Memorandum COSOR 87-34

Flexible Manufacturing Systems:
background, examples and models

by

W.H.M. Zijm
Flexible Manufacturing Systems: background, examples and models

W.H.M. Zijm #

Abstract

In this paper, we discuss recent innovations in manufacturing technology and their implications on the design and control of manufacturing systems. Recognizing the need to respond properly to rapidly changing market demands, we discuss several types of flexibility that can be incorporated in our production organisation to achieve this goal. We show how the concept of a Flexible Manufacturing System (FMS) naturally arises as an attempt to combine the advantages of traditional Job Shops and dedicated production lines.

The main body of the paper is devoted to a classification of FMS problem areas and a review of models developed to understand and solve these problems. For each problem area, a number of important contributions in the literature is indicated. The reader, interested in the applications of Operations Research models but not familiar with the technical background of FMS's, will find the descriptions of some essential FMS elements useful. Some final remarks and directions for future research conclude the paper.

# Nederlandse Philips Bedrijven B.V., Centre for Quantitative Methods, Building HCM-721, p.o. Box 218, 5600 MD - Eindhoven, The Netherlands,

and

1. Introduction

The need for a more flexible response to rapidly changing market demands has caused a rather dramatic change in the philosophy on the design and the layout of production systems, as well as on production organisation and even product development. Order leadtimes and production throughput times have to be shortened, final stock and work-in-process inventory levels have to be reduced and at the same time service levels are to be increased. The implications of these objectives can be observed at many different levels in the industrial organisation.

In this paper, we focus exclusively on the implications at the shopfloor level. This paper is based on Zijm[1987] where a large number of models for the design and control of Flexible Manufacturing Systems are treated in extenso. For a discussion of the relationship between the need for an increased flexibility and methods for integral goods flow control, we refer to Zijm[1988]. It should be understood that the two subjects are closely related to each other; it makes no sense to invest in the installation of flexible technology, without adapting at the same time the entire logistics organisation to the possibilities, induced by the increased flexibility. In other words: moving towards flexible manufacturing means more than only installing automated production systems.

The main objective of this paper is to review a number of important contributions in the literature on Operations Research models developed to evaluate FMS investment decisions as well as FMS performance. We present a classification of FMS models, covering justification, design and operational control problems.

The structure of the paper is as follows. In the next section, we highlight some background issues of flexible manufacturing. In particular, different types of flexibility are characterized. We dwell on the advantages and disadvantages of Job Shops and dedicated production lines and show how the development of the FMS-concept arises naturally.

Section 3 provides a reference framework for those readers not familiar with the technical background of an automated manufac-
turing system. A number of basic concepts and elements of Flexible Manufacturing Systems are briefly described.

In section 4, we present a classification of FMS problems and their mutual relationships. This classification is neither entirely new nor unique; earlier work has been published by e.g. Kusiak[1986], Stecke[1985], Van Looveren et. al.[1986] and Kalkunte et. al.[1986]. Future developments on the path towards the "Factory of the Future" are nicely described in Bullinger et. al.[1986] and Meredith[1987]. With respect to each problem area, a number of important contributions in the literature is indicated. This literature review is not exhaustive; despite of the fact that the area is relatively new, the number of papers, conferences and journals devoted to FMS's is growing rapidly.

In section 5, finally, some conclusions are formulated and directions for future research are indicated.

In this paper, we deal almost exclusively with flexible machining systems, not with flexible assembly systems. The nature of flexible assembly systems demands for a different approach in some important aspects, both in design and with respect to operational control. In Zijm[1987], flexible assembly systems are discussed separately for this reason.

A number of models, mentioned only briefly now but discussed more thoroughly in Zijm[1987], differ in important aspects from previous ones published. With respect to justification, it should be said that we do not believe in one Operations Research model to represent all possible trade-offs; however, in highlighting the benefits of a FMS, models can play a significant role.

Finally, we remark that the classification presented here reflects our idea to decompose problems as much as possible, in order to provide insights and to facilitate quick and, if necessary, repeated solution of these problems. All these issues are discussed more extensively in Zijm[1987], which forms the basis of this paper.
2. Flexibility and manufacturing

Flexibility in manufacturing means: a quick response to the market. To develop, manufacture and distribute, with short leadtimes, products of high quality, if possible on an order based planning rather than on a forecast based planning.

