Politecnico di Milano
Scuola di Ingegneria Industriale e dell'Informazione
Corso di Laurea Magistrale in Ingegneria Meccanica

A proactive approach to a lot-sizing and parallel machine scheduling problem under demand uncertainty

Relatore: Prof. Giovanni MIRAGLIOTTA
Correlatore: Prof. Suleyman KARABUK

Tesi di Laurea Magistrale di:
Paolo MINNOZZI Matr. 804356
Gianluca TEDALDI Matr. 799966

Anno Accademico 2013-2014

"Lavorare, studiare, è fare un uomo prima che fare un'opera."

E. Mounier

Contents

Abstract

The aim of this Thesis is to provide a proactive approach to a lot-sizing and scheduling problem under uncertainty of demand. High demand variability causes difficulties in developing efficient production schedules. Many authors studied this problem using deterministic approaches, meaning that they consider demand as "known"; the main problem related to these models is the so-called system "nervousness" in terms of machines configuration, which determines high scheduling costs. The actual demand, sometimes very different from the forecasts, induces planners to modify their schedules, and, as a consequence, a large number of machine changeovers occur. In this work we develop a deterministic model for the scheduling problem we are going to consider, and we refer to this as a benchmark. Our aim is to suggest a stochastic model in which uncertainty is taken into account in the model, and prove that this approach can be a solution to the problem. We develop a stochastic model based on a scenario tree representation; the very large number of variables involved in the model makes the problem too hard to solve in a complex environment. This forces us to develop an approximate stochastic model. We consider a textile manufacturing scheduling problem, that is one of the most difficult fields in which create production plans due to uncertainty of demand. Good results are obtained from the computational experiments, in terms of the purpose of this work.

Keywords

Lot-sizing and Scheduling - Parallel Machine - Demand Uncertainty - Rolling Horizon - Multistage Stochastic Programming

Sommario

L'obiettivo della Tesi è fornire un approccio proattivo a un problema di dimensionamento dei lotti e scheduling in condizioni di incertezza della domanda di prodotto finito. L'elevata variabilità della domanda comporta difficoltà nello sviluppo dei piani di produzione efficienti ed efficaci. Molti autori hanno studiato il problema, affrontandolo mediante approcci deterministici, ovvero a domanda "nota"; il problema principale legato all'utilizzo di questi modelli è il cosiddetto "nervosismo" in termini di configurazione delle macchine, che determina elevati costi. La domanda reale infatti, talvolta molto differente dalle previsioni, induce i planners a modificare i loro piani.
In questo lavoro sviluppiamo un modello deterministico, e lo utilizzeremo come base per il confronto con la nostra ipotesi di soluzione. Il nostro scopo è quello di implementare un modello stocastico, in cui l'incertezza è considerata nel modello, e provare che questo approccio può essere una possibile soluzione al problema. Sviluppiamo un modello stocastico basato sulla formulazione matematica di un albero degli scenari; il fatto di voler descrivere l'incertezza in modo soddisfacente causa un numero troppo elevato di variabili coinvolte nel modello, che rende il problema troppo difficile da risolvere. Questo ci impone di sviluppare un modello stocastico approssimato. Consideriamo un problema di scheduling di un'impresa del settore tessile, caratterizzato da elevata incertezza della domanda. Dagli esperimenti numerici si ottengono buoni risultati in termini di obiettivi del lavoro.

Keywords

Lot-sizing and Scheduling - Parallel Machine - Demand Uncertainty - Rolling Horizon - Multistage Stochastic Programming

Estratto in lingua italiana

Questo lavoro affronta un problema di lot-sizing e scheduling multi-stadio di un sistema produttivo multi-prodotto a macchine parallele e con capacità limitata, in un contesto caratterizzato da elevata incertezza della domanda di prodotto finito. Il sistema produttivo in esame si dice “multi-prodotto” e a “macchine parallele” in quanto esso comprende un insieme di macchine generiche che possono elaborare indistintamente una serie di prodotti. Inoltre, nel sistema considerato ogni prodotto si trasforma in prodotto finito con la sola trasformazione ottenuta con la macchina generica.
Questo tipo di problema richiede una modellizzazione matematica atta a definire “cosa”, “quando” e “quanto” produrre all’interno di un orizzonte temporale discretizzato in un certo numero di periodi. In questi problemi si pianificano i setup delle macchine e si dimensionano i lotti per ogni periodo in modo da ottimizzare il costo totale di setup, costo di inventario ed eventualmente il costo di mancanza/stock-out nell’orizzonte di pianificazione.
Il comune approccio allo scheduling in condizioni di incertezza è di tipo deterministico. L’incertezza della domanda fa sì che esso sia spesso applicato con una logica di tipo Rolling Horizon.

Approccio deterministico e logica Rolling Horizon

Solitamente si sviluppano piani di produzione per un orizzonte temporale per il quale si ha a disposizione una previsione piuttosto affidabile della domanda, mediante un approccio deterministico, ovvero che prescinde dall’incertezza. Tuttavia, in contesti caratterizzati da forte incertezza della domanda, lo scostamento tra la domanda effettiva e quella prevista risulta tale da richiedere un intervento di mo-

difica dei piani di produzione, poiché essi sono stati ottimizzati sulla base della domanda prevista.
Con il manifestarsi dell'incertezza nel corso del tempo, l'osservazione della domanda reale ci permette di acquisire nuove informazioni in termini di inventario e commesse non evase; queste informazioni spingono le aziende a effettuare periodicamente dei nuovi piani di produzione per rispondere in modo reattivo al proprio contesto.
Questa logica, che suggerisce di rivedere i propri piani ottimizzati prima che questi vengano realizzati completamente, è stata studiata ampiamente in letteratura e ha preso il nome di logica RHP (Rolling Horizon Procedure).
Se ad un tipo di schedulazione con approccio deterministico viene applicata la logica rolling horizon in un contesto di elevata incertezza, il risultato è quello di realizzare piani di produzione altamente instabili. In altre parole, il numero di setup che vengono realizzati in ogni periodo è molto alto, e il sistema si trova ad essere affetto da una certa forma di nervosismo ([9], [10], [11], [12], [13]). In particolare, per quanto riguarda i problemi di lotsizing e scheduling, Tiacci propone un algoritmo semplificato per la schedulazione che sfrutta la logica rolling horizon con l'obiettivo di ridurre il numero di setup [7], con un approccio deterministico.

Incertezza della domanda

Questo nervosismo del sistema produttivo, ampiamente trattato in letteratura, è tanto più importante quanto più è grande l'incertezza della domanda. In ogni caso, le principali soluzioni a questo problema sono due: ridurre l'incertezza e utilizzare una domanda deterministica o considerare l'incertezza nei modelli.

- Nel primo caso, si potrebbero sviluppare determinati meccanismi di previsione atti a migliorare la qualità delle previsioni della domanda in modo da ridurre il gap con la domanda reale. In questo caso potrebbe essere giustificata l'applicazione di programmi di produzione sviluppati con un approccio deterministico.

- Nel secondo caso, invece, si potrebbe considerare di tenere conto dell'incertezza della domanda in sede di pianificazione/schedulazione, senza puntare eccessivamente sulla qualità delle previsioni (a cui molto spesso è associato anche un costo significativo). In altre parole, in questo caso si tiene conto dell'aleatorità della domanda all'interno del modello matematico che genera il piano di produzione (approccio stocastico).

Per quanto riguarda la qualità delle previsioni esistono numerosi metodi statistico-matematici che, mediante la considerazione di dati storici e/o altri fattori, permettono di produrre delle previsioni della domanda che si avvicinano molto alla domanda reale. Ad esempio, esistono modelli basati sulle serie storiche che hanno avuto molta fortuna: il modello di Brown (exponential smoothing) [47], il modello di Holt-Winters [48], il modello di Box-Jenkins [49], altri modelli di regressione di serie storiche [50] e ARIMA (AutoRegressive Integrated Moving Average).

Tuttavia, in alcuni campi di applicazione l'incertezza non può essere ridotta in modo soddisfacente e l'unica strada percorribile sembra quella di considerare la domanda come una variabile stocastica e incorporarla nei modelli.

Approcci stocastici

Un tipo di programmazione stocastica multi-stadio sembra, quindi, essere la risposta a questo problema. In letteratura si trovano diversi modelli di programmazione stocastica multi-stadio basati sulla teoria dell'albero degli scenari. Una formulazione matematica utile per seguire questa teoria è stata sviluppata da [16].
Tuttavia, se l'incertezza caratterizza la domanda di ogni singolo prodotto, si dovrebbero considerare un elevato numero di scenari per ogni nodo dell'albero. Infatti, per ogni prodotto e per ogni periodo, vengono definiti due o più possibili realizzazioni della domanda; per ogni nodo dell'albero degli scenari perciò devono essere presenti tutte le possibili combinazioni di incertezza dei singoli prodotti, come si vede nelle figure (1) e (2).

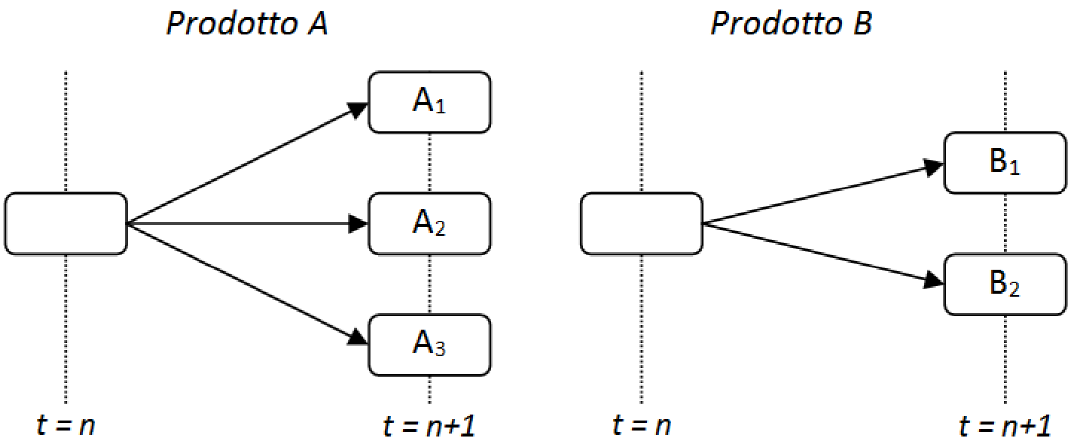


Figura 1: Scenari per nodo per ogni singolo prodotto

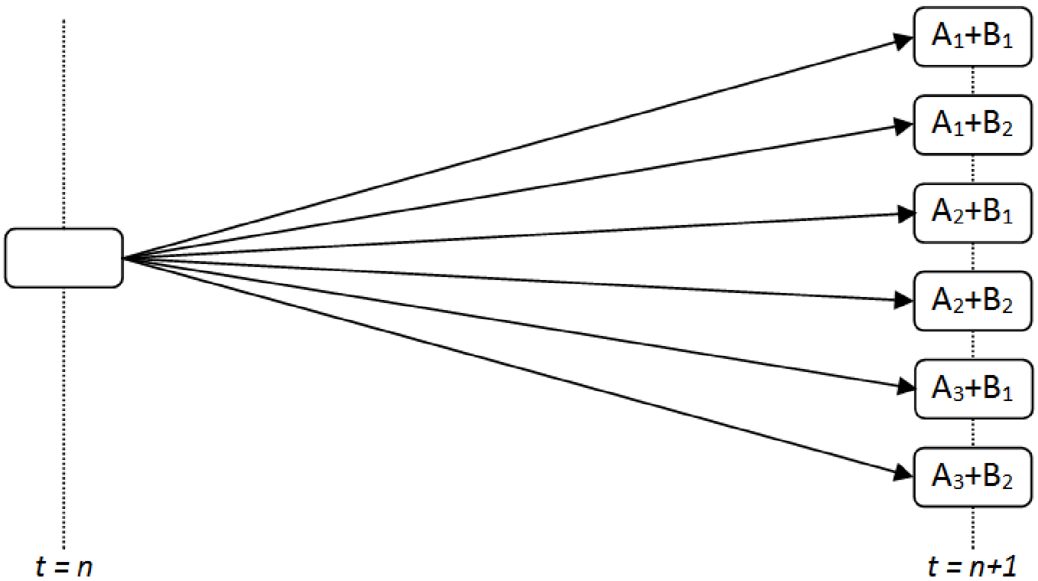


Figura 2: Scenari per nodo per l'albero degli scenari completo

Risolvere problemi di ottimizzazione con un albero degli scenari "completo" richiede un onere computazionale troppo elevato che ci permette di risolvere solo problemi molto semplici (pochi prodotti e pochi scenari per prodotto e per periodo).

Infatti il problema fondamentale della programmazione stocastica multi-stadio e multi-prodotto è diventato quello di contenere il numero di scenari che si considerano per ogni nodo, in modo da diminuire il numero di variabili che cresce esponenzialmente con il numero di realizzazioni per nodo. In figura (3) si vede un esempio di questo.

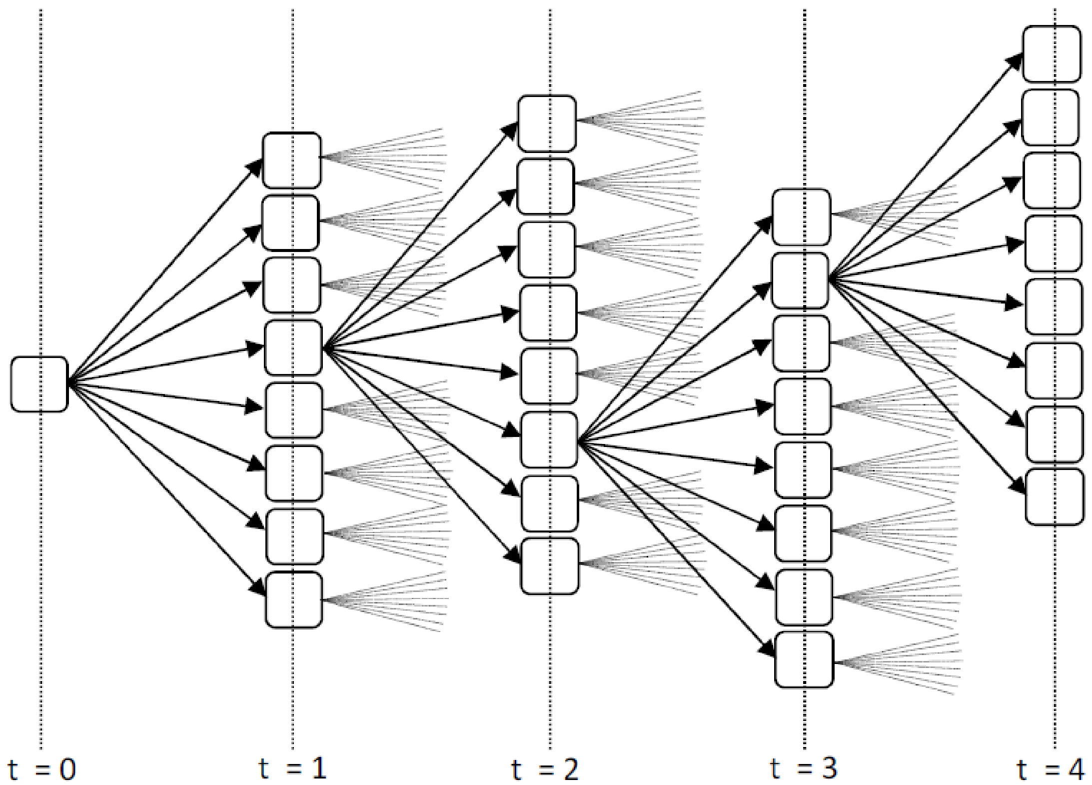


Figura 3: Esempio di albero degli scenari, con 8 scenari per nodo: 4096 scenari

A questo problema la ricerca ha provato a rispondere fondamentalmente in tre modi:

1. metodi di generazione di scenari: si creano scenari ad hoc mediante algoritmi complessi (tra gli altri [17] e [18]).

2. metodi di riduzione di scenari: partendo dal modello "completo" vengono elaborati algoritmi che riducono il numero di scenari da considerare. Rockafellar propone un metodo di aggregazione degli scenari [39], che viene applicato con buoni risultati nel caso di scheduling multi-stadio da Kensuke [6].

3. sviluppo di modelli approssimati che superano i modelli basati sulla teoria dell'albero degli scenari (questo non avviene per i primi due metodi, che invece mantengono la struttura ad albero). In quest'area di ricerca si trovano metodi, che, ad esempio, scompongono in sottoproblemi più semplici il problema principale posto dall'albero degli scenari "completo" [31].

L'approccio seguito in questo lavoro è quest'ultimo.

Nuovo modello stocastico approssimato

Nel modello basato sull'albero degli scenari le variabili in gioco sono tipicamente di tipo "Wait and See": ad ogni nodo dell'albero corrispondono variabili che dipendono dalla realizzazione dell'incertezza. In altre parole, al rivelarsi della domanda reale corrisponde un determinato stato del sistema (in termini di inventario, backlog) e determinate decisioni da prendere (variabili decisionali riguardanti la configurazione delle macchine).

La prima importante caratteristica che assume il nostro modello per diminuire il numero delle variabili in gioco è quella di considerare "Wait and See" solo le variabili riguardanti inventario e backorder, mentre le variabili decisionali diventano variabili di tipo "Here and Now" (non dipendono dalla realizzazione della domanda, quindi sono uguali per ogni scenario all'interno del medesimo periodo).

La seconda caratteristica, che costituisce la novità principale del nostro modello, è quella di considerare i valori attesi di Inventario e Backorder per ogni periodo e considerarli come "condizioni inziali" per il periodo successivo.

Per esempio:

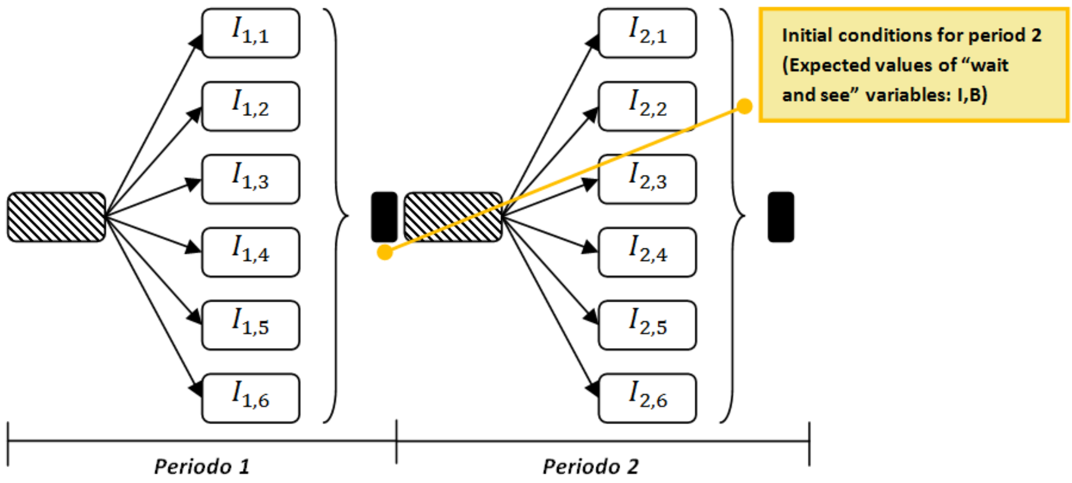


Figura 4: Valori attesi delle variabili Wait and See

$$E_1(I) = \sum_{s=1}^{6} I_{1,s} \cdot p_{1,s}$$

dove s è l'indice dello scenario e p è la probabilità associata a quel determinato scenario.
E' interessante notare che questi valori attesi di inventario e backorder non siano altro che funzioni lineari di variabili relative ai singoli scenari, pertanto non è necessario creare nuove variabili.

