet still operational. When operating, the system transits from stage i to stage $i + 1$ with rate λ_f, for $i = 0, 1,\ldots, d$. This type of transition will be referred to as "deterioration failure". The transition time from one deterioration stage to another is exponential. When in stage $d + 1$, the machine cannot produce, it is under repair, and it transits to stage 0 (good-as-new) with rate μ_r. Repair times are exponential and a repair incurs a cost of C_r monetary units.

The system operates under a *continuous-review* (s, S) control policy which is described by two integer parameters $s < S$. The level of the finished goods buffer is monitored continuously and, at the time when it drops to s, a new production lot (batch) is initiated. Each production lot has a fixed cost of C_s monetary units. Once the production of a new batch has commenced, the machine manufactures products up to the point where there are S items in the finished goods buffer and then switches to idle state.

In the special case where $s = S -1$, the system operates under a *base stock*, or order-up-to-S, or simply S control policy. A base stock policy is a simple, threshold-type control policy which is characterized by a single parameter, i.e. S: the machine idles as long as the finished goods inventory is greater than or equal to S, and produces otherwise. A manufacturing system controlled by the base stock policy will be called *base stock system*.

This stochastic model can be extended by considering periodic inspections of the machine's state and preventive maintenance. In the remainder of this section we define the additional elements for this extension.

Unless the machine has broken down ($i = d + 1$), the current deterioration stage is unknown. It is determined by conducting periodic inspections of the manufacturing system. An inspection can only be initiated when the machine is idling. The times between inspections are exponential and, when idling, the system transits to the "under inspection" state with rate λ_{in}. The cost for an inspection is C_{in} and its duration is an exponentially distributed random variable with mean $1/\mu_{in}$.

Upon completion of an inspection, the current deterioration of the machine is established. Based on that information, it can be decided to perform a *preventive maintenance* or do nothing. A preventive maintenance costs C_m monetary units, has an exponential duration with mean $1/\mu_m$ and restores the machine to stage 0 (good-as-new). Maintenance decisions are made on the basis of a *threshold-type maintenance policy*. It is noted that, when the system transits to deterioration stage $d + 1$, the machine breaks down and repair actions are initiated immediately.

7.2 Markov chain models

In this section we outline the continuous-time Markov chain model for the systems that were defined in section 7.1. The state of the system is described by the tuple (i, j, k), where $i = 0, 1,\ldots, d + 1$ is the deterioration stage of the machine, $j = 0, 1,\ldots, S$ denotes the number of finished products in stock, and k symbolizes the machines state:

$$k = \begin{cases} 0, & \text{if machine is idling} \\ 1, & \text{if machine is working} \\ 2, & \text{if machine is repaired} \\ 3, & \text{if machine is inspected} \\ 4, & \text{if machine is maintained} \end{cases} \tag{7.1}$$

It is noted that the machine is "repaired" when the current deterioration stage is $d + 1$, i.e. the machine does not have the ability to produce. On the other hand, the machine is "maintained" if it is deteriorated, yet operational. The investigated manufacturing system is a *discrete-event* system, i.e. its state changes only at the occurrence of specific types of events. The only events that can occur are “demand arrival”, “production completion”, “deterioration failure” and “repair completion”, in addition to the events “inspection start”, “inspection completion”, and “maintenance completion” if inspections/maintenance works are considered.

7.3 Performance metrics and cost function

The steady-state probabilities for all states are computed by solving the associated Markov chains. Several performance metrics can be defined as combinations of specific steady-state probabilities. In this research we examine the following:

- Average demand rejection rate
- Average rate with which new production lots are initiated
- Average rate with which hard failures occur
- Average finished goods inventory
- Average rate of inspections
- Average maintenance rate

Our goal is to minimize the total expected cost which is a linear combination of all cost components.

7.4 Outline of solution method – optimization algorithm

We investigated numerically the properties of the cost function in respect to the control parameters of the production/maintenance/inspection policies. These experiments gave indications that the objective function is unimodal, i.e. it has a unique minimum. Of course, this lacks the vigor of a mathematical proof but still, it justifies the use of a local search procedure to obtain a good approximation of the optimal solution.

The proposed search procedure is a steepest descent algorithm with minimal requirements in terms of computational burden. Initially, a feasible, candidate solution is set arbitrarily. Then, the neighborhood of the candidate solution is examined. The neighbor that offers the greatest reduction of the cost function is selected as the next candidate solution. The procedure continues to iterate up to the point where no further improvement can be achieved, in terms of the cost function value.