Hence, flexibility means more that just installing automated production systems; it has farreaching consequences also for product development, as well as for materials management and physical distribution management. In general, we distinguish between three different, but interrelated types of flexibility.

- **Volume flexibility** is related to the possibility to react to volume fluctuations in sales. The basic idea is that leadtimes should be so short that production will be able to follow these fluctuations. This in turn requires flexibility in capacity but the alternative is to keep large stocks of final products, a situation which is generally recognized to be highly undesirable by most industrial companies. An appropriate logistics organisation is a key factor in attaining volume flexibility.

- **Mix flexibility** refers to the rate at which changes with respect to the specifications of a product can be reflected on the shop floor. Throughputtimes and set-up times in the manufacturing process are the key parameters here.

- **Innovation flexibility** finally deals with the ability to get new products, or new versions of a product, on the market in time. Product life cycles become shorter and shorter; we are forced to reduce development leadtimes drastically. Also, when products are designed, care should be taken to make them "manufacturable". In fact, product and process development should go hand in hand.

Although all three types of flexibility have their impact on the design of production systems, Flexible Manufacturing Systems particularly seem to be the key element towards achieving mix flexibility.
A Flexible Manufacturing System is a production system capable of manufacturing a large variety of different parts or products through each other, while maintaining high overall production rates and short throughput times. Often, these systems are characterized by a high degree of automation, such as an automated Material Handling System (MHS), Computer Numerically Controlled (CNC) machines, while also the complete shopfloor process is under rigid computer control. The absence of large set-up times is a key feature of these systems. When operational, a FMS can be (re)planned with a high frequency (e.g. every four hours).

Other views on flexibility in manufacturing are described by various authors, see e.g. Buzacott[1982] and Slack[1983].

It is interesting to compare a FMS with more traditional manufacturing systems such as the classical Job Shop or, on the other hand, the dedicated production line, as found in traditional automobile factories. In the latter ones, a high productivity and short throughput times were achieved indeed, but at the cost of an extreme inflexibility with respect to variations in the output. On the other hand, the traditional Job Shop made it possible to supply a fairly varied range of products, but, due to large changeover times, at the cost of high intermediate stocks and extremely long throughput times.

Flexible Manufacturing Systems are ultimately an attempt to combine the advantages of both the Job Shop and the dedicated production line, i.e. a high degree of diversity and, at the same time, a high level of efficiency with short throughput times, and to avoid the disadvantages. In passing, we mention another advantage of short throughput times, with respect to quality: faults are detected earlier and information about these can therefore be fed back more quickly. High intermediate stocks, on the other hand, only help to conceal faults (Schonberger[1982]).

In the next section, we describe some key elements of a metal cutting FMS, to provide the reader with some basic understanding of the (computer controlled) hardware components, that are the key towards more flexibility.
3. Elements of Flexible Manufacturing Systems

This section is devoted to a description of some of the basic hardware components that can be found in a metal cutting Flexible Manufacturing System. Examples of these kind of systems are described in several papers (e.g. Stecke and Solberg[1981]). Examples of Flexible Assembly Systems can be found in Wittrock[1985] and Zijm[1987].

Let us start with the heart of a metal cutting FMS, a CNC-machine. Fig. 3.1 shows an integrated milling and drilling machine. The main difference, compared with conventional machinery, is the fact that this CNC-machine is capable to perform a, usually large, set of different operations, without human intervention. Each operation is described by a NC-program, specifying among others which tools (and in what sequence) are needed to perform that operation. On completion, the part is automatically unloaded and replaced by another. After identification, a new program is loaded and the next operation may start. It is emphasized that one operation may require the subsequent use of several different tools.

Coupled with the machine is a tool magazine, holding a large

Fig. 3.1. An integrated milling and drilling CNC-machine.
A variety of different milling cutters and drills. An automatic tool changer interchanges tools between the drive mechanism and the tool magazine when needed, in only a few seconds. The tool magazine is loaded with all tools, needed to perform operations in a fixed period of four hours, say. It is this tool magazine, together with the automatic tool changing mechanism, which gives the machine its inherent flexibility.

Modern machines are usually provided with an inspection mechanism, to check the state of all tools on operation. It is not uncommon that a tool is worn off after an exploitation of totally 30 minutes (using it intermittently).