Grazie a questi due aspetti del nostro modello il numero di variabili in gioco diminuisce drasticamente (come si vede nella tabella), permettendoci di studiare problemi più complessi rispetto al modello stocastico tradizionale, ovvero considerando un maggior numero di possibili realizzazioni dell'incertezza per ogni nodo.

Scenari per nodo	NODI		Variabili (TOT)	
	STOC	STOC APP.	STOC	STOC APP.
8	4680	32	70200(WS)	192(WS) + 36(H) = 228
9	7380	36	110700(WS)	216(WS) + 36(H) = 252
10	11110	40	166650(WS)	240(WS) + 36(H) = 276
11	16104	44	241560(WS)	264(WS) + 36(H) = 300
12	22620	48	339300(WS)	288(WS) + 36(H) = 324

WS = "wait and see" variables

H = "here and now" variables

L'introduzione dei valori attesi delle realizzazioni costituisce una significativa approssimazione del modello "completo", tuttavia si scoprirà che il modello fornisce buoni risultati. Lo scopo della nostra ricerca non è quello di quantificare l'approssimazione del nostro modello rispetto a quello "completo", bensì quello di quantificare i benefici che si hanno rispetto ad un modello deterministico, che riserva non pochi problemi in ambienti caratterizzati da incertezza.

Questo modello di programmazione stocastica si propone come modello proattivo per i problemi di lot-sizing e scheduling. Tuttavia, il modello da noi presentato, viene proposto anch'esso in una logica Rolling Horizon per conservare la dimensione reattiva della schedulazione.

Nella trattazione viene elaborato anche un modello deterministico, che fungerà da benchmark per misurare i vantaggi del nostro modello.

Caso di studio

Si considera come caso di studio un settore manifatturiero caratterizzato da un'elevata incertezza della domanda: il settore tessile.
Anche in questo campo, proprio a causa della variabilità della domanda, utilizzando un approccio a domanda deterministica unito a una logica rolling, si verifica il problema della modificazione reattiva degli schedules. Nel caso di studio consi-

derato, dato che la componente principale dei costi di scheduling è costituita dai costi di setup, si dà maggiore importanza ai risultati in termini di riduzione del numero dei setup nell'orizzonte di pianificazione.
Come detto prima, le possibili strade sono due:

- miglioramento del sistema di previsione della domanda e utilizzo della domanda deterministica;
- considerazione dell'incertezza in un modello stocastico.

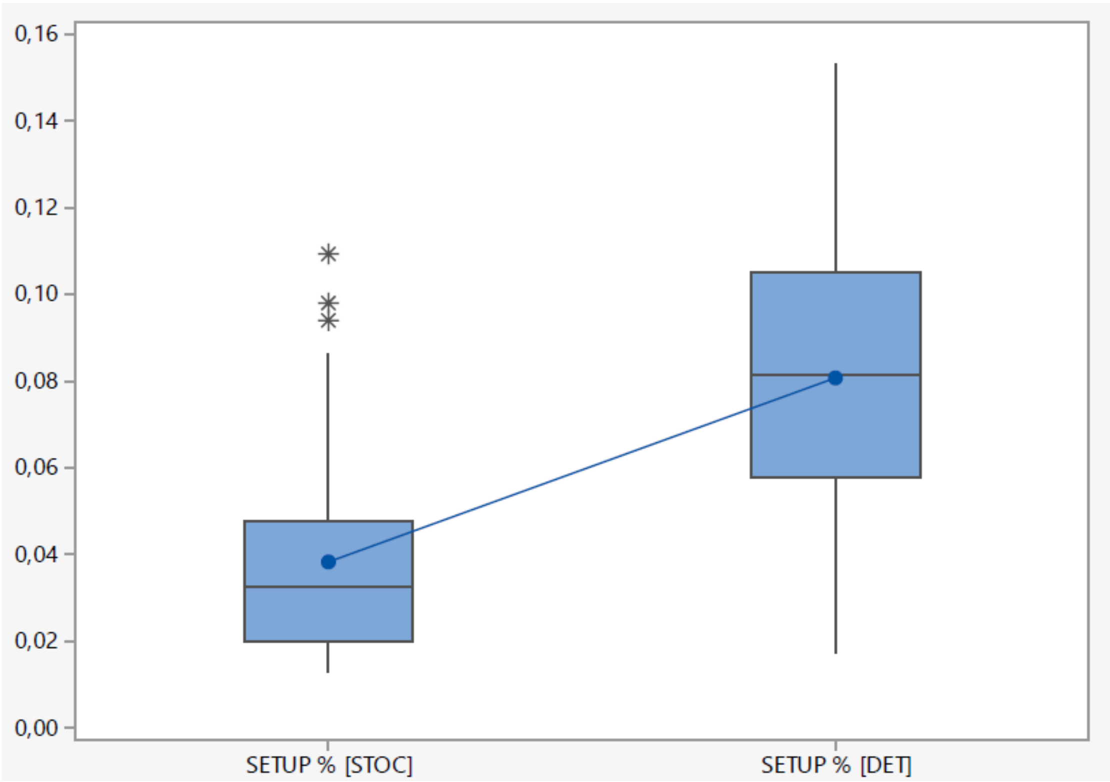
Molti metodi di previsione della domanda basati sulle serie storiche citati in precedenza non risultano efficaci in questo campo. Quindi, la maggior parte delle imprese utilizza tecniche avanzate per aumentare l'accuratezza delle previsioni di domanda, ma difficilmente si raggiungono risultati soddisfacenti.
Pertanto, come detto prima, si ritiene che l'applicazione del modello stocastico descritto precedentemente possa portare benefici.

Esperimenti e risultati

Per il caso di studio in esame vengono definite tutte le grandezze necessarie per identificare l'ambiente in cui realizzare gli esperimenti, tra cui la capacità dell'impianto, l'orizzonte temporale per lo scheduling, i prodotti e le relative caratteristiche (coefficiente di variazione della domanda, profili di domanda attesi, ...), lo shift factor (rapporto tra costi di setup e costi di inventario), ecc.
Vengono svolti quindi alcuni esperimenti per mostrare se il nuovo modello stocastico porta a dei benefici in termini di riduzione del numero di setup. Si risolvono entrambi i modelli con un certo orizzonte di pianificazione su un arco temporale di un anno, applicando la logica Rolling Horizon. La grandezza fondamentale in questo caso specifico, come detto in precedenza, è il numero di setup realizzati dai due modelli nell'anno.
Si ripete questo esperimento 100 volte considerando diverse realizzazioni della domanda, ma le stesse per entrambi i modelli. L'indice con cui si misura la prestazione in termini di riduzione di setup è SETUP [%]:

$$\text{SETUP [\%]} = \frac{\text{Numero totale di macchine con un attrezzaggio pianificato in un anno}}{\text{Numero di macchine} \cdot 52 \text{ settimane}}$$

Come si vede dal Box Plot, si può affermare che il numero di setup, a seguito dell'applicazione di un modello stocastico, si riduce di circa il 50%. Si esegue un t-test per evidenziare la differenza statistica tra il numero dei setup suggeriti dai due modelli.



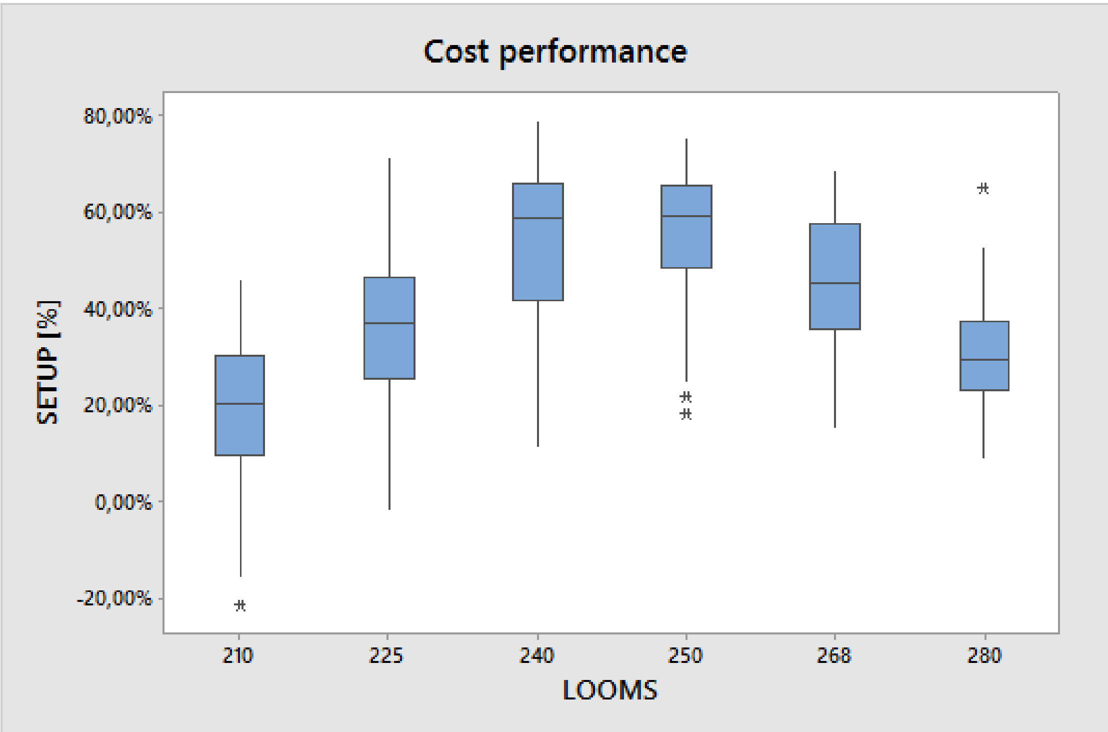
Nel corso degli esperimenti vengono studiati anche altri indici di prestazione riguardanti backorder, inventario e utilizzazione delle macchine. In particolare, si notano benefici anche rispetto al backorder e quindi al livello di servizio. Per quanto riguarda invece inventario e utilizzazione delle macchine non emergono significative differenze tra i due modelli.

Analisi di sensitività

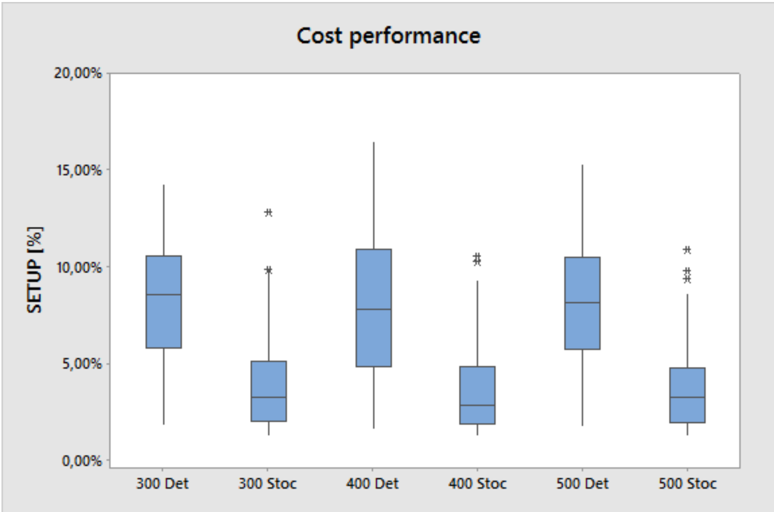
Si è voluto studiare la variazione del risultato visto in precedenza, al variare di alcuni parametri principali (capacità dell'impianto, shift factor e livello di incertezza della domanda), per studiarne la robustezza.

Variando il numero di telai si nota che il vantaggio ottenibile dall'applicazione del modello stocastico proposto presenta un picco per una certa capacità dell'impianto, mentre:

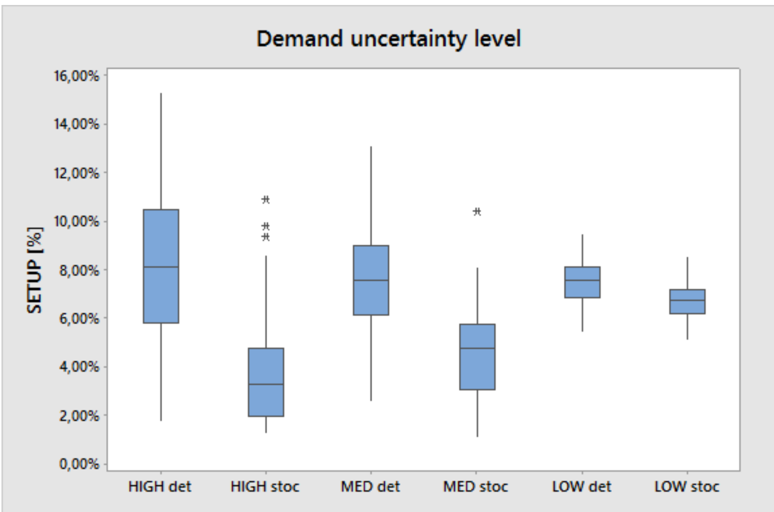
- nel caso di numero di macchine troppo basso rispetto alla domanda richiesta, il sistema di produzione lavora sotto pressione ed entrambi i modelli suggeriscono un elevato numero di setup. Quindi il vantaggio ottenibile con il nostro modello stocastico è meno rilevante;
- nel caso di numero di macchine molto elevato rispetto alla domanda richiesta, anche il modello deterministico suggerisce un modesto numero di setup. Per questo, un approccio proattivo alla gestione dei setup, sebbene efficace, non risulterebbe così significativo.



Variando lo shift factor la soluzione invece rimane piuttosto stabile. Vengono selezionati shift factor coerenti con il contesto di applicazione scelto.



Infine si può affermare che, all'aumentare dell'incertezza della domanda (e quindi dei coefficienti di variazione della domanda dei prodotti), l'applicazione del modello stocastico risulta più vantaggioso. Al contrario, in un contesto di domanda sempre meno incerta, il vantaggio si riduce progressivamente. In particolare, il risultato del modello deterministico rimane stabile al variare dell'incertezza, mentre quello dello stocastico perde gradualmente la propria efficienza.



Capitolo 1

Scheduling under uncertainty: the state of the art

The research problem we are going to study is about lot-sizing and production scheduling under demand uncertainty. A review of scheduling problems is provided by Graves [1]; according to his classification, we define the present problem as:

- **One-stage, parallel processors**. Each task requires a single processing step which may be performed on any of the parallel processors. It doesn't matter which machine a job is assigned to, but it cannot be processed on more than one machine at the same time. In particular the machines considered are identical processors, meaning that time required for the task is the same for all the machines.

- **Open shop**. An open shop is build to order, and the minimum possible inventory is stocked. In a closed shop the orders are filled from existing inventory.

- **Schedule cost**. The objective of the scheduling process is the minimization of the costs. In the following chapter we examine in depth the costs included in our instance.

Other important features of the problem are:

- Multi-item

- Multi-period
- Limited capacity
- Uncertainty of demand

Given an horizon consisting of a certain number of periods, the objective is to select a schedule that minimizes the total costs. We consider in the model setup costs and inventory carrying costs; we assume that setup costs are several times more important than inventory carrying costs (high shift factor), so that the main issue in this environment is the limitation of the number of changeovers.
Many authors studied this problem using deterministic approaches.
[2] and [3] are the main comprehensive survey papers on scheduling problems; the first paper reviewed the literature since the mid-1960s. Since the publication of that paper, there has been an increasing interest in scheduling problems with setup times (costs) with an average of more than 40 papers per year being added to the literature. The objective of the second paper was to provide an extensive review of the scheduling literature on models with setup times (costs) from then to date, covering more than 300 papers.

1.1 Deterministic models

Most of the work in this area has been limited to deterministic MILP models, wherein the problem parameters are assumed to be known with certainty. Examples of similar problem setting with respect to our work, in a single machine instance, can be seen in [4] and [5] (single machine) or in [6]. In this work we develop a deterministic Mixed Integer Linear Programming model (MILP) for the considered scheduling problem; this will be the starting point and the benchmark of the research aim.
The problem of the deterministic approaches, if a RHP (Rolling Horizon Procedure) is applied, is that they include the reactive modification of schedules upon realization of uncertainty. An application of the RHP method in scheduling problems can be found in [7].
With the reactive modification strategy there is the concrete possibility that too

many changes might be made to the schedule (this phenomenon is also present in other areas, such as MRP, Material Requirement Planning, and it is called system nervousness: see [9]).
Examples of this problem can be seen in [10], [11], [12] and [13] (all these papers refer to batch chemical plants scheduling problems).

Generally, planners don't want to handle a nervous system; therefore, in our instance, system nervousness causes a great number of changeovers, and, as a consequence, expensive schedules.
To address system nervousness, there has been increased concern in the development of different types of models that explicitly take into account uncertainties. The most important approach to optimization under uncertainty is stochastic programming. We refer to [14] and [15] as basic references for the theory and application of stochastic programming models.

1.2 Stochastic models

The aim of stochastic programming is to give solutions in terms of good decisions in problems in which there are some uncertain data. Stochastic is opposed to deterministic, meaning that some data are random, while programming means that the problem can be modeled as mathematical program. This field, also known as optimization under uncertainty, is increasing in interest rapidly in different research areas such as operations research, mathematics, economics, probability and statistics.
In this section we introduce stochastic programming theoretically (multistage stochastic models) and we discuss about the representation of uncertainty, according to the scenario tree mathematical formulation suggested by [16].