Provided that the cost function is unimodal and that the steps of the search are small enough, the proposed procedure will converge to the optimal solution. If the objective function is multimodal, however, the procedure might get trapped in a local minimum and produce sub-optimal results. The performance of the proposed procedure and of the cost function's properties were probe by analyzing an extended set of numerical experiments.

8. CONCLUSIONS

This chapter concludes the postdoctoral research "Modelling, control and optimization of stochastic production systems". We summarize the scientific findings for each of the research directions 1-4 that were stated in chapter 2. In the following sections we also outline avenues for future research.

8.1. Summary of scientific findings for chapter 4 and directions for future research

We developed the queueing network models of the CONWIP, Base Stock and CONWIP/Kanban Hybrid control policies for multi-product manufacturing systems. The conversion from a single-product to a multi-product system is rather straightforward however, the model complexity increases dramatically for more than one part types.

The results of the simulation experiments indicated that the defining characteristics of the CONWIP, Base Stock and CONWIP/Kanban Hybrid policies for single-model systems are largely retained in multi-product systems too. The Base Stock strategy outperforms CONWIP and CONWIP/Kanban hybrid when the manufacturing system is under heavy workload. However, when the system operates under a moderate workload, the differences between the alternative control policies are rather negligible. The CONWIP/Kanban Hybrid policy is by far the more sensitive in respect to changes in the average demand rate and setup time. On the other hand, the CONWIP policy was found to be the mostly affected by changes of the control parameters.

There are several ways to extend this research. A straightforward extension is to consider additional pull control policies and conduct larger scale simulation experiments. An even more interesting extension would be to study the synergy of applying specific priority rules for Work-In-Process sequencing and production control policies in mixed-model systems. This is because, in the field of pull type production control, the First-Come-First-Served queue discipline rule is almost invariably assumed. On the other hand, in the field of dynamic sequencing with sequence-dependent setups, coordination among production stages is typically overlooked.

8.2. Summary of scientific findings for chapter 5 and directions for future research

We examined a single-stage production/inventory system with parallel and identical machines that operates under an adaptive Kanban-type control policy. The optimal policy in respect to minimizing backorder and holding costs was obtained by formulating the problem as a Markov Decision Process and solving it with a Dynamic Programming approach. The properties of the optimal policy were investigated numerically. This investigation revealed that well-known, adaptive approaches such as the Adaptive Kanban mechanism can never be optimal for seasonal demand. The optimal policy was compared to three heuristic, adaptive policies and to the standard Kanban policy in a series of simulation experiments.

The optimal adaptive policy offered reductions in the expected average cost up to 44.08% compared to the standard Kanban policy, showcasing the benefits of dynamic card number adjustment. The performance of all heuristic control policies deteriorated substantially even with a moderate increase of the demand variability. All adaptive control policies cannot achieve near-optimal results when the variability of the inter-arrival times is significant. The Adaptive Kanban policy was found to be a good approximation of the optimal policy in all cases with the exception of situations where the backorders cost factor was relatively low. The Extended Kanban ranked second and it was also found to be a reasonable approximation of the optimal policy mostly in cases with low service rates, low backorder costs or, high Work-In-Process costs. The FSL policy marginally outperformed the standard Kanban policy in all cases and ranked third in this experimental trial.

There are several straightforward ways to extend this research. A plausible direction of future research is to study a problem where the average service rates vary dynamically with time due to, e.g. machine deterioration and breakdowns. An even more challenging extension would be to consider non-stationary production and arrival processes with unknown characteristics. In that case, statistical and/or signal processing techniques should be utilized so that the current system state becomes known to the decision-maker.

8.3. Summary of scientific findings for chapter 6 and directions for future research

This research proposes an approach for developing adaptive CONWIP-type control policies based on an optimization paradigm called Reinforcement Learning. The proposed approach is applied to manufacturing systems that consist of a number of stages in series. The aim is to minimize the long-run average holding and backorder costs. The proposed approach is compared to the standard CONWIP, the Dynamic CONWIP and a special case of the Generalized Kanban system.

Our experimental results show that the proposed approach outperforms all alternative control policies in all simulation cases. The Generalized Kanban policy is found to be a good approximation of the control policies that are derived by the RL approach in cases where the demand rate is relatively low.