Associated with the machine (or with a group of similar machines) is a (usually small) set of specially designed pallets that can hold one or more parts or workpieces. The correct positioning of the parts on the pallets is achieved by means of special fixtures or clamping devices. The pallet with the workpiece is placed and fixtured on a worktable which can be rotated on two or three axis.

The machine may be equipped with a number of special features such as (vision) systems for identifying the parts and checking their correct positions, while also the performance of each.

Fig. 3.2. A Flexible Manufacturing System with AGV transport.
operation is supervised continuously.

Fig. 3.2 shows a Flexible Manufacturing System, consisting of a number of the above described CNC-machines. The machines are basically similar, however, due to the fact that each machine may be loaded with a different set of tools, the set of operations assigned to each machine may be different (the assignment of operations to machines is one of our major control problems, compare the next section). The machines are in general connected by an automated Material Handling System (MHS), in our example an Automated Guided Vehicle System (AGVS). Other Material Handling Systems may be a railcart system, an automatic conveyor system or, when machines are physically placed together in a workcell, a handling robot. If buffer space is provided, it is usually integrated with the MHS. Pallets, fixtures, grippers and tools are shared by the machines.

The machines, including the tool magazines, the tool changers and the pallet positioning mechanisms, are usually controlled by local micro-processors. The interaction between different machines in the system and the operation of the MHS is supervised by a central computer. The latter one also communicates with the computers of the production planning department.

Stecke and Solberg[1981] describe a Flexible Manufacturing System at Caterpillar, consisting of three drilling machines, four milling machines, two vertical turret lathes and an inspection machine, connected by a railcart system. Other examples of FMS's in metal cutting are given by e.g. Groover and Zimmers[1984], Ranky[1983] and by Kearney and Tracker[1980], a large manufacturer of automated production systems. These systems are typically meant to produce an almost unlimited variety of parts in very small batches, on order. Special designs can be implemented by specifying a new set of programs. Typical products of such a system are e.g. gear boxes, compressor houses, etc.

In the next section, we will define a classification of FMS problem areas and discuss a number of relevant models, suggested in the literature. In order to get a clear picture of these problems and their mutual relationships, the background material presented above may serve as a reference to the reader.
4. A classification of FMS problems

The hierarchical classification, we will present in this paper, distinguishes between justification problems, design problems and operational problems. Each of these three groups can be subdivided again. Another important area is what we shall call interface problems.

Justification of an investment in a FMS should be based on clearly defined objectives such as a desired market share, reduction of stocks and leadtimes and the like, in particular when replacing traditional machinery. Quantitative models should yield estimations of the benefits, and of the reductions in the various types of production related costs, and balance these against implementation costs in order to select an acceptable alternative.

The design of a FMS is related to such questions as: central or local buffers, the operation of the MHS but also the parts spectrum to be produced on the system. Design conflicts should ultimately been solved by studying their impact on performance criteria, in view of the desired objectives.

Operational problems primarily deal with detailed shopfloor planning and scheduling decisions, on a daily basis. It is the responsibility of the shopfloor controller to operate the system in an optimal way, within the constraints imposed by the design and the layout of the system.

Interface problems finally consider the interaction with the environment of the FMS. Flexibility implies that one is able to produce a large variety of products in a short time. Material availability and modular product structures are then crucial factors. Another important factor relates to the adaptability of the organisation, in particular in the field of production planning and distribution, to an increased flexibility in the factories.

In the next subsections, justification problems are briefly discussed, after which we pay more detailed attention to design and operational problems. Interface problems are not discussed at all, for an elaborate treatment of this subject we refer to Zijm[1987].
4.1. **Justification problems**

A frequently raised argument to invest in flexible technology is the loss of market share in the long run if one doesn't. This idea is usually based on the observation that diversity in demand has increased, and will further increase, rapidly, leading to either an explosion of the set-up's (and hence a loss of capacity) in traditional machining centres, or long leadtimes and high work-in-process levels and, consequently, high final inventory levels, all in line with the usual economies-of-scale considerations. At the same time, decreasing product life cycles make it risky to keep large amounts of final products in stocks, apart from the fact that high interest rates make stock reductions inescapable. Therefore, we should focus on economies of scope, rather than on economies of scale (Goldhar and Jelinek[1983]).