1.2.1 Multistage stochastic programs with recourse

In the multistage setting, the uncertain data ξ_1, ξ_2, ..., ξ_T is revealed gradually over time, in T periods, and decisions should be adapted to this process. The decision process has the form

$$\text{decision}(x_1) \to \text{observation}(\xi_2) \to \text{decision}(x_2) \to ... \to \text{decision}(x_T).$$

The sequence $\xi_t \in \mathbb{R}^{d_t}$, with $t = 1, 2, ..., T$, of data vectors as a stochastic process: it is a sequence of random variables with a specified probability distribution. With $\xi_{[t]}$ we denote the history of the process up to time t.
The values of the decision vector x_t, chosen at stage t, may depend on the information (data) $\xi_{[t]}$ up to time t, but not on the results of future observations. This is the basic requirement of nonanticipativity. As x_t may depend on $\xi_{[t]}$, the sequence of decisions is a stochastic process as well.
We say that the process ξ_t is stagewise independent if ξ_t is stochastically independent of $\xi_{[t-1]}$, $t = 2, ..., T$. It is said that the process is Markovian if for every $t = 2, ..., T$, the conditional distribution of ξ_t given $\xi_{[t-1]}$ is the same as the conditional distribution of ξ_t given ξ_{t-1}. Of course, if the process is stagewise independent, then it is Markovian.
In a generic form of a T-stage stochastic programming model can be written in the nested formulation

$$\min_{x_1 \in X_1} f_1(x_1) + \mathbb{E}\left[\inf_{x_2 \in f_2(x_1, \xi_2)} + \mathbb{E}\left[... + \mathbb{E}[\inf_{x_T \in X_T(x_{T-1}, \xi_T)} f_T(X_T, \xi_T)]\right]\right]$$

driven by the random data process ξ_1, ξ_2, ..., ξ_T. Here $x_t \in \mathbb{R}^{n_t}$, with $t = 1, 2, ..., T$, are decision variables, $f_t : \mathbb{R}^{n_t} \times \mathbb{R}^{d_t} \to \mathbb{R}$ are continuous functions and $X_t : \mathbb{R}^{n_{t-1}} \times \mathbb{R}^{d_t} \to \mathbb{R}^{n_t}$, $t = 1, 2, ..., T$, are measurable closed valued multifunctions. The first-stage data, i.e., the vector ξ_1, the function $f_1 : \mathbb{R}^{n_1} \to \mathbb{R}$ and X_1 are deterministic. It is said that a multistage problem is linear if the objective functions and the constraint functions are linear. In a typical formulation,

$$f_t(X_t, \xi_t) = c_t^T x_t, \quad X_1 := \{x_1 : A_1 x_1 = b_1, x_1 \geq 0\}$$

$$X_1(x_{t-1}, \xi_t) := \{x_t : B_t x_{t-1} + A_t x_t = b_t, x_t \geq 0\}, t = 2, ..., T.$$

Here, $\xi_1 := (c_1, A_1, b_1)$ is known at the first-stage (and hence is nonrandom), and $\xi_t := (c_t, B_t, A_t, b_t) \in \mathbb{R}^{d_t}$, $t = 2, ..., T$ are data vectors, some (or all) elements of which can be random.

1.2.2 Scenario tree representation

Scenario-based stochastic linear programs optimize under uncertain future conditions by producing contingent decision over a number of future scenarios. A typical set of scenarios, arranged in a branching, probabilistic tree, is shown in figure (1.2):

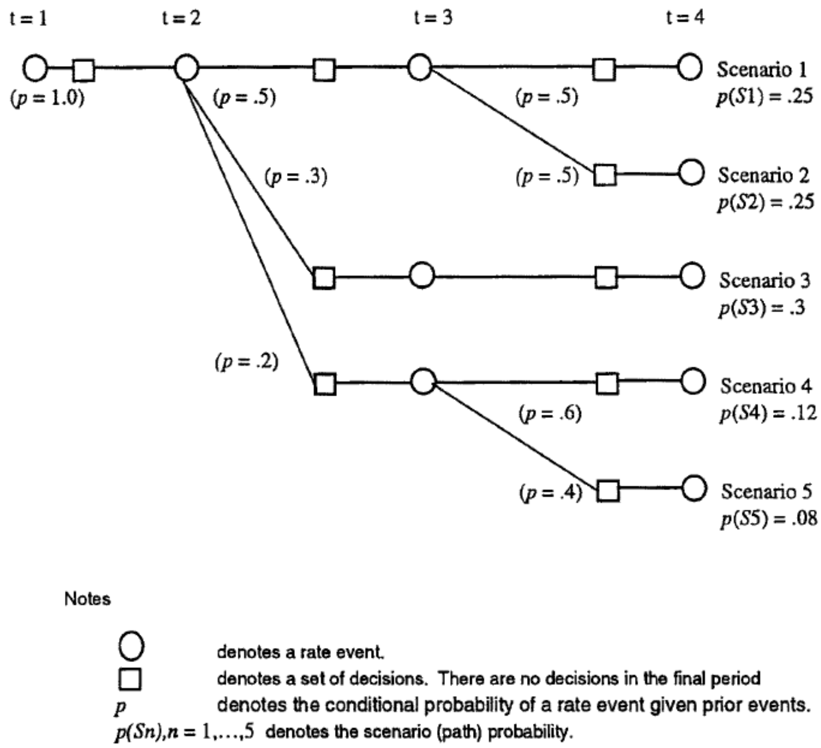


Figura 1.1: Example of branching rate probability tree

Each node in the tree corresponds to a time period with an associated problem state. At each branch point a random event occurs, and the next time period is associated with a number of new states based on the realization of the random variable with the discrete distribution described by the (conditional) branch probabilities. A scenario can be informally defined as a path through the tree from the root to a leaf; the scenario's probability is the probability of all its events occurring.

A scenario tree structure is defined by specifying the number of time periods, the scenario index set, and parent names, start time values and probabilities for all additional scenarios (the probabilities can be given as path probabilities or can be conditional on prior events in the parent scenario). There are five types of scenario structure:

1. arbitrary scenario structure with fixed horizon. The first type of problem assumes that random variable distributions are dependent on both the time period and prior history, which determines the current position in the tree. This can happen, for example, in planning problems when assumed scenarios reflect major, unique future events;

2. scenario structure with period to period independence of data value distributions;

3. scenarios with random walk or random walk with period independent drift. This type of problem uses random variables for changes in scenario data whose distributions are independent from period to period. Realizations in each scenario therefore depend on both the time period and prior history, although the increments can be specified by time period alone, regardless of scenario.

4. scenario trees in which random variable distributions of increments depend on prior events;

5. scenario trees in which the number of decision variables depends on the current scenario.

We focus the attention on the second type of scenario tree structure, showing what kind of parameters have to be defined. Considering the scenario tree represented above, the scenario structure can be defined following these steps:

1. definition of the parameter REALIZATIONS, that is the number of realizations or possible outcomes in the random variable's distribution;

2. definition of the parameter $\text{COND}_{\text{PROB}}$, indexed over time periods and realizations: the conditional probability of each outcome in the distribution;

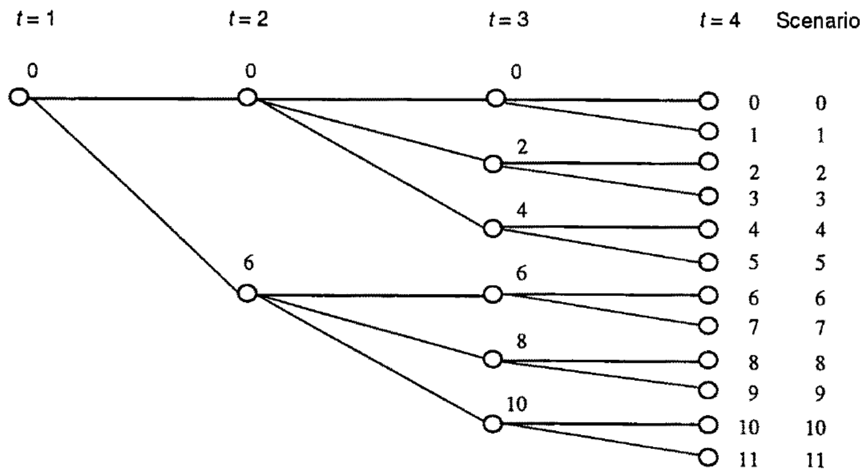


Figura 1.2: Scenario tree structure definition

3. definition of the parameter BRANCHES, indexed over time periods: the number of branches in the tree below a given time period;

4. definition of the index set SCENARIOS, calculated as all integers between 1 and the number of branches below time period 1;

5. definition of the parameter STARTTIME, indexed over scenarios, that is the start time of a certain scenario j;

6. definition of the parameter PARENT indexed over scenarios. The parent of a scenario j, is 0 for the base scenario and otherwise is the largest integer multiple of branches below the period prior to the start time of scenario j less than or equal to the scenario number;

7. definition of the set $\text{EVENT}_{\text{NODES}}$ and the parameter PREVIOUS, indexed over event nodes;

8. definition of the parameter $\text{RV}_{\text{ADDRESS}}$, indexed over event nodes, which associates with each node in the tree the number of the realization from the relevant distribution;

9. definition of the parameter $\text{PATH}_{\text{PROB}}$, indexed over event nodes: it is computed as the product of the conditional probability of each node of the path.

Considering the previous scenario tree we can compute some parameters, as an example:

Parameter	Stage 1	Stage 2	Stage 3	Stage 4
REALIZATIONS	1	2	3	2
$\text{COND}_{\text{PROB}}$	1	0.5	0.3333	0.5
BRANCHES	12	6	2	1

Parameter	Scenario 1	Scenario 3	Scenario 6	Scenario 10
STARTTIME	4	4	2	3
PARENT	0	2	0	6

1.3 Approximation of multistage stochastic models

In this paragraph we explain in a brief summary the approximation methodologies of multistage problems that authors used in the past.
Since we are going to consider a multistage stochastic model, the major issue in any application of multistage stochastic programming is the representation of uncertainty. [16], as we said in the previous paragraph, suggested a way to represent uncertainty using a scenario tree, but a key difficulty arising following this approach is in dealing with a very large number of scenarios (large-scale stochastic models), as reported in [8].
In our case, for example, a scenario-tree based stochastic model according to [16] is not possible because of the complexity of the environment we want to study (see chapter 3).
Many authors, even if in different research areas, suggested different ways to cope with this situation. We provide a short review about the different possible solutions for this problem, that can be found in the literature. All these solutions are about a reduction in the complexity of the model; this reduction can be reached with different approximation procedures:

1. reduction in the number of scenarios considered in the model using scenario generation methods;
2. reduction in the number of scenarios considered in the model using scenario reduction methods;
3. modification to the structure of the complete stochastic model (scenario tree based, [16]).

1.3.1 Scenario generation methods

One common solution is to apply a scenario generation method, consisting on generating a certain set of scenarios following some criteria. Many other researchers worked on this problem.

Dupocova et al. [17] recognized that a major issue in any application of multistage stochastic programming is the representation of the underlying random data process. They discussed the notion of representative scenarios (or representative scenario tree).
Kaut et al. [18] discussed the evaluation of quality/suitability of scenario generation methods for a given stochastic programming model. They formulated minimal requirements that should be imposed on a scenario generation method before it can be used for solving the stochastic programming model and they showed how to these requirements can be tested.
Mitra [20] tried to provide an overview of the main scenario generation methods. The main four scenario generation methods for multistage stochastic models are:

1. **Statistical Approaches.** The general idea behind applying statistical methods is to determine the values of particular statistical properties of some given (distribution of) data. Values of statistical properties can then in turn be used to determine the best fit theoretical distribution to data and so the theoretical distribution can be used for generating scenarios.

2. **Sampling.** This method takes sample from a given probability density function, thus providing scenario values. Sampling methods are a common approach for multistage stochastic programs, just as they are for two-stage models. Due to the exponential increase of the number of possible scenarios as the horizon length increases, multistage scenario generation approaches place a greater emphasis on reducing the number of required samples. The result is that the sampling procedure often involves considerable effort to ensure that the samples provide similar solution characteristics to a true underlying model. Main concerns are that the sample distribution has similar moments to the underlying distribution, that the sample distribution is not too distant from the underlying in terms of the probability of any event, and that the solution of the model using the sample distribution is consistent with practical limitations, such as the absense of arbitrage. Under mild conditions, these criteria can ensure that the sampling model solution converges asymptotically to a solution of the model with the underlying distribution.

3. **Simulation.** This involves the simulating of a process by inputting random numbers into its equation. The results give us the random variable's realizations, thus providing scenario values.

4. **Other methods (e.g. hybrid methods).** These include a variety of methods e.g. mixing scenario generation of sampling with moments matching.

1.3.2 Scenario reduction methods

A second solution to the problem is related to scenario reduction methods:

- Scenario reduction. This is a method for decreasing the size of a given tree. This method tries to find a scenario subset of prescribed cardinality, and a probability measure based on this set, that is closest to the initial distribution in terms of some probability metrics. The method is described in [21] and [22]. Furthermore, Kensu in [6] applies an interesting method of scenarios aggregation in a scheduling problem, proposed by [39].

- Internal sampling methods. Instead of using a pre-generated scenario tree, some methods for solving stochastic programming problems sample the scenarios during the solution procedure. The most important methods of this type are: stochastic decomposition [23], importance sampling with Benders' (L-shaped) decomposition ([24], [25] and [26]) and stochastic quasigradient methods ([27], [28]).

In addition, there are methods that proceed iteratively: they solve the problem with the current scenario tree, add or remove some scenarios and solve the problem again. Hence, at least in principle, the scenarios are added exactly where needed. The methods differ in the way decide where to add/remove scenarios: [29] uses dual variables from the current solution, while [30] measures the importance of scenarios by EVPI.

1.3.3 Stochastic model approximations

The first two solutions presented above basically keep the scenario tree structure and simply reduce the size of the scenario tree with different criteria. Another

possible idea consists on modifying the structure of the multistage stochastic model, in different approaches. The aim is to create an approximate stochastic model that takes into account demand uncertainty and that overcomes the computational expense associated with the solution of the large-scale stochastic multistage MILP model without losing too much information. This kind of solution that takes its basis not in the scenario tree representation does not have so many references in the literature.
Balasubramanian et al. [31] considered a problem of scheduling under demand uncertainty of a multiproduct batch plant. They presented a multistage stochastic MILP model, wherein certain decisions are made irrespective of the realization of the uncertain parameters (here and now variables) and some decisions are made upon realization of the uncertainty (wait and see variables). To cope with the computational issue associated with the solution of the large-scale stochastic multistage MILP for large problems, they examined an approximation strategy that relies on solving a number of smaller models, in particular it is based on the solution of a series of a two stage models within a shrinking horizon approach. Computational results indicate that the proposed approximation strategy provides an objective function within a few percent of the multistage stochastic MILP result in a fraction of the computation time and provides significant improvement in the expected profit over similar deterministic approaches.
The approximation of the stochastic model proposed by [31] consists on aggregating time periods in the horizon, so that the model that is going to be solved is a two-stage stochastic model. Another example of this kind of approximation is in [32]: the approach they studied consists of solving a sequence of two-stage stochastic programs with simple recourse, which can be viewed as an approximation to a multistage stochastic programming formulation of airplane revenue management problem.

The aim of this work is to present a proactive approach to setup management in the problem setting explained before, not reducing the multistage problem in a two-stage one as seen before. The idea is to perform an approximation strategy to solve a multistage stochastic model. In each stage some decisions have to be made according to demand forecasts of the products minimizing costs. A new feature is about how uncertainty is represented in the model. In scenarios the demand for each product is not represent by a point but by an interval. A scenario is thus a particular combination of demand intervals given by all the products. We choose to consider all the possibles scenarios (all the possible combinations of intervals) in each period and we compute wait and see variables for each scenario. Our approximation strategy consists on computing the expected value of these variables and consider them as an input (connection link with) for the following week.

A rolling horizon procedure is then applied at the end of each week: with the observation of the real value of the demand we can compute the real value of the wait and see variables (inventory and backorder). Thus, at the end of each period we make a new schedule for a certain number of periods (schedule horizon).
The idea of the expected values applied to wait and see variables has not ever been used in order to approximate the basic scenario tree described by Gassman, according to [15], [14].

In literature an example of expected values used in a scheduling problem, but in a two stage model, can be found in [35]. Baker used a stochastic model to solve a single machine scheduling problem, with processing times that follow normal distributions.

The objective is to use these expected values in a multistage model: the advantage we want to get is a better organization of the changeovers in the scheduling horizon, that leads into less expensive schedules, even in a rolling horizon approach.

Capitolo 2

Deterministic Model

The first scheduling method that we consider is a deterministic model that doesn't take into account the uncertainty of the demand. In this chapter we are going to describe in detail the mathematical model.
These are the main characteristics of the model:

- Mixed-Integer Linear Programming (MILP);
- Multi-product;
- Parallel machines. Every product can be processed on a general machine;
- Limited capacity;
- Costs included: setup costs and inventory carrying costs;
- Setup times are negligible: we assume that changeovers are made outside of working time;
- Demand uncertainty.

2.1 Description of the model

In the first part of this section we are going to describe the terms involved in the model while in the second part we explain the objective function and his relative constraints.

2.1.1 Terms of the model

The terms of the model are described in the table below.

Term	**Description**
P	Number of types of product: $j \in \{1, 2, ..., P\}$
T	Number of periods of the production schedule: $t \in \{1, 2, ..., T\}$
L	Number of parallel machines (capacity of the plant)
$Y[j, t]$	Number of machines that are configured to process product j in period t
$Z[j, t]$	Number of machines that process product j in period t
$\delta[j, t]$	Number of changeovers to process product j in period t
$I[j, t]$	Inventory amount for product j in period t
$B[j, t]$	Backorder amount for product j in period t
$d[j, t]$	Expected demand for product j in period t
$Y[j, 0]$	Machines configured to process product j in period $t = 0$
$I[j, 0]$	Inventory amount for product j in period $t = 0$
$B[j, 0]$	Backlog amount for product j in period $t = 0$
$c_Y[j]$	Setup costs for product j
$c_I[j]$	Inventory carrying costs for product j
BF	Backorder Factor

Some clarifications need to be done regarding the terms table above.
First of all, demand is expressed as the number of machines that need to produce for product j in period t. We can explain the demand as:

$$d[j,t] = \frac{\text{Quantity of product requested in period t}}{\text{Production rate}} \quad \left[\frac{\frac{\text{Unit}}{\text{Period}}}{\frac{\text{Unit}}{\text{Machine}\cdot\text{Period}}}\right] \tag{2.1}$$

$$= \text{Machines to run in period t for product j} \quad [\text{Machine}] \quad \forall j \in 1, ..., T$$

We simply assume that we have the demand already expressed as number of machines to run, not starting from the quantity of product requested. This way to express the demand simplifies significatively the description so that we don't need to make any considerations about production rates of the machines (also because this is not the aim of our research topic).

As a consequence of this, all the other variables are expressed in this way. For example, the Z variable doesn't represent a quantity of product, but the number of machines that need to run for a certain product in a specific period.
The same applies to the inventory and backorder, that are expressed as the number of looms that need to run in a given period to produce certain amounts for a specific product.

Regarding the variables type we can make some considerations:

- Z variables are allowed to be continuous in order to help softwares in finding the optimal solution during the optimization. Z variables doesn't need to be integer because, for example, $Z = 15.5$ means that we have to run 15 machines for a whole period and one machine has to run for a half period.
- Both I and B variables are continuous too, for the reason just explained above.
- Of course, Y and δ can't be allowed to be continuous variables, because it's not possible to make a number of changeovers different from an integer number.