This research can be extended by studying the properties of the optimal or near-optimal policies which are derived by the Reinforcement Learning approach. Insights can be gained by this analysis, and on that basis improved heuristics can be proposed. A more detailed performance evaluation of existing control policies would also be interesting. Finally, another direction of research would be to study adaptive production policies in the context of non-stationary manufacturing systems.

8.4 Summary of scientific findings for chapter 7 and directions for future research

We examined a deteriorating production/inventory system under different production/maintenance policies. Some of the findings, that resulted from the interpretation of the experimental results, can be generalized and provide guidelines for establishing efficient control policies in arbitrary, manufacturing settings.

It was observed that periodic inspections/preventive maintenance can generally improve the systems performance. However, if the repair rate and/or the repair costs are "sufficiently" high and low, respectively, the benefits from inspecting/maintaining the system could be counterbalanced by the associated costs.

A straightforward way to extent this line of research is to provide additional theoretical results in respect to the total expected cost function. For example, it would be interesting to obtain a closed-form expression for the cost function in respect to control policy parameters. A theoretical analysis of how to optimize the control parameters of the examined production/maintenance policies based on mathematically proven properties of the cost function would also be helpful. Another plausible direction for future research is the examination of multi-product and/or multi-stage systems. Nonetheless, in such a case, the complexity of the manufacturing system would dramatically increase, rendering analytical approaches intractable. For the examination of such systems, resorting to simulation would probably be inevitable.

REFERENCES

[1] Axsater, S. (2015). *Inventory Control*. New York: Springer.

[2] Belisario LS, Pierreval H (2015) Using genetic programming and simulation to learn how to dynamically adapt the number of cards in reactive pull systems. Expert Syst Appl42: 3129-3141.https://doi.org/10.1016/j.eswa.2014.11.052

[3] Bertsekas DP(1995)Dynamic Programming and Optimal Control.Athena Scientific, Belmont, Massachusetts

[4] Buzacott JA, Shanthikumar JG (1993) Stochastic models of manufacturing systems. Prentice-Hall, Englewood Cliffs, NJ

[5] Cekyay, B., Ozekici, S. (2012). Performance measures for systems with markovian missions and aging. *IEEE Transactions on Reliability*, 61(3), 769-778.

[6] Cheng, T.C.E., Gao, C., Shen, H. (2011). Production and inventory rationing in a make-to-stock system with a failure-prone machine and lost sales. *IEEE Transactions on Automatic Control*, 56(5), 1176-1180.

[7] Dallery Y, Liberopoulos G (2000) Extended Kanban control system: Combining Kanban and base stock. IIE Trans32(4): 369–386. https://doi.org/10.1023/A:1007651721842

[8] Framinan JM, Gonzalez PL, Ruiz-Usano R (2006) Dynamic card controlling in a Conwip system. Int J Prod Econ99: 102-116.https://doi.org/10.1016/j.ijpe.2004.12.010

[9] Golmakani, H.R., Fattahipour, F. (2011). Optimal replacement policy and inspection interval for condition-based maintenance. *International Journal of Production Research*, 49(17), 5153-5167.

[10] Gosavi, A., Das, T.K., Sarkar, S. (2004). A simulation based learning automata framework for solving semi-Markov decision problems under long-run average reward. *IIE Transactions*, 36(6), 557-567.

[11] Gupta SM, Al-Turki YAY, Perry RF (1999) Flexible Kanban System.Int J Oper Prod Man19(10): 1065-1093.https://doi.org/10.1108/01443579910271700

[12] Hopp WJ, Roof ML (1998) Setting WIP levels with statistical throughput control (STC) in CONWIP production lines. Int J Prod Res36(4): 867-882.DOI: 10.1080/002075498193435

[13] Iravani, S.M.R., Duenyas, I. (2002). Integrated maintenance and production control of a deteriorating production system. *IIE Transactions*, 34(5), 423-435.

[14] Jiang, A., Dong, N., Tam, K.L., Lyu, C. (2018). Development and optimization of a condition-based maintenance policy with sustainability requirements for production system. *Mathematical Problems in Engineering*, DOI: 10.1155/2018/4187575

[15] Kaelbling LP, Littman ML, Moore AW. Reinforcement Learning: A survey. Journal of Artif Intell Res. 1996; 4: 237-285.

[16] Kazaz, B., Sloan, T.W. (2013). The impact of process deterioration on production and maintenance policies. *European Journal of Operational Research*, 227, 88-100.