Despite all this, the range of OR-models that address in particular an investment analysis of FMS's, in view of these operational cost characteristics, is very limited. One reason may be the difficulty in determining future effective demand, by the interaction between manufacturing technology and market characteristics (Burstein and Talbi[1985]). Also, sound measurements of market competitiveness, by means of for example portfolio analysis methods, are usually based on doubtful assumptions, if developed at all.

Whether or not classical capital budgeting techniques should be used to evaluate investment decisions in flexible technology, is the subject of an ongoing debate in the literature. Burstein and Talbi[1985], Michael and Millen[1985] and Primrose and Leonard[1986] express the opinion that these traditional approaches fail to capture adequately the essential features (and, in particular, the benefits) of a FMS; some alternative approaches are suggested. On the other side of the spectrum we find for instance Kulatilaka[1985] who exploits a risk-adjusted discounted cash flow model in calculating net present values. Strategic issues with respect to FMS-investments are furthermore discussed by Gaimon[1985],[1986] and Srinivasan and Millen[1986]. An interesting scoring model for flexible manufacturing systems project selection is presented by Nelson[1986]. Fine and
Freund [1986] develop a convex quadratic programming method to decide upon a portfolio of flexible and nonflexible capacity. Papers that address the reduction of specific operational costs are those of Porteus [1985], [1986] and of Van Beek and Van Putten [1987]. These authors do not try to develop an overall investment analysis method but merely try to quantify integral effects of, for instance, reductions in set-up costs or overall leadtimes. Unfortunately, they all consider the single product situation whereas the benefits of a FMS particularly arise from the possibility to produce a large variety of different items with a negligible loss because of set-up times, hence in small batches and with short throughput times. Zijm [1987] advocates the use of a slightly extended version of a model, developed originally by Silver [1975] (see also Silver and Peterson [1985]) to quantify the effects of reductions in changeover costs in a Group Technology environment (cf. Burbidge [1975]). In particular, a distinction is made between changeover costs from one part family to another one and changeover costs within one family. Modern Group Technology approaches, combined with sophisticated machine loading techniques, enable planners to define large part families, where the inter-family changeover costs drop to zero. The advantages of these achievements are demonstrated in Zijm [1987].

4.2. Design problems

Once a decision has been made with respect to a possible investment in a Flexible Manufacturing System, the question arises what is the best system layout and what basic operation rules should be implemented in order to exploit the installed equipment in an optimal way. For example, choices have to be made about sizes and locations of buffers, release rules for dispatching products to the system should be developed, etc. Performance indices may be a desired throughput and/or throughput times, along with the desired work-in-process levels. Note that we are still in a situation where actual production needs are not specified; at the best, we have a global indication about the long term production ratio's for families of items. A careful
long-term part family selection, based on similar operating characteristics as well as on similar geometrical and technological attributes (Group Technology) should be included in the design phase.

With respect to system layout and system performance, queueing network models have provided valuable insights into a number of design issues. In particular, closed queueing networks or restricted open queueing networks yield an adequate description of a FMS, in principle at least. The ratio behind this is that in most FMS's the number of parts in the system at any moment is limited, due to the limited number of available pallets or because of a strict workload control.

Recently, we have seen a tremendous increase in papers dealing with queueing network models applied to a FMS-environment. Let us describe the basics of the approach. A FMS is described as a network of single or multi-server stations where each multi-server station denotes a workcell of identical machines. The material handling system (MHS) is usually modelled as a pure delay, i.e. as an infinite server (e.g. in the case of an automatic conveyor belt) but sometimes also as a multi-server station (e.g. if only a limited number of AGV's is available). A basic result in queueing network theory states that under certain assumptions the steady state probabilities of finding \( n_i \) jobs at workcell \( i \) have a product form, i.e. they separate into a normalization constant and factors that depend exclusively on the characteristics of one workcell (Jackson[1963], Gordon and Newell[1967], Baskett et. al.[1975]). In workcells with a FCFS (First Come First Serve) discipline, exponentially distributed service times have to be assumed in order to make the model exact, while furthermore buffer capacities are unlimited in principle. However, several approximations have been developed for more complicated networks (e.g. Whitt[1983a], [1983b], Yao and Buzacott[1985b], [1986a]).