As a recap, the variable type are shown in the following table:

Variable	**Type**
Y	Integer
Z	Continuous
δ	Integer
I	Continuous
B	Continuous

Before describing the objective function and constraints we show the variables involved in the model in a simple case (2 SKUs, 4 weeks).

2.1.2 Objective function and constraints

The model is characterized by two objective functions:

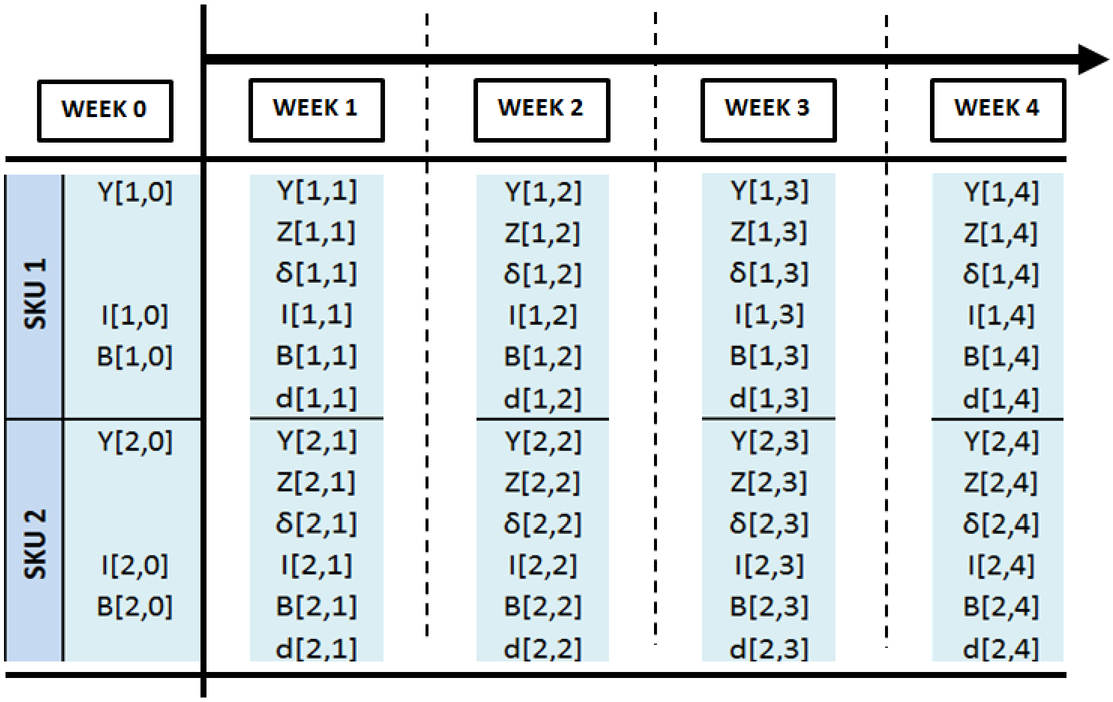


Figura 2.1: Example of variables involved in the model

1. minimization of backorder amount. The output of this part is the minimum backorder amount possible (B_{MIN}), considering the capacity, the demand of the products and the initial conditions of inventory and backorder.

2. minimization of total costs, taking into account the minimum backorder amount that is possible to reach.

The model involves two objective functions because the first one provides the second with important information about feasibility in terms of backorder.
Thus, the purpose of the first objective function is to find the minimal possible backorder amount:

$$\text{minimize:} \quad \sum_{j=1}^{P} B[j,T] \tag{2.2}$$

It's interesting to notice that this minimization is about the total backorder amount of all the products at the end of the scheduling horizon.

There are four constraints related to the first objective function, and they are

explained below.
Constraints (2.3) impose the total machines availability limits. Constraints (2.4) measure the positive changes in the number of machines dedicated to a product from one period to the next. This, together with (2.3) ensures that changeovers are measured correctly.

$$\sum_{j=1}^{P} Y[j,t] = L \quad \forall t \in \{1,2,...,T\} \tag{2.3}$$

$$Y[j,t] - Y[j,t-1] - \delta[j,t] \leq 0 \quad \forall j \in \{1,2,...,P\}, \forall t \in \{1,2,...,T\} \tag{2.4}$$

Constraints (2.5) are inventory balance equations for each product.

$$Z[j,t-1]+I[j,t-1]-B[j,t-1]-I[j,t]+B[j,t] = d[j,t] \quad \forall j \in \{1,2,...,P\}, \forall t \in \{1,2,...,T\} \tag{2.5}$$

The total number of machines that are configured for product j limits the total number of machines that can be operated to produce it. This is ensured by:

$$Z[j,t] \leq Y[j,t] \quad \forall j \in \{1,2,...,P\}, \forall t \in \{1,2,...,T\} \tag{2.6}$$

From this first minimization we get some important values that we use in one of the constraints of the second objective function. We define

$$B_{\text{MIN}}[j,T] = B[j,T] \quad \forall j \in \{1,2,...,P\}$$

where $B[j,T]$ comes from the solution of the first optimization.

In the second part of the model we find the second objective function:

$$\text{minimize:} \ \sum_{j=1}^{P}\sum_{t=1}^{T} c_Y \cdot \delta[j,t] + \sum_{j=1}^{P}\sum_{t=1}^{T} c_I \cdot I[j,t] \tag{2.7}$$

This second objective function is about minimizing the total of changeover costs and inventory carrying costs over the scheduling horizon. Backorders are not con-

sidered as a cost in the objective function because we assume that it's not possible to assign a monetary value to them. However, backorder is handled as a constraint in terms of quantities.
The constraints for this objective function are the same of the first one. Therefore, there is one additional contraint about a limitation of the backorder amount. Before showing this constraint we need to introduce the "Backorder factor" (BF), an input parameter that represents how much backorder is allowed for each product with respect to the total demand over the periods of the horizon, with:

$$0 < BF < 1.$$

The percentage of demand satisfaction can be defined as "Service Factor" (SF):

$$SF = 1 - BF.$$

Now we can introduce the last constraint, that is:

$$B[j,T] \leq B_{\text{MIN}}[j,T] + \text{BF} \cdot \left(\sum_{t=1}^{T} d[j,t] \right) \quad \forall j \in \{1, 2, ..., P\} \tag{2.8}$$

Setting BF = 0 (corresponding to SF = 1) we are asking the model to provide the best schedule in terms of service performance. On the contrary, increasing BF means that we are relaxing this service constraint decreasing the value of SF. In other words, we are giving to the objective function more freedom to minimize the costs (and we are consequently moving towards a cost performance). The behavior of the model on varying BF is qualitatively described in figure (2.2).

Figura 2.2: Cost and service performance

In conclusion, we propose a recapitulatory scheme of the model in figure (2.3).

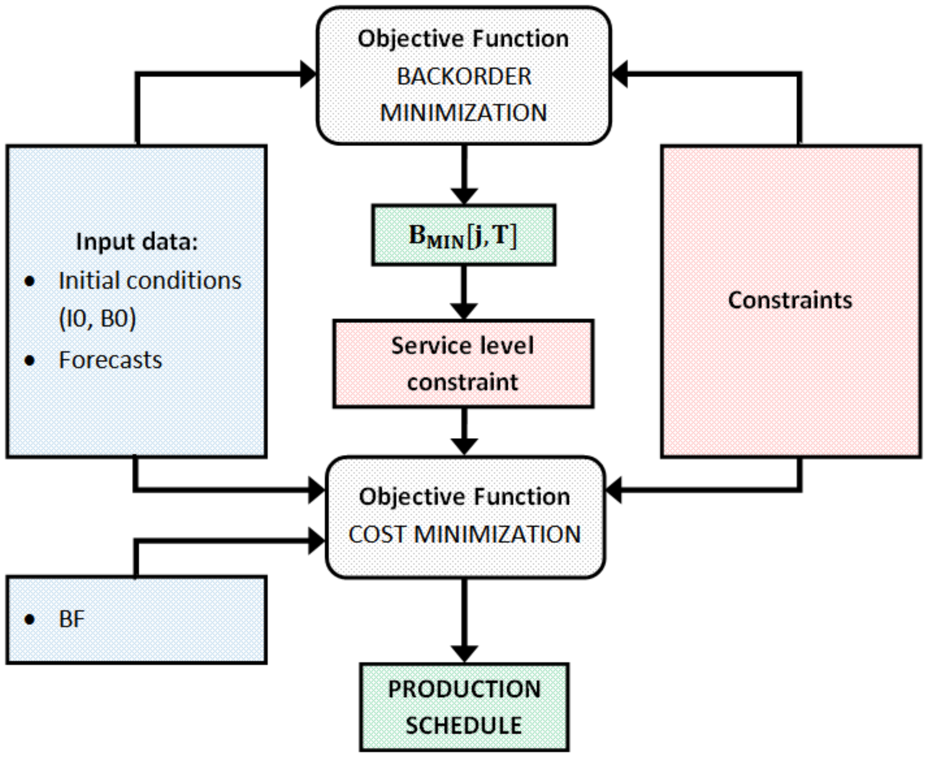


Figura 2.3: Scheme of the deterministic model

Capitolo 3

Stochastic model

In this chapter we present a stochastic model, in which demand uncertainty is taken into account within the optimizing model through the implementation a scenario tree, according to [16]. We start from the mathematical formulation of the model (terms of the model, objective function and constraints), we present the main problems of this approach and we suggest, as an answer, an approximated stochastic model.
These are the main characteristics of the model:

- Mixed-Integer Linear Programming (MILP);
- Multi-product;
- Parallel machines. Every product can be processed on a general machine;
- Limited capacity;
- Costs included: setup costs and inventory carrying costs;
- Setup times are negligible: we assume that changeovers are made outside of working time;
- Demand uncertainty.

3.1 Description of the model

In the first part of this section we are going to describe the terms involved in the model while in the second part we explain the objective function and his relative constraints.

3.1.1 Terms of the model

Term	Description
P	Number of types of product: $j \in \{1, 2, ..., P\}$
T	Number of periods of the production schedule: $t \in \{1, 2, ..., T\}$
L	Number of parallel machines (capacity of the plant)
S	Number of scenarios: $s \in \{1, 2, ..., S\}$
$Y[j, s, t]$	Number of machines that are configured to process product j under scenario s in period t
$Z[j, s, t]$	Number of machines that process product j under scenario s in period t
$\delta[j, s, t]$	Number of changeovers to process product j under scenario s in period t
$I[j, s, t]$	Inventory amount for product j under scenario s in period t
$B[j, s, t]$	Backorder amount for product j under scenario s in period t
$d[j, s, t]$	Expected demand for product j under scenario s in period t
$Y[j, 0]$	Machines configured to process product j in period $t = 0$
$I[j, 0]$	Inventory amount for the product j in period $t = 0$
$B[j, 0]$	Backlog amount for the product j in period $t = 0$
$c_Y[j]$	Setup costs for product j
$c_I[j]$	Inventory carrying costs for product j
N^1	Set of all the nodes except the leaf nodes
N^2	Set of all the nodes except the root node
$p[s, t]$	Conditional probability of the path from root to (s, t)
BF	Backorder Factor

Some observations need to be done regarding the terms of the table above, as we did in chapter 2 for the deterministic model.

Variables Y and Z are related to the configuration of the machines: we introduce Z variable because it's possible to configure a machine to produce a certain product without running it if it's not necessary.
Demand is expressed as the number of machines that need to produce for product j in period t. We can explain the demand as:

$$d[j,s,t] = \frac{\text{Quantity of product requested in period t under scenario s}}{\text{Production rate}} \quad \left[\frac{\frac{\text{Unit}}{\text{Period}}}{\frac{\text{Unit}}{\text{Machine·Period}}}\right] \tag{3.1}$$

$$= \text{Machines to run in period t for product j under scenario s} \quad [\text{Machine}] \quad \forall j \in 1, ..., T$$

We simply assume that we have the demand already expressed as number of machines to run, not starting from the quantity of product requested. This way to express the demand simplifies significatively the description so that we don't need to make any considerations about production rates of the machines (also because this is not the aim of our research topic).

As a consequence of this, all the other variables are expressed in this way. For example, the Z variable doesn't represent a quantity of product, but the number of machines that need to run for a certain product in a specific period.
The same applies to the inventory and backorder, that are expressed as the number of looms that need to run in a given period to produce certain amounts for a specific product.

Regarding the variables type we can make some considerations:

- Z variables are allowed to be continuous in order to help softwares in finding the optimal solution during the optimization. Z variables doesn't need to be integer because, for example, $Z = 15.5$ means that we have to run 15 machines for a whole period and one machine has to run for a half period.
- Both I and B variables are continuous too, for the reason just explained above.
- Of course, Y and δ can't be allowed to be continuous variables, because it's not possible to make a number of changeovers different from an integer

number.

As a recap, the variable type are shown in the following table:

Variable	**Type**
Y	Integer
Z	Continuous
δ	Integer
I	Continuous
B	Continuous

3.1.2 Objective function and constraints

The model is characterized by two objective functions:

1. minimization of backorder amount. The output of this part is the minimum backorder amount possible (B_{MIN}), considering the capacity, the demand of the products and the initial conditions of inventory and backorder.

2. minimization of total costs, taking into account the minimum backorder amount that is possible to reach.

The model involves two objective functions because the first one provides the second with important information about feasibility in terms of backorder.
Thus, the purpose of the first objective function is to find the minimal possible backorder amount:

$$\text{minimize: } \sum_{j=1}^{P} B[j, s, T] \tag{3.2}$$

It's interesting to notice that this minimization is about the total backorder amount of all the products at the end of the scheduling horizon.

There are four constraints related to the first objective function, and they are explained below.
Constraints (3.3) impose the total machines availability limit. Constraints (3.4) measure the positive changes in the number of machines dedicated to a product from one period to the next. This, together with (3.3) ensures that changeovers are measured correctly.

$$\sum_{j=1}^{P} Y[j,s,t] = L \quad \forall (s,t) \in N^1 \tag{3.3}$$

$$Y[j,s,t] - Y[j,\text{PREV}(s,t),t-1] - \delta[j,s,t] \leq 0 \quad \forall j \in \{1,2,...,P\}, \forall (s,t) \in N^1 \tag{3.4}$$

Constraints (3.5) are inventory balance equations for each product.

$$Z[j,\text{PREV}(s,t),t-1] + I[j,\text{PREV}(s,t),t-1] - I[j,s,t] + \tag{3.5}$$

$$-B[j,\text{PREV}(s,t),t-1] + B[j,s,t] = d[j,s,t] \quad \forall j \in \{1,2,...,P\}, \forall (s,t) \in N^2$$

The total number of machines that are configured for product j limits the total number of machines that can be operated to produce it. This is ensured by:

$$Z[j,s,t] \leq Y[j,s,t] \quad \forall j \in \{1,2,...,P\}, \forall (s,t) \in N^1 \tag{3.6}$$

From this first minimization we get some important values that we use in one of the constraints of the second objective function. We define

$$B_{\text{MIN}}[j,s,T] = B[j,s,T] \quad \forall j \in \{1,2,\dots P\} \ \forall (s,t) \in N^2$$

where $B[j,s,T]$ comes from the solution of the first optimization.

In the second part of the model we find the second objective function:

$$\text{minimize:} \quad \sum_{(s,t)\in N^1} p[s,t] \left(\sum_{j\in P} c_Y \cdot \delta[j,s,t] \right) + \sum_{(s,t)\in N^2} p[s,t] \left(\sum_{j\in P} c_I \cdot I[j,s,t] \right) \tag{3.7}$$

This second objective function is about minimizing the total of changeover costs and inventory carrying costs over the scheduling horizon. Backorders are not considered as a cost in the objective function because we assume that it's not possible to assign a monetary value to them. However, backorder is handled as a constraint in terms of quantities.
The constraints for this objective function are the same of the first one. Therefore,

there is one additional contraint about a limitation of the backorder amount, based on backorder factor (BF) previously defined in chapter 2.

$$B[j,s,T] \leq B_{\text{MIN}}[j,s,T] + \text{BF} \cdot \left(\sum_{t=1}^{T} d[j,s,t] \right) \quad \forall j \in \{1,2,...,P\} \; \forall (s,t) \in N^2 \tag{3.8}$$

3.2 Limitations for the scenario tree based stochastic model

The main issue for the stochastic model is the representation of the uncertainty. As we said in chapter 1, the most used approach is the one suggested by [16], the scenario tree based stochastic model.
We try to define this scenario tree structure in a general way according to the scheduling problem considered in this work, considering the following hypothesis:

- We assume that stages are independent from one period to another. In other words, events that modify the demand of the products are not related to the previous events.

- We also assume that we know (for example, from historical data) the variability of the demand for each product: demand of a single product can be described by an expected value function $E(t)$ and by a definition of its standard deviation σ.

3.2.1 Scenario tree for the single products

For each product we identified a possible discrete function for the expected demand over time. We assume that in each period the demand for the product is characterized by an expected value and a standard deviation, for example (3.1):
We can generate two or more scenarios discretizing the probability distribution, identifying some demand values (points) and certain probabilities. In the simple case of two scenarios (in period $t = 1$), the discretization can be done identifying two demand values with the same probability at a certain distance (for example,

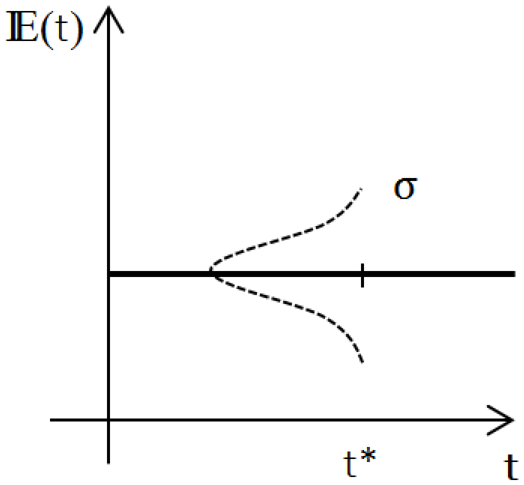


Figura 3.1: Expected demand and standard deviation for a single product

0.7σ) from the expected demand value, as we can see in figure (3.2):

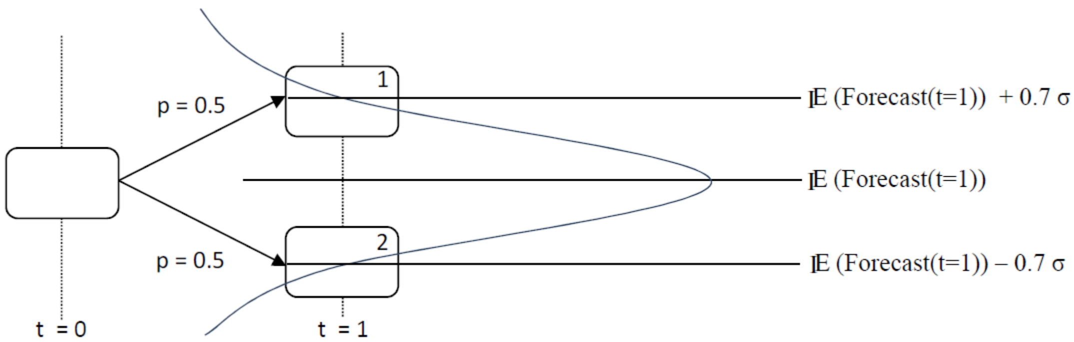


Figura 3.2: Discretization of a single product demand probability distribution function

Since we are going to prepare schedules for a certain number of periods, we have to consider more periods and more scenarios. See figure (3.3) for an example of a single product, with two outcomes for each period and with $t = 4$ periods.