[17] King D. H., Harrison H. S. (2013). “Open-source simulation software JaamSim”, In: R. Pasupathy, S.H. Kim, A. Tolk, R. Hill, M. E. Kuhl, eds, *Proceedings of the 2013 Winter Simulation Conference*, pp. 2163-2171, IEEE.

[18] Korugan A, Gupta SM (2014) An adaptive CONWIP mechanism for hybrid production systems.Int J Adv Manuf Technol 74:715-727.https://doi.org/10.1007/s00170-014-5977-2

[19] Koulouriotis D. E., Xanthopoulos A. S., Tourassis V. D. (2010), “Simulation optimization of pull control policies for serial manufacturing lines and assembly manufacturing systems using genetic algorithms”, *International Journal of Production Research*, Vol. 48, No. 11, pp. 2887-2912.

[20] Koutras, V.P., Malefaki, S., Platis, A.N. (2017). Optimization of the dependability and performance measures of a generic model for multi-state deteriorating systems under maintenance. *Reliability Engineering and System Safety*, 166, 73-86.

[21] Lee, T.H. (2009). Optimal production run length and maintenance schedule for a deteriorating production system. *International Journal of Advanced Manufacturing Technology*, 43, 959-963.

[22] Li, R., Ma, H. (2017). Integrating preventive maintenance planning and production scheduling under reentrant job shop. *Mathematical Problems in Engineering*, DOI: 10.1155/2017/6758147

[23] Liang, Z., Parlikad, A.K. (2015). A tiered modelling approach for condition-based maintenance of industrial assets with load sharing interaction and fault propagation. *IMA Journal of Management Mathematics*, 26, 125-144.

[24] Liberopoulos G. (2013), “Production release control: paced, WIP-based or demand driven? Revisiting the push/pull and make-to-order/make-to-stock distinctions”. In: J. Smith, B. Tan, eds, *Handbook of stochastic models and analysis of manufacturing system operations*, International Series in Operations Research & Management Science, Vol. 192, New York: Springer.

[25] Liu Q, Huang D (2009) Dynamic card number adjusting strategy in card-based production system. Int J Prod Res 47(21): 6037-6050.https://doi.org/10.1080/00207540802192134

[26] Mousavi, S.M., Shams, H., Ahmadi, S. (2017). Simultaneous optimization of repair and control-limit policy in condition-based maintenance. *Journal of Intelligent Manufacturing*, 28, 245-254.

[27] Najid, N.M., Alaoui-Selsouli, M., Mohafid, A. (2011). An integrated production and maintenance planning model with time windows and shortage cost. *International Journal of Production Research*, 49(8), 2265-2283.

[28] Onyeocha C. E., Khoury J., Geraghty J. (2015a), “Evaluation of multi-product lean manufacturing systems with setup and erratic demand”, *Computers & Industrial Engineering*, Vol. 87, pp. 465-480.

[29] Onyeocha C.E., Khoury J., Geraghty J. (2015b), “Robustness analysis of pull strategies in multi-product systems”, *Journal of Industrial Engineering and Management*, Vol. 8, No. 4, pp. 1125-1161.

[30] Onyeocha, C. E., Wang, J., Khoury, J., Geraghty, J., 2015. A comparison of HK-Conwip and BK-Conwip control strategies in a multi-product manufacturing system. Operations Research Perspectives, 2, 137-149.

[31] Pavitsos, A., Kyriakidis, E.G. (2009). Markov decision models for the optimal maintenance of a production unit with an upstream buffer. *Computers & Operations Research*, 36, 1993-2006.

[32] Peng, H., van Houtum, G.J. (2016). Joint optimization of condition-based maintenance and production lot-sizing. *European Journal of Operational Research*, 253, 94-107.

[33] Prajapati, A., Bechtel, J., Ganesan, S. (2012). Condition based maintenance: a survey. *Journal of Quality in Maintenance Engineering*, 18(4), 384-400.

[34] Prajapati, A., Ganesan, S. (2013). Application of statistical techniques and neural networks in condition-based maintenance. *Quality and Reliability Engineering International*, 29, 439-461.

[35] PutermanM(1994)Markov Decision Processes: Discrete Stochastic Dynamic Programming. John Wiley and Sons, New York.

[36] Rao, P.N.S., Naikan, V.N.A. (2009). An algorithm for simultaneous optimization of parameters of condition-based preventive maintenance. *Structural Health Monitoring*, 8(1), 83-94.