Closed queueing networks depict the situation where each finished job is immediately replaced by a new one. This situation can be modelled explicitly by defining a special single server load/unload station. The throughput-time, calculated by Little's formula, is then the time between two successive visits of a job
to the load/unload station in the closed queueing network. Special algorithms have been developed to calculate these different performance measures in an effective way, the two most well-known being the convolution method (Reiser and Kobayashi[1975], Reiser[1977]) and the Mean Value Analysis (MVA) algorithm (Reiser and Lavenberg[1980], Reiser[1981]).

Solberg[1977], [1981] was the first author who applied closed queueing network techniques to model and analyze Flexible Manufacturing Systems. Approximate MVA algorithms for FMS's are described in e.g. Suri and Hildebrant[1984] and Shalev-Oren, Seidman and Schweitzer[1985].

The infinite buffer assumption appeared to be a serious obstacle to model realistic FMS's (with usually small buffers). However, in one special case, particularly relevant to FMS's, a nice solution exists. A central server model (CSM) describes the situation where, between any two stations, access is needed to a central server. In most FMS's, the MHS does play the role of such a central server. Under the assumption of exponential service time distributions, these models appear to be time reversible (Keilson[1979], Kelly[1979], even when buffers are assumed to be finite. Note that, due to the limited buffer assumption, blocking may occur when a job attempts entry from the MHS to a work station. The capacity of the MHS on the other hand is assumed large enough to accomodate all jobs in the system. These models have been analyzed by Yao and Buzacott[1985 a], [1986 b]; even in certain state-dependent routing cases, product form solutions could be obtained.

In more general limited capacity models (without the central server assumption), blocking phenomena prevent an exact analysis of problems of realistic sizes. Approximation methods have been developed by Yao and Buzacott[1985 b], Altiok and Perros[1986], Dallery and Yao[1986], Van Dijk[1987] and, using a fluid model approach, by De Koster[1986] and Wessels et. al.[1986]. An interesting alternative method, based on a perturbation analysis of one long simulation experiment, is presented by Ho et. al.[1979]. A detailed analysis of a buffer control model with priorities in the routing of the jobs is provided by Repkes and Zijm[1988].
Furthermore, we advocate the use of aggregation methods (see e.g. Chandy et. al.[1975]), for the situation where some part of a larger network behaves like a closed queueing network (because of a workload dependent admittance rule), and of renewal approximations (Whitt[1983 a], [1983 b]) or an exponentialization method (Yao and Buzacott[1986]) for the case of general service time distributions. Buzacott and Yao[1986] provide a detailed overview of queueing network approaches to model FMS’s.

The long term part family selection problem is usually solved by clustering techniques, based on operating characteristics as well as on technological considerations. Since basically the same approach is used for short term FMS planning, in particular when grouping operations for a certain period, we postpone its discussion to section 4.3.

4.3. Operational problems

Once the detailed design and, subsequently, the installation of the FMS have been completed, one should develop a framework for the solution of a number of frequently returning planning and scheduling problems. In this section, we define four, hierarchically coupled, levels of decision making in an operational environment (compare also Kusiak[1985]).

- Planning

Prepare a list of production orders, based on internal or external demand, for instance on a weekly basis.

- Operations Grouping

Each job involves a sequence of operations, each one described by a NC-program. Different operations may require different fixtures or even pallets. Divide the first planning period into a number of time windows. For each time window, varying for example from four to eight hours, select a set of appropriate operations for processing in that window, taking into account eventual
precedence constraints, such that refixtureing within such a time window is not necessary.

- Tool Loading

Given the outcome of an Operations Grouping procedure for a time window, assign operations and (hence) tools to the tool magazines of the CNC-machines.

- Sequencing

Within one time window, develop dispatching rules for the set of operations, given the outcome of the Tool Loading procedure.

Below, we discuss some approaches to these problems.

4.3.1. Planning

Suppose a long term family structure has been determined for the set of all part types, based particularly on operating characteristics such as required tools, fixtures, etc. In a Flexible Manufacturing System this means that in particular set-up costs only play a role at the level of a complete family, i.e. when changing production from one family to another one. Within one family, set-up costs are not relevant any more, moreover, inventory holding costs do not depend on the particular item within a family structure.