For a T periods model and considering S scenarios for each product and for each

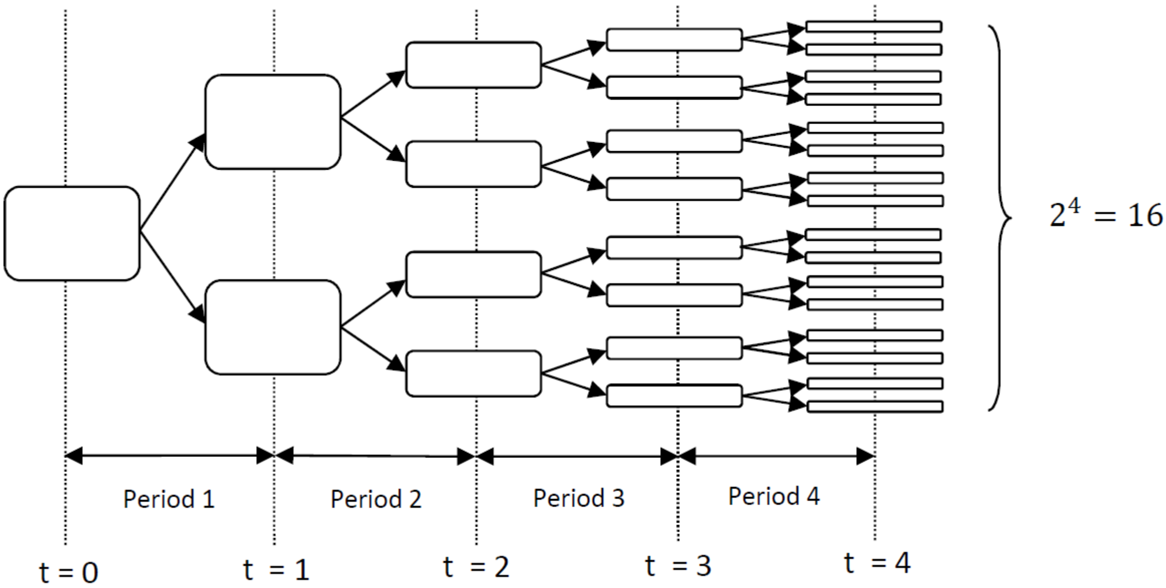


Figura 3.3: Scenario tree for a single product, considering $t = 4$ and two outcomes for each period

period, the total number of scenarios at the end of the periods is:

$$\text{Number of scenarios} = S^T \tag{3.9}$$

Indeed, in this case (3.3) there are 16 scenarios. Computing the number of nodes of the tree is possible to compute the total number of variables involved in the model.

$$\text{Number of nodes} = \sum_{t=0}^{T} S^t \tag{3.10}$$

This means that in our case there are 31 nodes.

3.2.2 Scenario tree for all the products

For each product can be defined a scenario tree like in figure (3.3), but the complete scenario tree is a combination of all the single trees. Thus, it's important to choose in a proper way the number of products to be considered and the number of scenarios for each product and for each period.

For example, considering three products (A, B and C) and two scenarios for each product and for each period, for each node of the tree there will be a situation like in the following figure (3.4). In the same figure we can see the complete scenario tree, with $T = 4$ periods; only some of the involved nodes are underlined.

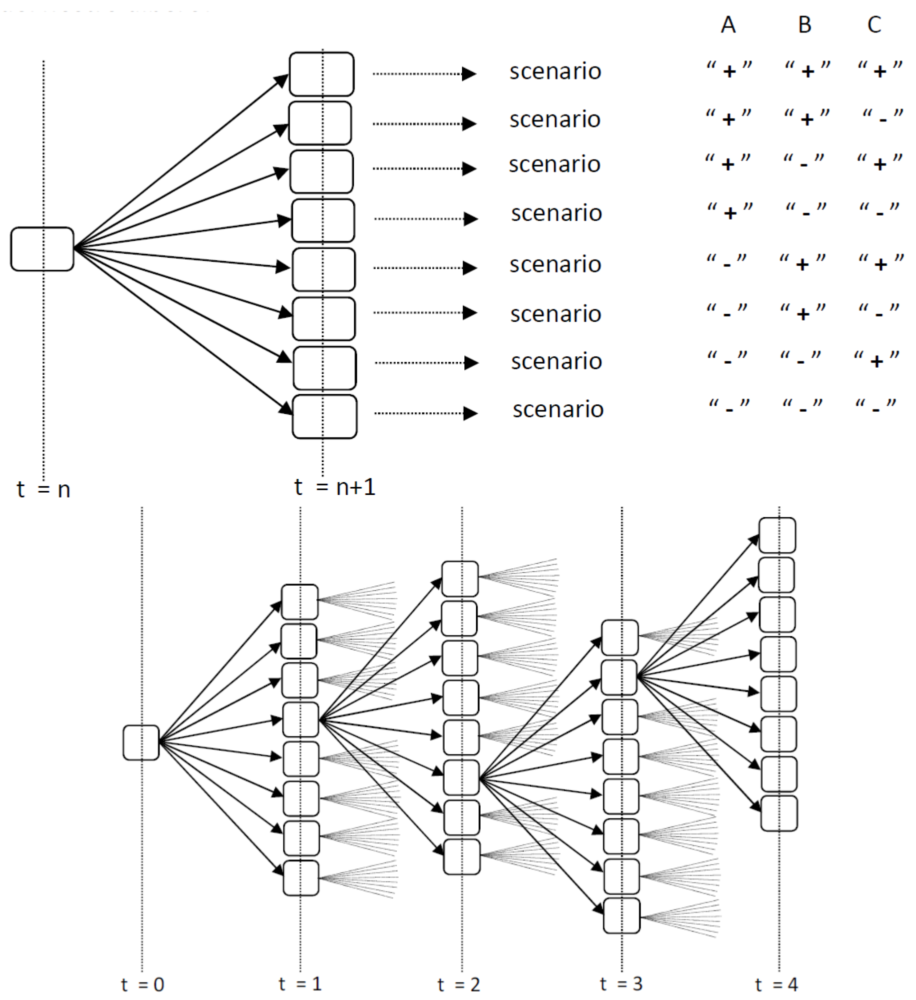


Figura 3.4: Scenarios for each node and complete scenario tree (four periods) with three products and two outcomes for each product and for each period.

In figure (3.4) the total number of nodes is 4680. Now we are going to compute how many variables are involved in a problem with these data ($T = 4$, two outcomes per period and per product). All the variables in the models are wait and see (they depend on the scenario): this means that every single term in the model has 3 products $\cdot$ 4680 nodes = 14040 variables. Since there are 5 terms involved in the model (Y, Z, delta, I, B) the total number of variables is $14040 \cdot 5 = 70200$. This causes computational problems, since the model operates with a large number of integer variables.
In the following table we compute the total number of variables that the model has to handle with, considering a certain number of scenarios per node.

Scenarios per node	**# Nodes**	**Total number of variables**
8	4680	70200
9	7380	110700
10	11110	166650
11	16104	241560
12	22620	339300

We underline the fact that this large number of scenarios (nodes of the tree) has been reached with only 3 products and with only 2 scenarios for each product and for each period. It's our aim to consider more scenarios in the models, because we don't want to omit too many instances of demand; this fact doesn't let us to use a complete scenario tree based stochastic model, but it's necessary to make some changes. The authors suggest not to reduce the number of scenarios for each period, but to create an approximate stochastic model without omitting information in the representation of the uncertainty.

3.3 Stochastic approximated model

Since the number of variables for the scenario tree-based stochastic model is critical, it's necessary to develop a model that takes in account the computational issues. The model we are going to present, which we can call approximated stochastic model, tries to cope with this situation.

3.3.1 Introduction to the new model

In this section we focus on the new elements introduced in the stochastic model.

In the scenario tree based stochastic model all the variables are Wait and See (WS), meaning that their value is defined for each node in the scenario tree, or, in other words, that they depend on the realization of the uncertainty. Our approximation consists on considering Wait and See variables only I and B. This means that Y, Z and δ are Here and Now (HN), that is to say that they don't depend on the observation of the actual demand. This is the first new element of the approximated stochastic model.

The second new element is about the structure of the multistage problem. Instead

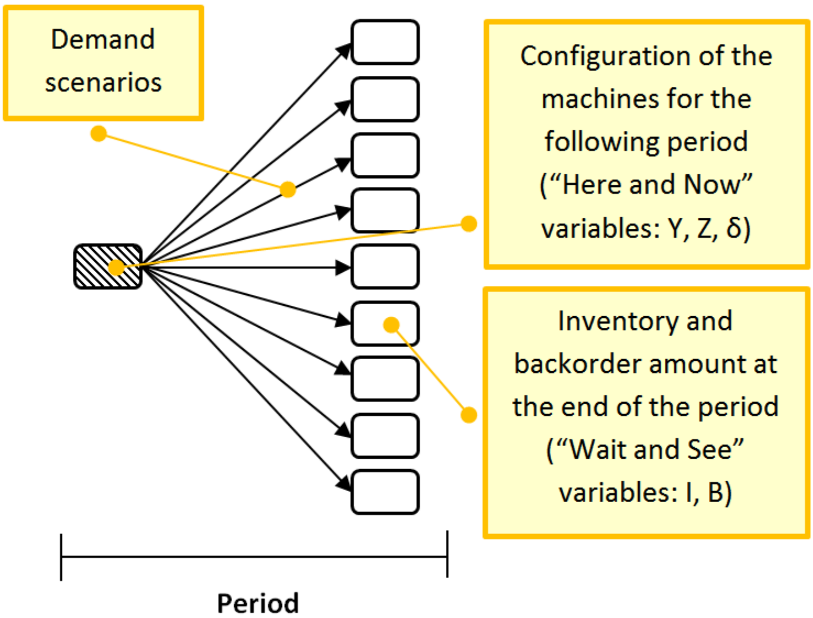


Figura 3.5: Here and Now and Wait and See variables

of developing every single node, like a scenario tree, we compute the expected values of the Wait and See variables (inventory and backorder), that they become the initial conditions for the following week. In this way the number of scenarios doesn't grow exponentially. In figure (3.6) we show better this concept.

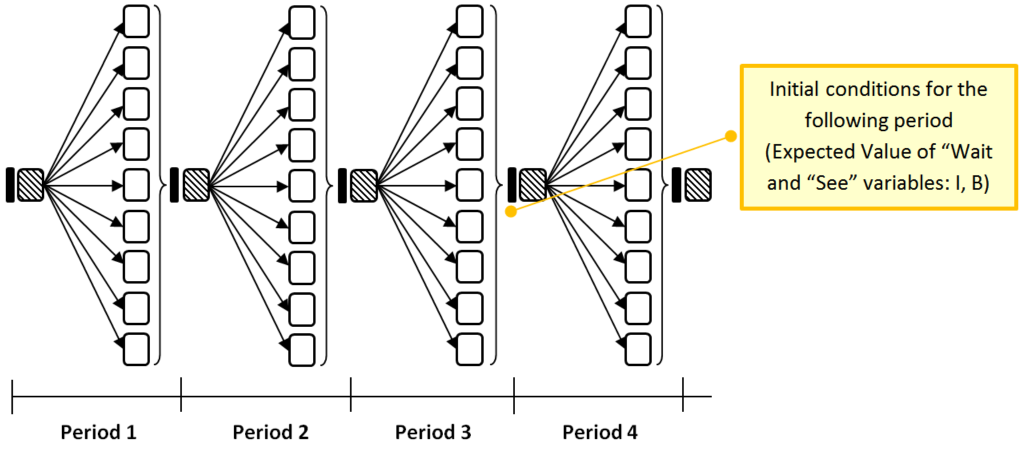


Figura 3.6: Expected value of the Wait and See variables

The number of variables, in this new model, is significatively reduced. Considering the same example of paragraph (3.2.2) (3 products, 4 periods and 2 outcomes for each period and for each product):

Stochastic Model	Scenarios per node	Total number of variables
Complete	8	70200 (WS)
Approximated	8	192 (WS) + 36 (HN) = 228

Since the number of variables is very low, it is possible to define more scenarios for each period and consider more instances of demand realization. To do this, we consider the probability distribution function of the products and, assuming to know (for example, from historical data) their mean and standard deviation, we follow these steps:

- we consider the interval $[\mu - 2.5\sigma; \mu + 2.5\sigma]$. The probability to obtain an actual demand for this product in the interval $(-\infty; -2.5\sigma)$ is equal to 0.0062. Similarly, the probability to obtain a demand in the interval $(+2.5\sigma; +\infty)$ is the same. This means that the interval chosen covers the 98.76% of all the instances;

- we divide this interval in a certain number of equal sub-intervals (for example, 5);

- we compute the probability of each sub-interval;

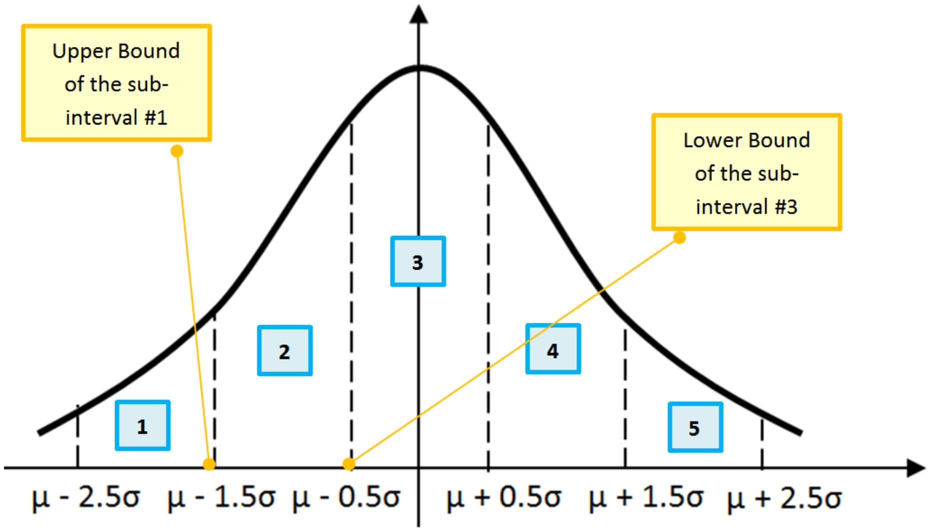


Figura 3.7: Demand partition

- we re-do the first three steps for all the products;
- we generate all the scenarios, that are all the possible combinations of the sub-intervals previously defined.

In figure (3.8) an example of scenario is shown. In this case (3 products, 5 outcomes for each product for each period, 4 periods) the number of variables becomes), the number of variables becomes:

Stochastic Model	**Scenarios per node**	**HN variables**	**WS variables**
Complete	125	0	$\approx 3.7 \cdot 10^9$
Approximated	125	36	3000

Even if we compare the number of variables of the approximated model just showed (3 products, 5 outcomes for each product for each period, 4 periods) with the complete model described before (3 products, 2 outcomes for each product for each period, 4 periods) the number of the variables is still smaller for the approximated.

Stochastic Model	**Scenarios per node**	**HN variables**	**WS variables**
Complete	8	0	70200
Approximated	125	36	3000

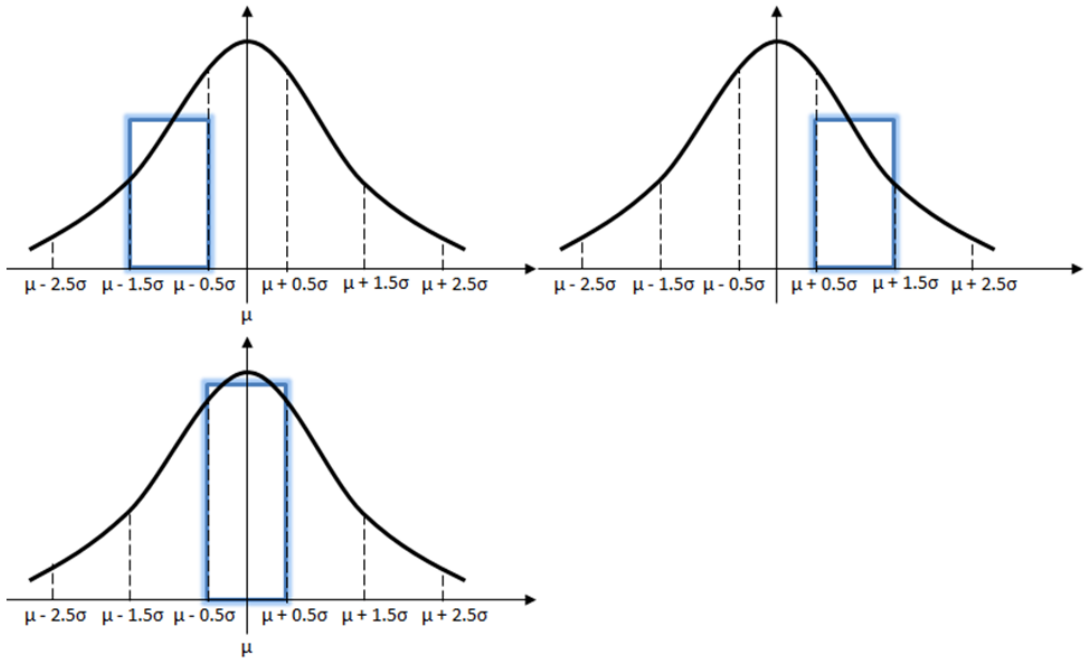


Figura 3.8: Example of scenario

It seems that the computational problem has been overcome with our new model.

Note that in this approximated stochastic model demand is not as a point value, but an interval with associated a certain probability. This is the third and last new element of this model.