[37] Rausch, M., Liao, H. (2010). Joint production and spare part inventory control strategy driven by condition based maintenance. *IEEE Transactions on Reliability*, 59(3), 507-516.

[38] Renna P, Magrino L, Zaffina R (2013) Dynamic card control strategy in pull manufacturing systems. Int J Comp Integ M 26(9): 881-894.https://doi.org/10.1080/0951192X.2013.799783

[39] Renna P (2015) A fuzzy control system to adjust the number of cards in a CONWIP-based manufacturing system. International Journal of Services and Operations Management, 20(2): 188-206.https://doi.org/10.1504/IJSOM.2015.067476

[40] Renna P. (2018), "BK-Conwip adaptive control strategies in a multi-product manufacturing system", In: Y. Khojasteh, ed., *Production Management: Advanced Models, Tools, and Applications for Pull Systems*, New York: CRC Press, Taylor & Francis Group, pp. 45-62.

[41] Sutton RS, Barto AG. Reinforcement learning: An introduction. Cambridge: MIT press; 1998.

[42] Ruschel, E., Santos, E.A.P., Loures, E.D.F.R. (2017). Industrial maintenance decision-making: A systematic literature review. *Journal of Manufacturing Systems*, 45, 180-194.

[43] Soemadi, K., Iskandar, B.P., Taroepratjeka, H. (2014). Optimal overhaul-replacement policies for a repairable machine sold with warranty. *Journal of Engineering Science and Technology*, 46(4), 465-483.

[44] Takahashi K (2003) Comparing reactive Kanban systems. Int J Prod Res 41(18): 4317-4337.https://doi.org/10.1080/0020754031000152569

[45] Takahashi K, Nakamura N (2002) Decentralized reactive Kanban system. Eur J Oper Res 139: 262-276.https://doi.org/10.1016/S0377-2217(01)00358-7

[46] Takahashi K, Doi Y, Hirotani D, Morikawa K (2014) An adaptive pull strategy for remanufacturing systems.J Intell Manuf 25: 629-645.https://doi.org/10.1007/s10845-012-0710-1

[47] Tardif V, Maaseidvaag L (2001) An adaptive approach to controlling kanban systems. Eur J Oper Res 132: 411-424.https://doi.org/10.1016/S0377-2217(00)00119-3

[48] Van, P.D., Berenguer, C. (2012). Condition-Based Maintenance with Imperfect Preventive Repairs for a Deteriorating Production System. *Quality and Reliability Engineering International*, 28, 624-633.

[49] Wolter, A., Helber, S. (2016). Simultaneous production and maintenance planning for a single capacitated resource facing both a dynamic demand and intensive wear and tear. *Central European Journal of Operations Research*, 24(3), 489-513.

[50] Xanthopoulos, A.S., Koulouriotis, D.E., Botsaris, P.N. (2015). Single-stage Kanban system with deterioration failures and condition-based preventive maintenance. *Reliability Engineering & System Safety*, 142, 111-122.

[51] Xanthopoulos AS, Koulouriotis DE, Gasteratos A (2017) Adaptive card-based production control policies. Comput Ind Eng 103: 131-144.https://doi.org/10.1016/j.cie.2016.11.019

[52] Xanthopoulos AS, Koulouriotis DE, Gasteratos A, Ioannidis S (2016) Efficient priority rules for dynamic sequencing with sequence-dependent setups. International Journal of Industrial Engineering Computations 7(3): 367-384.DOI: 10.5267/j.ijiec.2016.2.002

[53] Xanthopoulos A. S., Koulouriotis D. E. (2014), "Multi-objective optimization of production control mechanisms for multi-stage serial manufacturing-inventory systems", *International Journal of Advanced Manufacturing Technology*, Vol. 74, pp. 1507-1519.

[54] Xu, Q., Xu, J. (2016). Heuristics for the economic production quantity problem under restrictions on production and maintenance time. *Mathematical Problems in Engineering*, DOI: 10.1155/2016/8704750

[55] Zhang, R., Sun, X. (2018). Integrated production-delivery lot sizing model with limited production capacity and transportation cost considering overtime work and maintenance time. *Mathematical Problems in Engineering*, DOI: 10.1155/2018/1569029

[56] Zhou, H., Wang, S., Qi, F., Gao, S. (2019). Maintenance modeling and operation parameters optimization for complex production line under reliability constraints. *Annals of Operations Research*, https://doi.org/10.1007/s10479-019-03228-9