Production planning in the case of a machine shop as the one, described in section 3, is usually based on an explosion from a Master Production Schedule (MPS) on end-item level. This is a very common approach, known as Materials Requirement Planning (MRP), see e.g. Orlicky[1975], Silver and Peterson[1985] or Vollman, Berry and Whybark[1984]. In our case, the explosion can be on a family level basis, due to the flexibility of the FMS, and details with respect to part types can be specified at the last moment. This reflects the concepts of Hierarchical Produs-
tion Planning (HPP), developed by Hax and Meal[1975] and Bitran, Haas and Hax[1981], [1982], see also Hax and Candea[1984], ch. 6.

The reader should recognize that, due to the MRP-type approach, the machine shop is confronted with an essentially deterministic "demand" pattern. However, limited capacity and economic considerations based on set-up and inventory holding costs force the planning department to smooth this pattern and to determine batch sizes and the like. To convert the production requirements to a detailed production plan, we propose the use of a capacitated lotsize models at a family level. These problems are known to be NP-hard (Florian et. al.[1980]). Heuristics have been proposed by a number of authors, e.g. Dixon and Silver[1981], Dogramaci et. al.[1981] and Lambrecht and Vanderveken[1979]. Numerical comparisons have been presented by Maes and Van Wassenhove[1986 a],[1986 b], a more extensive treatment can be found in Maes[1987].

4.3.2. Operations Grouping

Suppose that for the first week, we have an adequate production plan on a family level basis. In general, the total number of tools needed to execute such a plan, exceeds the total capacity of the tool magazine. Furthermore, each job can in general be split up into a number of operations, where each operation needs special fixtures and sometimes even pallets; such an operation is in general completely specified by a NC program. Recall that one operation should not be identified with one single tool. In general, one operation requires a set of appropriate tools, the cardinality of this set however never exceeds the capacity of one tool magazine. Hence, one operation can (and will) be completely performed by one machine.

Due to the fact that subsequent operations of the same job require in general different fixtures or clamping devices and sometimes different pallets, it seems reasonable to split up the planning period into a number of time windows such that in each time window only operations are planned that do not require some intermediate refixturing. These time windows may vary, dependent on the nature of the underlying machining process, from four to
eight hours (for the unmanned night shift), unless automatic refixturing robots are available.

It follows that for each time window, we have to specify a set of operations for immediate and simultaneous processing, taking into account precedence constraints and the total capacity of the tool magazines. Hence, reloading of tools in the magazines, coupled with the machines, only takes place at the end of a time window. In order to minimize the number of time windows (or to maximize the average size of the windows), it is appropriate to select "schedulable" operations such that the corresponding sets of required tools have maximal intersections, where the union to these sets should not exceed the total available capacity. This approach leads in a natural way to a clustering formulation.

King[1980] and King and Nakornchai[1986] have developed a Rank Order Clustering (ROC) algorithm in order to solve these types of problems in a more traditional Group Technology (GT) environment (for an introduction into GT concepts, see Burbidge[1975]). Although certainly not the most efficient one, the matrix formulation proposed by these authors gives a nice insight into the way tool requirements influence the operations grouping. Extended versions of these type of algorithms (MODROC) have been presented by Chandrasekharan and Rajagopalan[1986].

Kusiak[1987] and Kumar, Kusiak and Vannelli[1986] present several algorithms for clustering parts according to the tool types needed. These methods are based on similarity measures between different operations (the p-median method) or on graph-theoretical approaches (in particular by trying to find so-called k-decompositions).

From the resulting set of clusters (where each cluster corresponds to a set of operations and a set of tools), we select as many as total tool magazine capacity allows. If the intersections of different clusters with respect to tool requirements is non-void, double or triple loading of similar tool types may be needed. Another reason for multiple loading of tool types may be to provide extra flexibility in the ultimate scheduling problem.
4.3.3. Tool loading

Once a set of operations, and a corresponding set of tools, is specified by the Operations Grouping procedure for execution in the next time window, the question arises how to assign these tools to the tool magazines coupled with the CNC machines. In particular, one may wonder what objectives should be pursued when assigning these tools.