3.3.2 Terms of the model

Term	Description
P	Number of types of product: $j \in \{1, 2, ..., P\}$
T	Number of periods of the production schedule: $t \in \{1, 2, ..., T\}$
L	Number of parallel machines (capacity of the plant)
O	Set of the o outcomes (scenarios): $o \in O$
$Y[t, j]$	Machines, in period t, that are configured to process product j
$Z[t, j]$	Machines, in period t, that produce product j
$\delta[t, j]$	Number of changeovers, in period t, to produce product j
$I[t, j, o]$	Inventory amount in period t for product j under outcome o
$B[t, j, o]$	Backlog amount in period t for product j under outcome o
$\text{LB}[t, j, o]$	Lower bound of the demand interval
$\text{d}[t, j]$	Demand forecast for period t for product j
$\text{UB}[t, j, o]$	Upper bound of the demand interval
$Y[0, j]$	Machines configured to process product j in period $t = 0$
$I[0, j]$	Inventory amount for product j in period $t = 0$
$B[0, j]$	Backlog amount for product j in period $t = 0$
$c_Y[j]$	Setup costs
$c_I[j]$	Inventory carrying costs for product j
$p[o]$	Probability of the outcomes
BF	Backorder Factor

Some clarifications need to be done regarding the terms table above.
First of all, demand is expressed as the number of machines that need to produce for product j in period t. We can explain the demand as:

$$d[j, t] = \frac{\text{Quantity of product requested in period t}}{\text{Production rate}} \quad \left[\frac{\frac{\text{Unit}}{\text{Period}}}{\frac{\text{Unit}}{\text{Machine}\cdot\text{Period}}}\right] \tag{3.11}$$

$$= \text{Machines to run in period t for product j} \quad [\text{Machine}] \quad \forall j \in 1, ..., T$$

We simply assume that we have the demand already expressed as number of machines to run, not starting from the quantity of product requested. This way to express the demand simplifies significatively the description so that we don't need to make any considerations about production rates of the machines (also because this is not the aim of our research topic).

As a consequence of this, all the other variables are expressed in this way. For example, the Z variable doesn't represent a quantity of product, but the number of machines that need to run for a certain product in a specific period.
The same applies to the inventory and backorder, that are expressed as the number of looms that need to run in a given period to produce certain amounts for a specific product.

Regarding the variables type we can make some considerations:

- Z variables are allowed to be continuous in order to help softwares in finding the optimal solution during the optimization. Z variables doesn't need to be integer because, for example, $Z = 15.5$ means that we have to run 15 machines for a whole period and one machine has to run for a half period.
- Both I and B variables are continuous too, for the reason just explained above.
- Of course, Y and δ can't be allowed to be continuous variables, because it's not possible to make a number of changeovers different from an integer number.

As a recap, the variable type are shown in the following table:

Variable	**Type**
Y	Integer
Z	Continuous
δ	Integer
I	Continuous
B	Continuous

3.3.3 Objective function and constraints

The model is characterized by two objective functions:

1. minimization of backorder amount. The output of this part is the minimum backorder amount possible ($\mathbb{E}[B_{\text{MIN}}]$), considering the capacity, the demand of the products and the initial conditions of inventory and backorder.

2. minimization of total costs, taking into account the minimum backorder amount that is possible to reach.

The model involves two objective functions because the first one provides the second with important information about feasibility in terms of backorder.
Thus, the purpose of the first objective function is to find the minimal possible backorder amount:

$$\text{minimize: } \sum_{j=1}^{P} \mathbb{E}[B[T, j]] \tag{3.12}$$

It's interesting to notice that this minimization is about the total backorder amount of all the products at the end of the scheduling horizon.

There are four constraints related to the first objective function, and they are explained below.
Constraints (3.13) impose the total machines availability limits. Constraints (3.14) measure the positive changes in the number of machines dedicated to a product from one period to the next. This, together with (3.13) ensures that changeovers are measured correctly.

$$\sum_{j=1}^{P} Y[t, j] = L \quad \forall t \in 1, ..., T \tag{3.13}$$

$$Y[t, j] - Y[t-1, j] - \delta[t, j] \leq 0 \quad \forall j \in \{1, 2, ..., P\}, \forall t \in 1, ..., T \tag{3.14}$$

Constraints (3.15) and (3.16) are inventory balance equations for each SKU.

$$Z[t, j] + \mathbb{E}[I[t-1, j]] - \mathbb{E}[B[t-1, j]] - I[t, j, o] + B[t, j, o] \geq \text{UB}[t, j, o] \tag{3.15}$$

$$\forall j \in \{1, 2, ..., P\}, \forall t \in \{1, 2, ..., T\}$$

$$Z[t,j] + \mathbb{E}[I[t-1,j]] - \mathbb{E}[B[t-1,j]] - I[t,j,o] + B[t,j,o] \leq \text{LB}[t,j,o] \quad (3.16)$$

$$\forall j \in \{1,2,...,P\}, \forall t \in \{1,2,...,T\}$$

The total number of machines that are configured for product j limits the total number of machines that can be operated to produce it. This is ensured by:

$$Z[t,j] \leq Y[t,j] \quad \forall j \in \{1,2,...,P\}, \forall t \in \{1,2,...,T\} \quad (3.17)$$

From this first minimization we get some important values that we use in one of the constraints of the second objective function. We define

$$\mathbb{E}[B_{\text{MIN}}[T,j]] = \mathbb{E}[B[T,j]] \quad \forall j \in \{1,2,\ldots P\}$$

where $\mathbb{E}[B[T,j]]$ comes from the solution of the first optimization.

In the second part of the model we find the second objective function:

$$\text{minimize:} \; \sum_{j=1}^{P} \sum_{t \in T} c_Y \cdot \delta[t,j] + \cdot \sum_{j=1}^{P} \sum_{t \in T} c_I \cdot p[o] \cdot I[t,j,o] \quad (3.18)$$

This second objective function is about minimizing the total of changeover costs and inventory carrying costs over the scheduling horizon. Backorders are not considered as a cost in the objective function because we assume that it's not possible to assign a monetary value to them. However, backorder is handled as a constraint in terms of quantities.
The constraints for this objective function are the same of the first one. Therefore, there is one additional contraint about a limitation of the backorder amount.

$$\mathbb{E}[B[t,j]] \leq E[B_{\text{MIN}}[T,j]] + \text{BF} \cdot \left(\sum_{t=1}^{T} d[t,j] \right) \quad \forall j \in \{1,2,...,P\} \quad (3.19)$$

In conclusion, we propose a recapitulatory scheme of the model in figure (3.9).

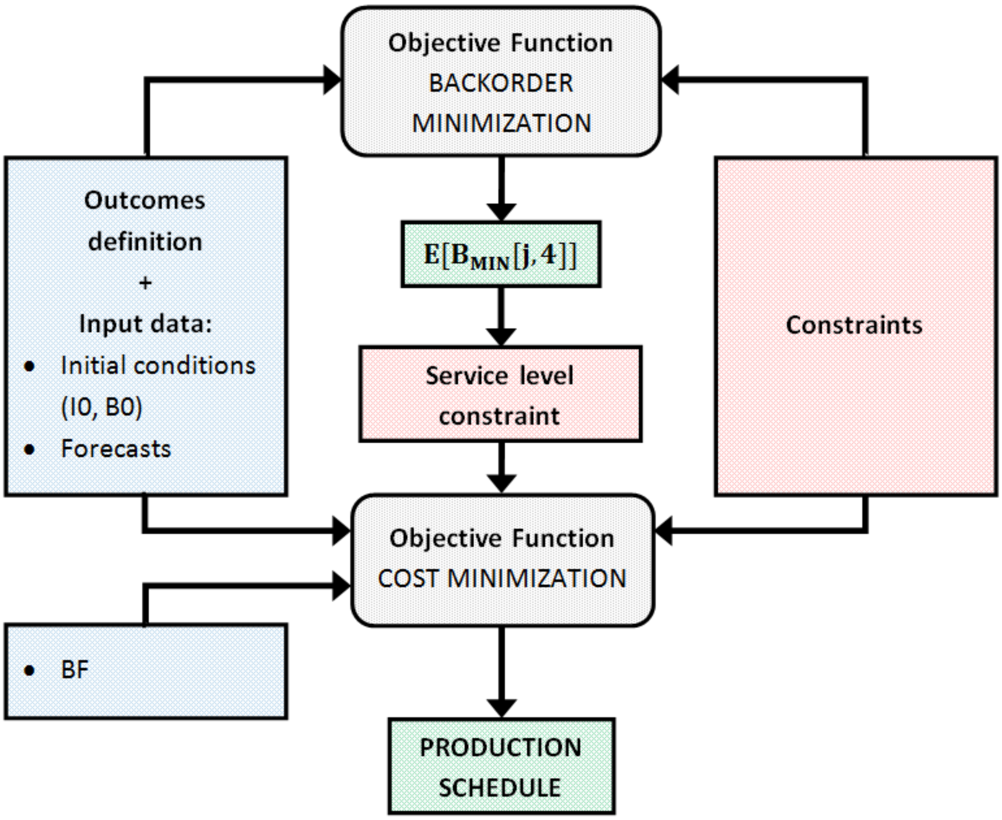


Figura 3.9: Scheme of the stochastic model

Capitolo 4

Textile manufacturing scheduling problem

In this chapter we present the particular instance in which we want to check how the models presented in the previous chapters work; we want to see if a stochastic model is a good answer to the reactive (nervous) modification of production schedules due to the uncertainty of demand, that is obtained if a deterministic approach is used together with a rolling horizon procedure.
We consider a textile manufacturing scheduling problem, because textile field is highly unstable due to long manufacturing lead time, change in consumer preferences, cyclic demand, fashion changes and sensitivity to general economic climate [19]. During manufacturing lead time, demand for a product can increase or decrease unexpectedly,
This chapter is organized as follows: in the first section a description of the process is presented, followed by a focus on the demand uncertainty in the textile environment. In the third section we specify what are the products we are going to consider and in the last section we provide the two models adapted to the new context.

4.1 Textile manufacturing process

Textile manufacturing process consists of three stages: yarn manufacturing (1), fabric formation (2) and finishing and dyeing processes (3), as we can see in the figure (4.1). Yarn is produced by *blending*, *combing*, *carding*, *roving* and *spinning*

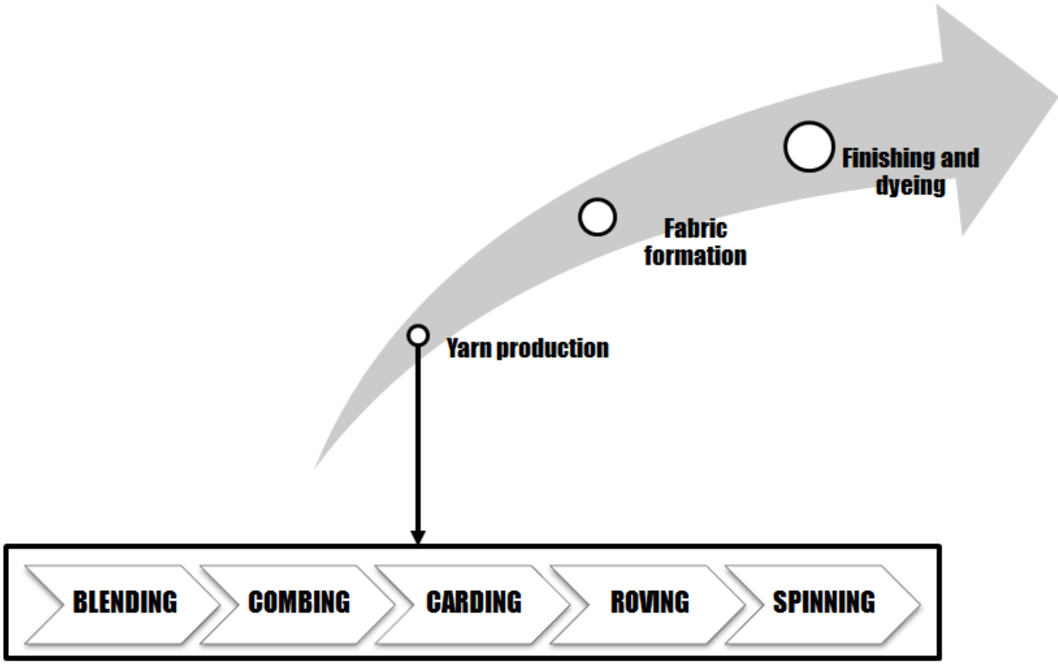


Figura 4.1: Textile manufacturing process

of natural or man-made fibres. Fibres first go through a blending and cleaning machine. This process cleans impurities, mixes fibres of different kinds and feeds to *combing*, where short fibres strands start to become aligned. At this point the material is called slyver. The next process, called *carding*, combs and twists slyver to increase its strenght and to reduce its thickness. Some blends directly go through a group of roving machines. The *roving* process further twists and strenghtens yarn to a certain volume and thus it is ready for the *spinning* process. Through *spinning*, yarn is differentiated one step further with respect to its thickness. *Spinning* is performed on spinning machines (frame), which twist the yarn at high speeds and give to it its final strength.
The following steps in the supply chain after yarn production are fabric formation (object of our research) and finishing and dyeing processes.

We focus on the fabric formation context, and we are going to apply our scheduling models within this field. In fabric formation plants the length/size of the period is usually the week. Changeovers occur outside the working time, generally on the weekend. In this kind of production system setup management is really important, because changeovers are generally very expensive compared to inventory carrying costs. In addition to that, they partially affect what and how much is going to be produced in the following period, becoming important also in terms of backorder. The real challenge is to produce minimizing the total number of changeovers, providing the customer with a high service level.

4.2 Demand uncertainty in textile industry

Textile industry is a very complex field as regards demand uncertainty, due to the long time-to-market which contrasts with the short life cycle of the products, due to seasonal sales, exogenous variables, high variety of products and volatile consumer demands (fashion effect). This leads to very challenging and difficult forecasting processes.

The complexity associated with the sales forecasting in textile industry induced many authors to work on this topic.

According to [36], the main characteristics which should be taken into account to design a sales forecasting system for the textile industry are:

1. **Forecasting horizon**. Higher is the horizon, better is the anticipation but higher are the errors of forecasts, so the definition of the horizon becomes very important.

2. **Product Life Cycle**. The life cycle, for most the products, consists of 4 phases (launch, rise, maturation and decline). Life cycle in textile industry is quite short [46].

3. **Product aggregation**. The product variety is an important characteristic to take into account: thus, the SKU (Stock Keeping Unit) management is very hard. One of the biggest issues is to determine the level and the right criteria for the aggregation.

4. **Seasonality**. Most of the products, especially in the apparel market, present a seasonal trend within the year. Therefore, weather causes modifications on the sales during the year.

5. **Exogenous variables** (end-of-season sale, promotions, ...).

The most used techniques for sales forecasting in textile industry are time series forecasting methods, for example: exponential smoothing [47], Holt Winters model [48], Box & Jenkins model [49], regression models [50] or ARIMA (AutoRegressive Integrated Moving Average). Commercial softwares have been implemented to apply these techniques; the results (in some instances) are satisfactory, according to [51]. Despite the existence of these softwares, they are seldom used in textile-apparel field. If they are used, companies use them only as a basis for their forecasting systems.
Most of textile companies use more advanced techniques to increase the accuracy of sales forecast. These advanced techniques can be divided in two groups:

1. **forecasting methods with historical data.** Among these techniques, NN (Neural Networks) are the most used, especially for short-term forecast, where the main problem is to be reactive to the last known sales [52]. If the demand of the products is not seasonal and non variable this method performs well. Thus, it's not suitable for the textile environment. Other techniques for sales forecasting have been implemented successfully in textile industry, such as Fuzzy logic and Fuzzy Inference Systems (FIS), that are commonly used to model uncertain data.

2. **forecasting methods without historical data.**, that is the case when the products are sold during only one season and it's not possible to have historical data, since it's a new product. New product forecasting is one of the most difficult forecasting problem, according to [54]. Some methods have been used to cope with this situation; the most used consists of a two-step method:

   (a) Cluster and classify new products to forecast their sales profile (mid-term forecast);

(b) Adapt and readjust this profile according to the first periods of sales (short-term forecast).

If no historical data are available, similar products information sold in previous seasons can be used. So it's important to take into account some descriptive attributes (price, life span, sales period, style, ...) of historical and new products, with the aim to model a relationship between these attributes and the sales, so that it's possible to use these relationships to forecast future sales for new products with some of the same attributes of old products. This methodology has been implemented with good results for the fashion industry in [55] and [56].

The aim of this research is to provide a solution for the reactive modification of production schedules. To do this, two ways of solution can be taken:

1. improvement of the forecasting methodology, in order to reduce the uncertainty, so that deterministic demand (known) is considered in the models;
2. consideration of the uncertainty of demand in the model. Demand in this context is not deterministic but stochastic, meaning that statistical parameters (such as mean, standard deviation , ...) are associated to demand. In other words, statistical dimension is within the programming model, and it is strictly related to scheduling decisions.

In our work we are not interested in finding the best accurate forecasting method for the textile environment we are considering. We simply assume that textile environment is such uncertain that this uncertainty cannot be overcome by optimal forecasting systems. Thus, in this work we take into account uncertanty, using a stochastic model.

4.3 Definition of SKUs

The SKU (Stock Keeping Unit) for our instance is the fabric. In this paragraph we describe how the SKUs in the models have been divided and classified.
End-products in textile field can be analyzed according to different classifications

[53]. From the point of view of their permanence in the collections (business classification) we can distinguish:

- **fashion items**: items sold in a particular season, with a very short life cycle. They are temporary items, related to haute couture.

- **basic items** (Carry Over or Never Out of Stock items): items sold in more seasons; their availability in stocks has to be guaranteed with frequent replenishments and the preparation of proper safety stocks.

According to the traditional classification of revenue based on the ABC Pareto chart, items can be divided in:

- A class (high sales)

- B class (medium sales)

- C class (low sales)

From the point of view of the ease of forecasting from the planners, products can be divided in:

- regular items (multi-season). Historical data are available. They can present seasonalities or trends;

- sporadic items (multi-season). Sales are not frequent and they are variable in terms of quantity (for example: luxury items);

- periodic items (multi-season). Items with sales concentrated in a small number of weeks, periodically over the years.

- mono-season.

We choose to handle with four types of SKUs of different material (we could choose also different blends of fabric: in the figures (4.2) and (4.3) a classification of the fibres from which the fabrics are produced is presented) and we consider for them four different demand patterns, which represent the expected value function of the demand. The products are basic multi-season items; two of them have stationary demand, while the other two present a periodic pattern.

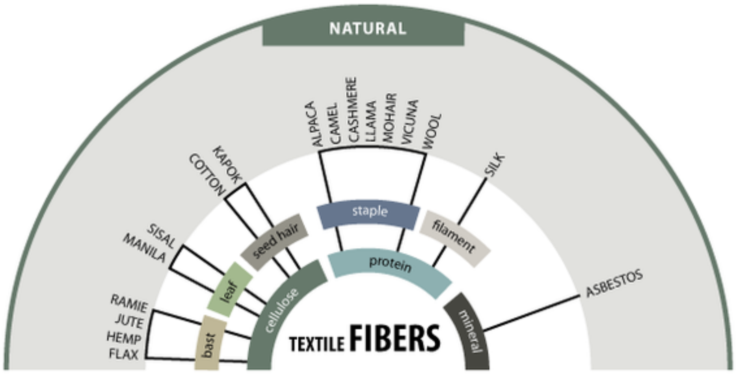


Figura 4.2: Natural textile fibres

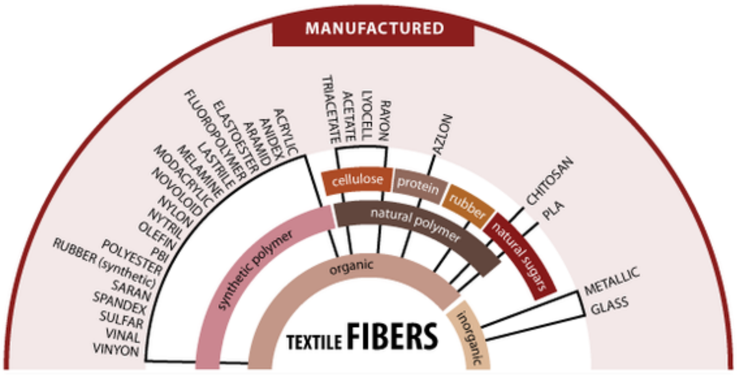


Figura 4.3: Manufactured textile fibers

SKU	Type	Demand	Material and use
SKU_1	Regular item	Stationary	Cotton for home furnishing
SKU_2	Regular item	Stationary	Silk for multi-season clothing
SKU_3	Periodic item	Seasonal	Linen for summer clothing
SKU_4	Periodic item	Seasonal	Polyesther for winter clothing

Comments:

- The looms we are going to consider can process each type of SKU chosen.
- A changeover occurs each time that a loom needs to produce a different SKU, because it has a different material.
- In the real context a certain fabric can have two or more different demand patterns due to the fact that it has different uses (for example: cotton for

sheets has a stationary demand pattern, while cotton for summer clothing has a seasonal demand pattern). In this case all the patterns of the same type of fabric have to be summed together.

4.4 Deterministic model for the fabric formation scheduling problem

In this section we present the deterministic model, developed in a general way in the previous chapters, for this particular instance. The schedule suggested by this model will be compared with the one suggested by the stochastic model.

4.4.1 Terms of the model

The terms of the model are described in the table below.

Term	**Description**
P	Number of types of SKU: $j \in \{1, 2, ..., P\}$
T	Number of weeks of the production schedule: $t \in \{1, 2, ..., T\}$
L	Number of looms (capacity of the plant)
$Y[j,t]$	Number of looms that are configured to process SKU j in period t
$Z[j,t]$	Number of looms that process product j in week t
$\delta[j,t]$	Number of changeovers to process SKU j in week t
$I[j,t]$	Inventory amount for SKU j in week t
$B[j,t]$	Backorder amount for SKU j in week t
$d[j,t]$	Expected demand for SKU j in week t
$Y[j,0]$	Machines configured to process SKU j in week $t = 0$
$I[j,0]$	Inventory amount for the SKU j in period $t = 0$
$B[j,0]$	Backlog amount for the SKU j in period $t = 0$
$c_Y[j]$	Setup costs for SKU j
$c_I[j]$	Inventory carrying costs for SKU j
BF	Backorder Factor

Also here the values of the demand are expressed as the number of looms that need to run for SKU j in week t. Regarding to this, some considerations are made in chapter 2.

Indeed we assume that we can turn the demand values (expressed in quantity of product j requested in week t) into number of machines to run in week t for SKU j. As we said in chapter 2, this way to express the demand simplifies significatively the description so that we don't need to make any considerations about production rates of the machines.

As a consequence of this, all the other variables are expressed in this way. For example, the Z variable doesn't represent a quantity of product, but the number of looms that need to run for a certain product in a specific period.
The same applies to the inventory and backorder, that are expressed as the number of looms that need to run in a given period to produce certain amounts for a specific product.

Regarding the variables type, all the considerations made in chapter 2 are valid. Thus we have also in this case:

Variable	**Type**
Y	Integer
Z	Continuous
δ	Integer
I	Continuous
B	Continuous

4.4.2 Objective function and constraints

The model involves two objective functions. As we well described in chapter 2, the first one provides the second with important information about feasibility in terms of backorder.
Thus, the purpose of the first objective function is to find the minimal possible backorder amount:

$$\text{minimize:} \quad \sum_{j=1}^{P} B[j,4] \tag{4.1}$$

The first minimization is about the total backorder amount of all the SKU at the end of the fourth week.

There are four constraints related to the first objective function, and they are explained below.
Constraints (4.2) impose the total looms availability limits. Constraints (4.3) measure the positive changes in the number of looms dedicated to a SKU from one week to the following one. This, together with (4.2) ensures that changeovers are measured correctly.

$$\sum_{j=1}^{P} Y[j,t] = L \quad \forall t \in 1, ..., \{1,2,3,4\} \tag{4.2}$$

$$Y[j,t] - Y[j,t-1] - \delta[j,t] \leq 0 \quad \forall j \in \{1,2,3,4\}, \forall t \in \{1,2,3,4\} \tag{4.3}$$

Constraints (4.4) are inventory balance equations for each SKU.

$$Z[j,t-1]+I[j,t-1]-B[j,t-1]-I[j,t]+B[j,t] = d[j,t] \quad \forall j \in \{1,2,3,4\}, \forall t \in \{1,2,3,4\} \tag{4.4}$$

The total number of looms that are configured for SKU j limits the total number of looms that can be processed to produce it. This is ensured by:

$$Z[j,t] \leq Y[j,t] \quad \forall j \in \{1,2,3,4\}, \forall t \in \{1,2,3,4\} \tag{4.5}$$

From this first minimization we get some important values that we use in one of the constraints of the second objective function. We define

$$B_{\text{MIN}}[j,T] = B[j,4] \quad \forall j \in \{1,2,3,4\}$$

where $B[j,T]$ comes from the solution of the first optimization.

In the second part of the model we find the second objective function:

$$\text{minimize:} \sum_{j=1}^{P}\sum_{t=1}^{4} c_Y \cdot \delta[j,t] + \sum_{j=1}^{P}\sum_{t=1}^{4} c_I \cdot I[j,t] \tag{4.6}$$

This second objective function is about minimizing the total of changeover costs and inventory carrying costs over the scheduling horizon. Backorders are not considered as a cost in the objective function because we assume that it's not possible to assign a monetary value to them. However, backorder is allowed and it is handled as a constraint in terms of quantities.
The constraints for this objective function are the same of the first one. Therefore, there is one additional contraint about a limitation of the backorder amount. Before showing this constraint we recall the Backorder factor (BF), described introduced in chapter 2. It is an input parameter that represents how much backorder is allowed for each product with respect to the total demand over the periods of the horizon, with:

$$0 < BF < 1.$$

The percentage of demand satisfaction can be defined as Service Factor (SF):

$$SF = 1 - BF.$$

Now we can introduce the last constraint, that is:

$$B[j,4] \leq B_{\text{MIN}}[j,T] + \text{BF} \cdot \left(\sum_{t=1}^{4} d[j,t]\right) \quad \forall j \in \{1,2,3,4\} \tag{4.7}$$

4.5 Stochastic approximated model for the fabric formation scheduling problem

In this section we present the stochastic approximated model, developed in a general way in the final section of chapter 3, for this particular instance.

4.5.1 Terms of the model

Term	Description
P	Number of types of SKU: $j \in \{1, 2, ..., P\}$
T	Number of weeks of the production schedule: $t \in \{1, 2, 3, 4\}$
L	Number of looms (capacity of the plant)
O	Set of the o outcomes (scenarios): $o \in O$
$Y[t, j]$	Looms, in week t, that are configured to process SKU j
$Z[t, j]$	Looms, in week t, that produce SKU j
$\delta[t, j]$	Number of changeovers, in week t, to produce SKU j
$I[t, j, o]$	Inventory amount in week t for SKU j under outcome o
$B[t, j, o]$	Backlog amount in week t for SKU j under outcome o
$\text{LB}[t, j, o]$	Lower bound of the demand interval
$\text{d}[t, j]$	Demand forecast for week t for SKU j
$\text{UB}[t, j, o]$	Upper bound of the demand interval
$Y[0, j]$	Machines configured to process SKU j in week $t = 0$
$I[0, j]$	Inventory amount for SKU j in week $t = 0$
$B[0, j]$	Backlog amount for SKU j in week $t = 0$
$c_Y[j]$	Setup costs
$c_I[j]$	Inventory carrying costs for SKU j
$p[o]$	Probability of the outcomes
BF	Backorder Factor

Demand values are expressed as the numbers of machines that need to produce for SKU j in week t. With regard to this, all the considerations are made in chapter 3 where we describe the generalized stochastic approximated model.
As a consequence of this, all the other variables are expressed in this way. For example, the Z variable doesn't represent a quantity of product, but the number

of machines that need to run for a certain SKU in a specific week.
The same applies to the inventory and backorder, that are expressed as the number of looms that need to run in a given week to produce certain amounts for a specific product.

The type of the variables involved in this model are the same of the generalized model described in chapter 3. The table below recalls the variables type:

Variable	**Type**
Y	Integer
Z	Continuous
δ	Integer
I	Continuous
B	Continuous

4.5.2 Objective function and constraints

The objective functions and the constraints are the same of those one involved in the generalized model described in chapter 3.

The purpose of the first objective function is to find the minimal possible backorder amount:

$$\text{minimize: } \sum_{j=1}^{P} \mathbb{E}[B[4, j]] \tag{4.8}$$

It's interesting to notice that this minimization is about the expected value of the total backorder amount of all the products at the end of the fourth period.

There are four constraints related to the first objective function, and they are explained below.
Constraints (4.9) impose the total looms availability limits. Constraints (4.10) measure the positive changes in the number of looms dedicated to a SKUt from one week to the following one. This, together with (4.9) ensures that changeovers are measured correctly.

$$\sum_{j=1}^{P} Y[t, j] = L \quad \forall t \in \{1, 2, 3, 4\} \tag{4.9}$$

$$Y[t,j] - Y[t-1,j] - \delta[t,j] \leq 0 \quad \forall j \in \{1,2,3,4\}, \forall t \in \{1,2,3,4\} \tag{4.10}$$

Constraints (4.11) and (4.12) are inventory balance equations for each SKU.

$$Z[t,j] + \mathbb{E}[I[t-1,j]] - \mathbb{E}[B[t-1,j]] - I[t,j,o] + B[t,j,o] \geq \text{UB}[t,j,o] \tag{4.11}$$

$$\forall j \in \{1,2,...,P\}, \forall t \in \{1,2,3,4\}$$

$$Z[t,j] + \mathbb{E}[I[t-1,j]] - \mathbb{E}[B[t-1,j]] - I[t,j,o] + B[t,j,o] \leq \text{LB}[t,j,o] \tag{4.12}$$

$$\forall j \in \{1,2,...,P\}, \forall t \in \{1,2,3,4\}$$

The total number of looms that are configured for product j limits the total number of looms that can be operated to produce it. This is ensured by:

$$Z[t,j] \leq Y[t,j] \quad \forall j \in \{1,2,...,P\}, \forall t \in \{1,2,3,4\} \tag{4.13}$$

From this first minimization we get some important values that we use in one of the constraints of the second objective function. We define

$$\mathbb{E}[B_{\text{MIN}}[T,j]] = \mathbb{E}[B[T,j]] \quad \forall j \in \{1,2,3,4\}$$

where $\mathbb{E}[B[T,j]]$ comes from the solution of the first optimization.

In the second part of the model we find the second objective function:

$$\text{minimize:} \ \sum_{j=1}^{P} \sum_{t \in 4} c_Y \cdot \delta[t,j] + \cdot \sum_{j=1}^{P} \sum_{t \in 4} c_I \cdot p[o] \cdot I[t,j,o] \tag{4.14}$$

This second objective function is about minimizing the total of changeover costs and inventory carrying costs over the four weeks. Backorders are not considered as a cost in the objective function because we assume that it's not possible to assign a monetary value to them. However, backorder is allowed and it is handled as a constraint in terms of quantities.

The constraints for this objective function are the same of the first one. Therefore, there is one additional contraint about a limitation of the backorder amount.

$$\mathbb{E}[B[t,j]] \leq E[B_{\text{MIN}}[4,j]] + \text{BF} \cdot \left(\sum_{t=1}^{4} d[t,j] \right) \quad \forall j \in \{1,2,...,P\} \tag{4.15}$$

where BF is the Backorder Factor introduced previously in the deterministic model.

4.6 Rolling Horizon approach: scheduling storyline

In our specific context we can describe a storyline of what happens into the reality. On friday the planner has the information about conditions of inventory and backorders, and we assume that he has available forecasts for the following four weeks. He has to decide which and how many changeovers to make on the weekend (outside of the working time), and how much to produce for each product the following week. He realizes an optimal plan minimizing the total cost function over the following four weeks considering the forecasts. In many lot-sizing and scheduling problems, production is planned using the rolling horizon procedure. Essentially companies realize plans based on unreliable forecasts for a certain number of periods, but they make new plans after one or more periods without realizing completely the original plan.
They normally use this procedure making new plans every period in order to be reactive to the market as much as possible, using new information, when available. In other words, they realize plans for a certain number of periods, but only the decisions related to the first one are implemented. In (4.4) an example of how the rolling horizon procedure can be applied to our fabric formation problem is shown. In this particular case we have three plans generated within an horizon of two weeks, and every week a new plan is made. Starting on friday, the planner decides changeovers and quantities to be produced for both the following periods. On following friday he observes $I0$ and $B0$ (actual values of inventory and backorder) and makes a new plan, and so on.
The rolling horizon procedure applied to our case can be sketched like in (4.5).

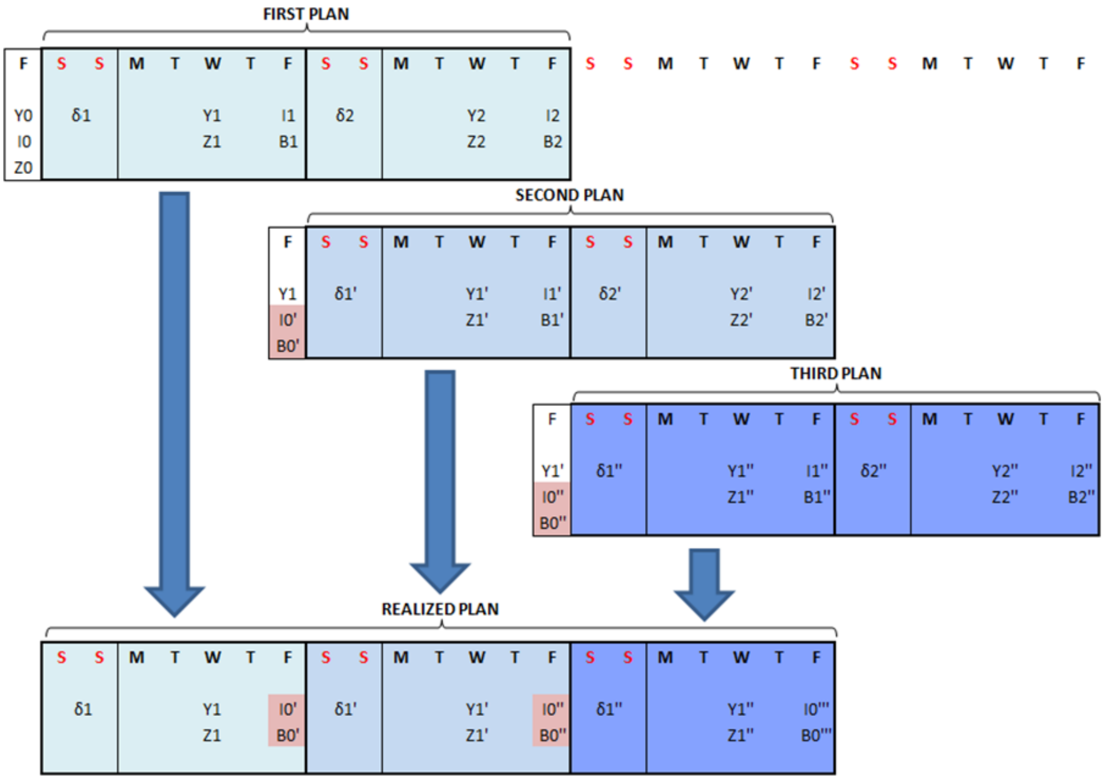


Figura 4.4: Rolling Horizon approach: variables involved

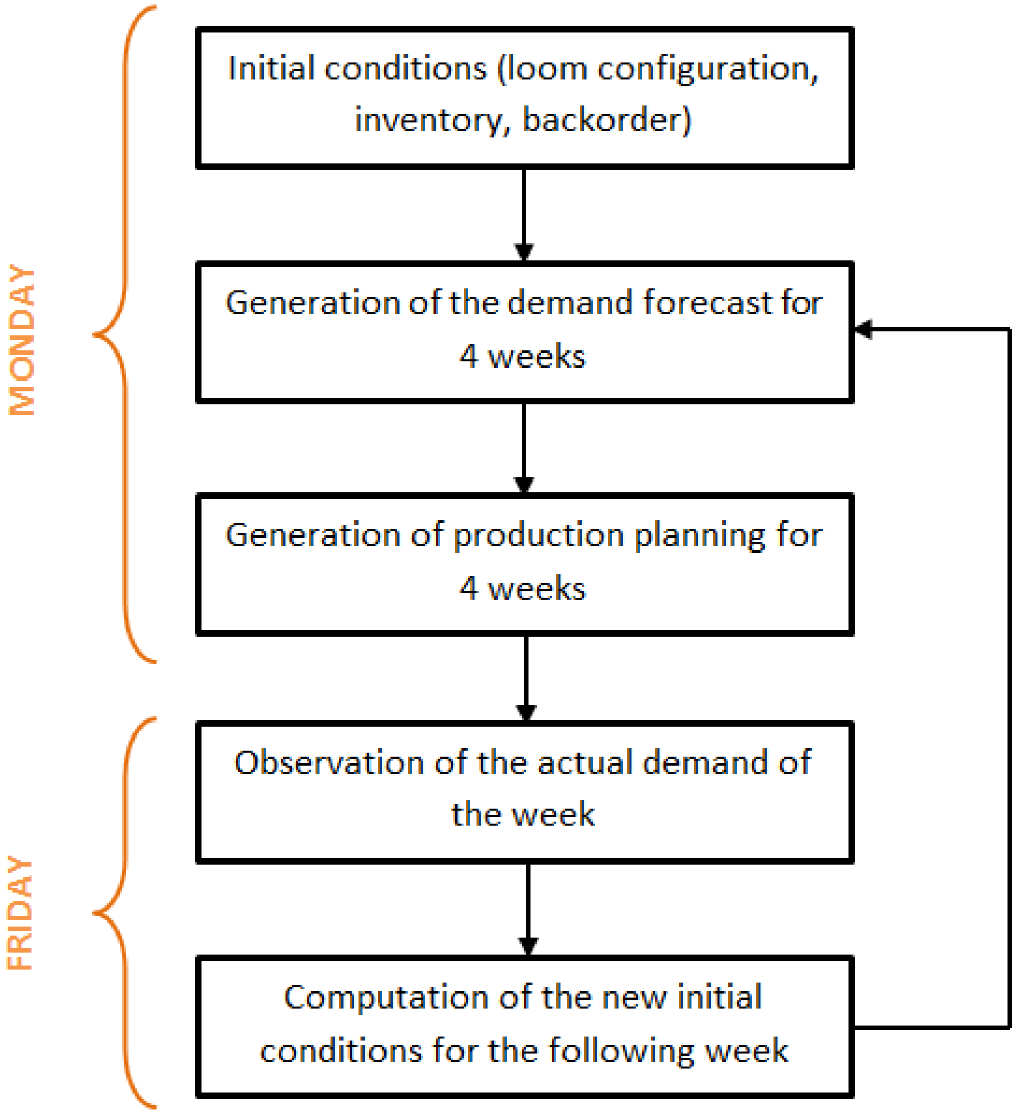


Figura 4.5: Rolling Horizon approach: application in the case of study

Capitolo 5

Preliminary experiments

Using the models developed in the previous chapter about the fabric formation scheduling problem, we conduct a series of preliminary experiments in order to get a general idea of how these models work, and try to understand the main features of the experiments we want to perform later in the dissertation.