Stecke[1983] presents a number of integer programming formulations and develops heuristics to solve these tool loading problems. Central in her approach is the concept of pooling. A set of machines is said to be pooled, if they can perform exactly the same set of operations. Note that this does not exclude the situation where an operation is assigned to more than one pooled group of machines. However, in general one can say it is better to have a minimum of groups of pooled machines (compare also Stecke and Morin[1985], Stecke and Solberg[1985]).

What kind of criterion function should be taken depends on the problem situation. Several criteria have been investigated in Stecke[1983]. For example, one may wish to balance the machine workloads, or to minimize the number of movements from machine to machine, or to fill the tool magazines as densely as possible.

Quite a different formulation is given in Zijm[1987], based on work of Zeestraten[1987]. The objective is here to maintain a maximum of flexibility in the underlying sequencing problem. The idea is that, for each machine, one should:
- maximize the amount of time, needed for operations that may be performed by that machine, and
- minimize the amount of time, needed for operations that have to be performed by that machine exclusively.

Another formulation has been presented by Rajagopalan[1986].

4.3.4. Scheduling

As a result of the tool loading procedure, we know exactly which operations can be performed on which machines. The final task left is now to schedule the operations such that for example the total makespan is minimized. The resulting scheduling problem
can be formulated as a mixed integer program but is in general extremely difficult to solve exactly. Also, the fact that there may be a number of stochastic interrupts, due to for instance breakdown of tools or other unexpected failures, leads to the use of simple priority rules such as SPT (Shortest Processing Time First) or SPT/TOT (compare e.g. Stecke and Solberg[1984]). However, recent results in job shop scheduling theory justify the exploitation of more advanced algorithms in our FMS scheduling problem.

To illustrate the difficulties of our scheduling problems, the reader should recall that an operation can be performed on more than one machine in general. This introduces parallelism but the parallel machine structure is highly operation dependent. For example, operation A may be performed on either machine 1 or machine 2, operation B on either machine 2 or 3, operation C on either machine 1 or 3, etc. To further complicate matters, we have precedence constraints between operations that actually belong to the same job and are performed in the same time window (in the case that refixturing is not necessary). The resulting problem is a strong generalization of the classical job shop problem.

In the case where machines are pooled together (compare section 4.3.3.) and no operation is assigned to more than one group of pooled machines, we still have a more natural, generalization of the job shop problem. Only when each pool contains only one machine, the classical job shop formulation arises.

If, on the other hand, pools contain more machines, but no consecutive operations of one job are scheduled in the same time window (hence no precedence constraints exist), we are left with a number of parallel machine problems for which good heuristics are developed, e.g. the LPT (Longest Processing Time First) rule or the MULTI FIT method (cf. Coffman, Garey and Johnson[1978]).

Promising new methods for the job shop scheduling problem have been presented by Adams, Balas and Zawack[1986]. They use an iterative procedure with Carlier's algorithm for the one machine scheduling problem with release and due dates as a building block (compare Carlier[1982]).
The literature related to scheduling problems in FMS's is rapidly growing although most problem formulations are highly simplified versions of the one discussed here. We mention in particular Morton and Smunt[1986] who present a more general framework for planning and scheduling and Shaw and Whinston[1986] and Subramanyam and Askin[1986] who exploit results from Artificial Intelligence. Finally, although meant for Flexible Assembly Systems (which differ in many important aspects from Flexible Manufacturing Systems, see also Zijm[1987]), we mention the work of Wittrock[1985], who develops a periodic scheduling algorithm for a printed circuit board assembly line, actually a heuristic procedure to minimize cycle time. Pinedo et. al.[1986] also discuss a periodic scheduling algorithm for a flexible assembly line, but include a finite buffer constraint.

This concludes our subsection on operational problems and thereby also the complete hierarchical classification. As the reader may have noticed, we believe that stochastic network methods are particularly suitable for design problems, whereas operational problems lend themselves more to a mathematical programming formulation. We will come back to this point in the final section.
5. Conclusions and directions for future research

There is certainly a lack of Operations Research models that address in particular the justification of investments in flexible technology. As mentioned before, most of the papers use capital budgeting techniques such as Discounted Cash Flow (DCF) analysis (see e.g. Lutz[1982]). However, most of the supposed benefits of a FMS are treated as given parameters in these models. We believe that exactly this quantification of possible benefits is a major problem, for which more research is needed. Also, the cost of installing a FMS are in general not easy to estimate, being not only dependent on hard- and software, but also on for example education and training. Total installation costs are therefore in general not linear in the required capacity. Some kind of break-even analysis (see e.g. Thuesen[1982]) may be necessary to make a total cost-benefit analysis.

As mentioned before, we believe that queueing network techniques are suitable to analyze in particular design problems whereas operational problems can be better formulated by using mathematical programming models. Solutions to design problems have to be robust with respect to a large and varying set of operational conditions, which cannot even be specified at this design phase. This leads in a natural way to a stochastic formulation. On the other hand, the short term planning environment, although highly diverse and complex, is in general deterministic in nature.

It is in general not very realistic to model a Flexible Manufacturing System as a product form network, due to finite buffer constraints, state-dependent routing and job priorities, together with the fact that "aggregate" service times are in general not exponential. More research on approximation methods which capture effectively these and other limitations, inherently present in FMS's, is needed. Also, the need for good traffic control rules for the MHS justifies further research.

With respect to operational problems, we feel that the impact of different grouping strategies on the loading problem, and of different loading strategies on the final scheduling problem, is not well understood. Furthermore, as sketched in the preceding
section, the loading and scheduling problems itselfs are extremely difficult; there is still a lot of room for better heuristics.

Several problems we did not mention at all. Machine control problems belong to the field of CAM (Computer Aided Manufacturing). For example, if operations are very short, the sequence in which the tools are loaded in the tool magazines becomes important. This leads to a formulation, known as the Travelling Salesman Problem. On the other side of the spectrum we have what we called FMS interface problems. Material procurement becomes a very critical factor when exploiting a Flexible Manufacturing System. Also, due to the flexibility of these production systems, production planning can be much more of a hierarchical nature, where details are filled in literally at the last moment. The ultimate success of a transfer to flexible technology is highly dependent on a good appreciation of these factors. The need for Flexible Management Systems to fully exploit the benefits of Flexible Manufacturing Systems is recognized by now, but research on these interface problems has hardly begun.
References


Gaimon, C.[1986], The continuous acquisition of automation to reduce production and in-process inventory costs, in Kusiak[1986 a], pp. 29-43.


Hax, A.C. and D. Candea[1984], Production and Inventory Management, Prentice-Hall, Inc., Englewood Cliffs, NJ,
Hax, A.C. and H.C. Meal[1975], Hierarchical Integration of Production Planning and Scheduling, in M. Geisler (ed.) [1975], TIMS Studies in Management Science, Vol. 1, Logistics, North Holland/American Elsevier, New York,
Ho, Y.C., M.A. Eyler and T.T. Chien, A gradient technique for general buffer storage design in a serial production line, International Journal of Production Research 17 (6), pp.557-580,
Jackson, J.R.[1963], Jobshop-like Queueing systems, Management Science 10 (1), pp. 131-142,
Kearney and Tracker Corporation[1980], Understanding Manufacturing systems, Vol. 1, Milwaukee, WI,
Keilson, J.[1979], Markov Chain models: Rarity and Exponentiality, Springer Verlag, New York,
Kelly, F.P.[1979], Reversibility and Stochastic Networks, Wiley, New York,
King, J.R. and V. Nakornchai[1986], An interactive data-clustering algorithm, in A. Kusiak[1986], pp. 285-300,
Koster, M.B.M. de[1986], Approximate Analysis of Production Systems, Report BDK/ORS/86/09, Department of Industrial Engineering, Eindhoven University of Technology, The Netherlands,


Porteus, E.L.[1985], Investing in reduced set-up in the EOQ model, Management Science 31 (8), pp.998-1010.


Schonberger, R.J.[1982], Japanese Manufacturing Techniques, The Free Press (MacMillan), New York,
Silver, E.A.[1975], Modifying the Economic Order Quantity (EOQ) to Handle Coordinated Replenishments of two or more items, Production and Inventory Management, Vol. 16 (3), pp. 26-38.


Van Beek, P. and C. van Putten[1987], OR contributions to flexibility improvement in production/inventory systems, European Journal of Operational Research 31, pp. 52-60.


Wessels, J., J. Hontelez and W.H.M. Zijm[1986], Eine Produktionslinie mit unzuverlässigen Komponenten (in German), presented at the 15th DGOR-meeting, Ulm, FRG.