These preliminary experiments are called static setting experiments; in these experiments we compare the behavior of the production schedules of the deterministic and the stochastic model in a certain period of time (one month, characterized by the same demand forecasts), without modifying the schedules when new information about the actual observation of demand is available, meaning not applying a rolling horizon approach. The models will produce different configurations of the looms and different number and timing of changeovers within the month.
Considering a set of a very large number of scenarios of the actual demand for each week of the month and for each SKU (same scenarios for both models), we can compute how much inventory and backorder would be produced with the schedules suggested by the models that cannot be changed over time. From this set of experiments we want to get some general idea of the difference between the models we are testing: the main variables we are interested in are the number of changeovers, inventory and backorder amount, number of looms running every week.

This chapter is organized as follows: the first paragraph is about the input data considered in the experiments (both preliminary experiments and the ones in chapter 6), while the second is about the static setting experiments (description of the experiments, additional input data for static experiments and results).

5.1 Data setting

In this section we describe the input data of the experiments. Previously we described our models but we didn't set the necessary parameters. Of course, experiments results will be affected by the parameters setting of the simulation environment. Indeed, it's necessary to proceed with an accurate and precise discussion about all the parameters involved in our models.

First of all, we decide the length of the scheduling horizon. Textile business uncertainty let us know demand forecasts not for a long horizon, that we assume to be one month; since, as we said in chapter 4, scheduling decisions have to be made weekly, the models we are going to consider become four-stages models ($T = 4$ weeks).

As we said before in chapter 4, we decided to study a products mix composed by 4 SKUs:

SKU	**Type**	**Demand**	**Material and use**
SKU_1	Regular item	Stationary	Cotton for home furnishing
SKU_2	Regular item	Stationary	Silk for multi-season clothing
SKU_3	Periodic item	Seasonal	Linen for summer clothing
SKU_4	Periodic item	Seasonal	Polyesther for winter clothing

In order to describe these SKUs we need to assign a definite expected value function of the demand forecast and a level of uncertainty for each SKU, expressed by the parameter CV (Coefficient of Variation), defined as:

$$CV[j] = \frac{\sigma[j]}{\mu[j]}$$

with $j \in P$.
We assume that a stationary demand and a seasonal demand can be described

respectively by a constant function and a sine function. It's not our aim to handle with the closest to reality demand patterns, we just want to create a simple environment in which we have different types of demand pattern. We chose for each SKU a mean value of the expected demand function and a coefficient of variation, as we can see in the table below:

SKU	**Expected demand function**	**CV**
SKU_1	$E_1(t) = 100$	$CV_1 = 0.1$
SKU_2	$E_2(t) = 30$	$CV_2 = 0.2$
SKU_3	$E_3(t) = 50[1 + \sin(\omega t + \phi_3)]$	$CV_3 = 0.4$
SKU_4	$E_4(t) = 35[1 + \sin(\omega t + \phi_4)]$	$CV_4 = 0.3$

As regards SKU_3 and SKU_4 we have to define the angular frequencies and the phases. We assume that both the SKUs have the same period of 52 weeks (one year).

$$\omega = \frac{2\pi}{T} = \frac{2\pi}{52} \left[\frac{1}{\text{week}}\right]$$

Considering the phases:

$$\phi_3 = \frac{3}{2}\pi + \frac{2\pi}{52} \cdot (\text{IP} + \xi)$$

$$\phi_4 = \frac{3}{2}\pi + \frac{2\pi}{52} \cdot \text{IP}$$

The term ξ represents the phase shift (in weeks) between SKU_3 and SKU_4, that are seasonal SKUs (respectively summer and winter clothing). We assume that this phase shift is 15 weeks, so $\xi = 15$. IP (Initial Position) is a parameter that moves the sine functions over time and it is expressed in number of weeks. For example, with $IP = 0$ and $\xi = 15$, the expected demand functions for the SKUs are:

As regards costs, the cost of a single changeover doesn't depend on the sequence of the SKUs; this data has been provided by Milliken & Co. About inventory carrying costs, it is reasonable to compute an inventory carrying cost on the basis of the production cost and the IRR (Internal Rate of Return) as follows:

$$c_I = \frac{\text{Production Cost} \cdot \text{IRR}}{52} \quad \left[\frac{\$}{\text{looms} \cdot \text{week}}\right]$$

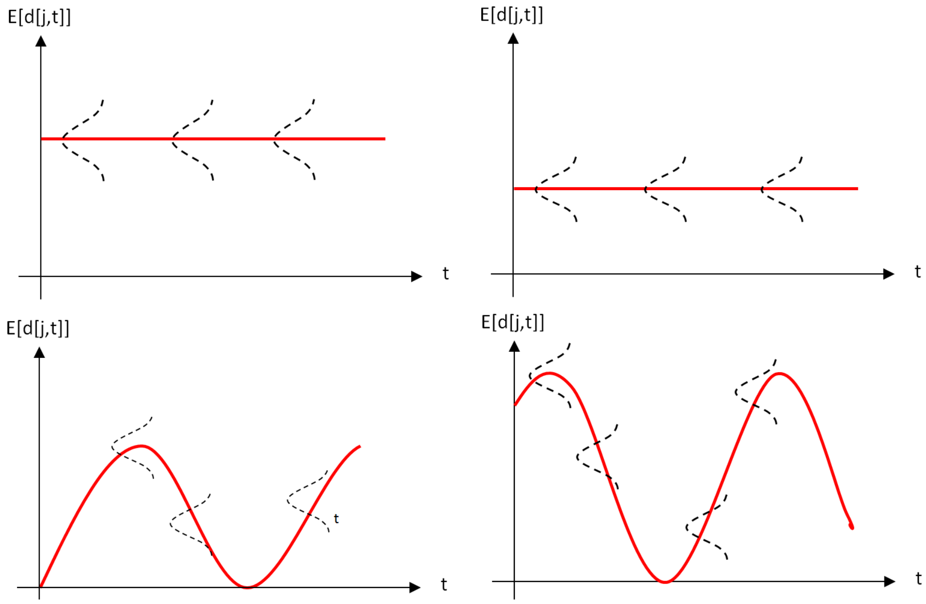


Figura 5.1: Expected demand functions for the single SKUs

where production cost is assumed to be equal to 150 $\frac{\$}{\text{looms week}}$ and it doesn't depend from the SKU type (this data has been provided by Milliken & Co.), while IRR is supposed to be equal to 20%. IRR is set quite high in order to give more weight to the inventory cost in the objective function.

$$c_I = \frac{150 \cdot 0.2}{52} = 0.5769 \left[\frac{\$}{\text{looms} \cdot \text{week}}\right]$$

In summary:

Setup costs	$c_Y = 500 \frac{\$}{\text{setup}}$
Inventory carrying costs	$c_I = 0.5769 \frac{\$}{\text{looms} \cdot \text{week}}$

5.2 Static setting experiments

5.2.1 Description of the experiments

As we said before, the first set of experiments are called static setting experiments. Static means that the production schedules realized by the models for one month ($T = 4$) don't change when new information about actual demand is available; in other words, a rolling horizon approach is not applied. The models develop their schedules according to the same demand forecasts $d[j, t]$. The models we want to analyze are:

- deterministic model;
- stochastic model 3 intervals (with three intervals of demand for each SKU);
- stochastic model 5 intervals (with five intervals of demand for each SKU).

The main steps followed to perform these experiments are:

1. **Solution of the models according to the forecasts**. The information we want to get are $Y[j, t]$ (looms configured for SKU j in week t), $Z[j, t]$ (looms working for each SKU and for each week) and $\delta[j, t]$ (changeovers due for SKU j in week t). We want to discover especially if a stochastic approach suggests a reduced number of changeovers.
   The other variables ($I[j, t]$ and $B[j, t]$) are computed too, but we are not interested in them, because they represent inventory and backorder amount only if the actual demand is equal to the forecasts;

2. **Generation of a large number of scenarios (outcomes)**. A scenario is a set of values of actual demand for each SKU and for each of the four weeks of the month. Actual demand values are generated randomly starting from a normal distribution with the value of the forecast as mean, and a standard deviation evaluated on the basis of the CV that we previously set. In the figure below (5.2) we can see an example of scenario of actual demand, with values that sometimes are very different from forecasts because of the CVs. These information, as we said before, don't interfere with the schedules

DEMAND FORECAST				
Products/periods	1	2	3	4
SKU1	100	100	100	100
SKU2	30	30	30	30
SKU3	93	90	85	80
SKU4	40	44	48	53

ACTUAL DEMAND				
Products/Periods	1	2	3	4
SKU_1	95.4	93.96	115.12	108.23
SKU_2	35.46	29.09	24.78	35.92
SKU_3	102.16	112.36	76.31	99.14
SKU_4	49.55	48.07	49.19	39.66

Figura 5.2: Forecast and actual demand values for the considered periods

provided by the models (the schedule is freezed in terms of configuration of the looms). This large number of outcomes is stored in a database and contains 100000 scenarios. One at a time will be used for all the models (see the following step).

3. **Computation of the actual values of $I[j,t]$ and $B[j,t]$.** We compute the real values of these variables, considering the schedule suggested by each model, so that we can see what are the differences among the models regarding inventory and backorder amount.

From this set of experiments we want to get some general idea of the difference between the models we are testing.

In these experiments we are going to see the difference between the models in different conditions of demand with respect to capacity of the plant. The conditions we are going to study are:

- average demand forecast lower than capacity
- average demand forecast equal to capacity

- average demand forecast higher than capacity

5.2.2 Additional input data for the static experiments

In order to perform these experiments, we still need to define the number of looms for each instance (and so the percentage of the average demand with respect to capacity):

- in the instance of average demand forecast lower than capacity [85%], the number of the looms is 225.
- in the instance of average demand forecast equal to capacity [100%], the number of the looms is 265.
- in the instance of average demand forecast higher than capacity [115%], the number of the looms is 305.

We decide to change the number of the looms (capacity of the plant) and not the demand, so that we use the same scenario database in every run. Since sinusoidal

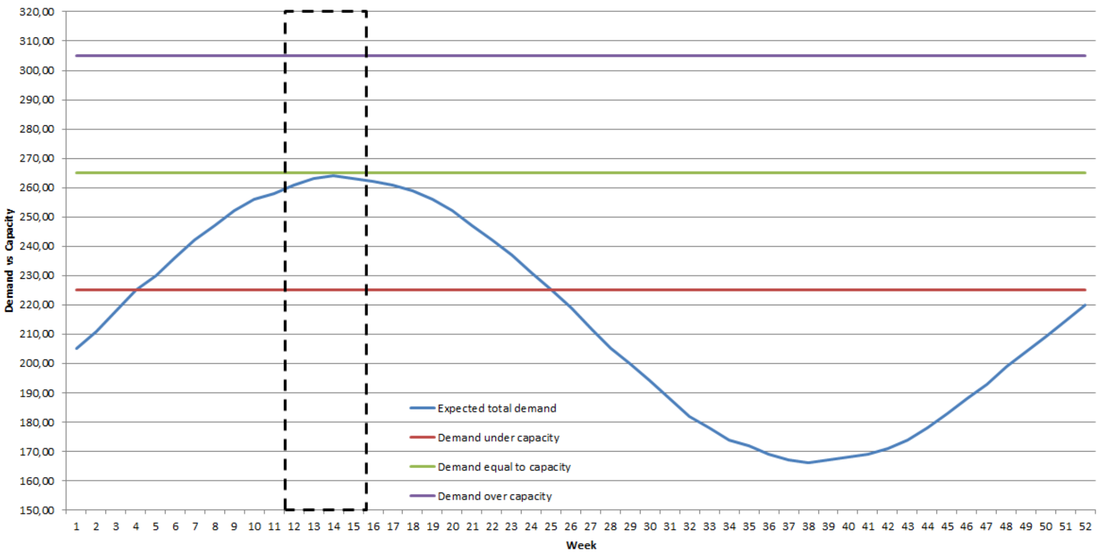


Figura 5.3: Different conditions analyzed

demand functions are involved in the environment, we need to choose, within the year, a particular period we want to study in this set of experiments. We choose

a certain period of four weeks, underlined in the figure, setting the parameter IP = 12, explained above.
In this period we note that the SKU_3 expected demand function is decreasing, while the SKU_4 is increasing. The other two SKUs have stationary demand.

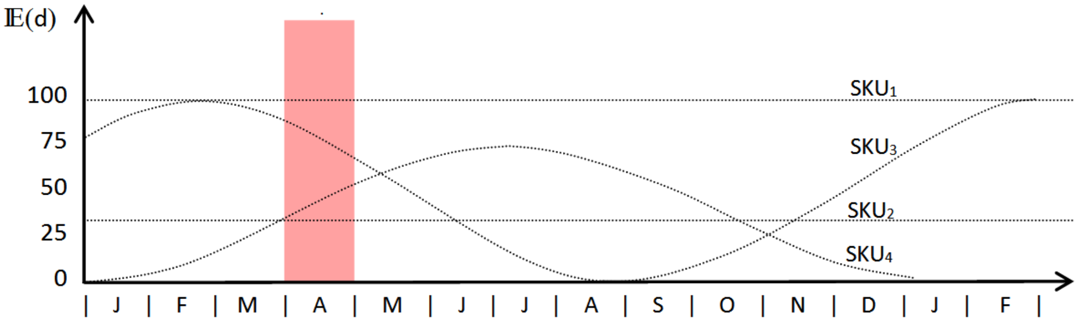


Figura 5.4: Period considered for the static setting experiments

5.2.3 Results

Average demand forecast lower than capacity [85%]

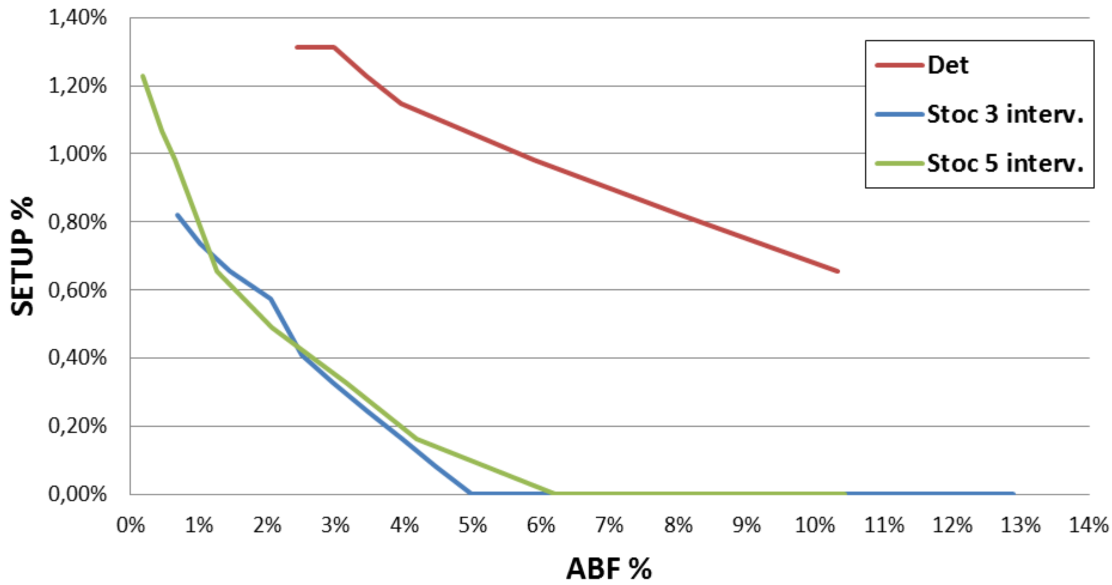


Figura 5.5: Demand lower than capacity: setup

We compute SETUP [%] varying BF (Backorder Factor) or, in other words, relaxing the service level constraint previously defined. In figure (5.5) are represented:

- on the x-axis → ABF [%] (Actual Backorder Factor), defined as:

$$\text{ABF} = \frac{\text{Cumulated actual backorder in 4th week}}{\text{Total demand of the 4 weeks}}$$

- on the y-axis → SETUP [%], defined as:

$$\text{SETUP [\%]} = \frac{\text{Number of changeovers}}{\text{4 weeks} \cdot \text{Number of looms}}$$

The information that can be discovered from this plot are:

- for each ABF we see that the stochastic models suggest a smaller number of changeovers with respect to the deterministic one (this fact leads to a significative reduction of costs). For example, with $ABF = 4\%$:

$$\text{DET : Number of changeovers} = \frac{1.15}{100} \cdot 265 \cdot 4 \sim 13 \text{ Changeovers}$$

$$\text{STOC : Number of changeovers} = \frac{0.2}{100} \cdot 265 \cdot 4 \sim 3 \text{ Changeovers}$$

- the minimum ABF [%] is obtainable with the stochastic models. This means that with a deterministic approach we reach a lower service level;

- there is no a big difference between the two different partitions of the demand. This fact, since computational time is higher in the stochastic model with 5 intervals, let us to consider just the 3 intervals partition for the dynamic setting experiments, without losing a lot of advantages in cost reduction (from now cost performance) and service level (from now service performance).

In this condition, we show other two plots (5.6) and (5.7). The first one is about inventory:

- x-axis → ABF [%] (Actual Backorder Factor), defined as before.

- y-axis → INVENTORY [%], defined as:

$$\text{INVENTORY [\%]} = \frac{\text{Total inventory amount in the 4 weeks}}{\text{Total demand of the 4 weeks}}$$

In the first plot we can see that, for each ABF, the average weekly inventory is lower in the stochastic case. With a stochastic approach (either 3 or 5 intervals of demand for each SKU in each week) it's possible to reach higher performances in terms of service level.

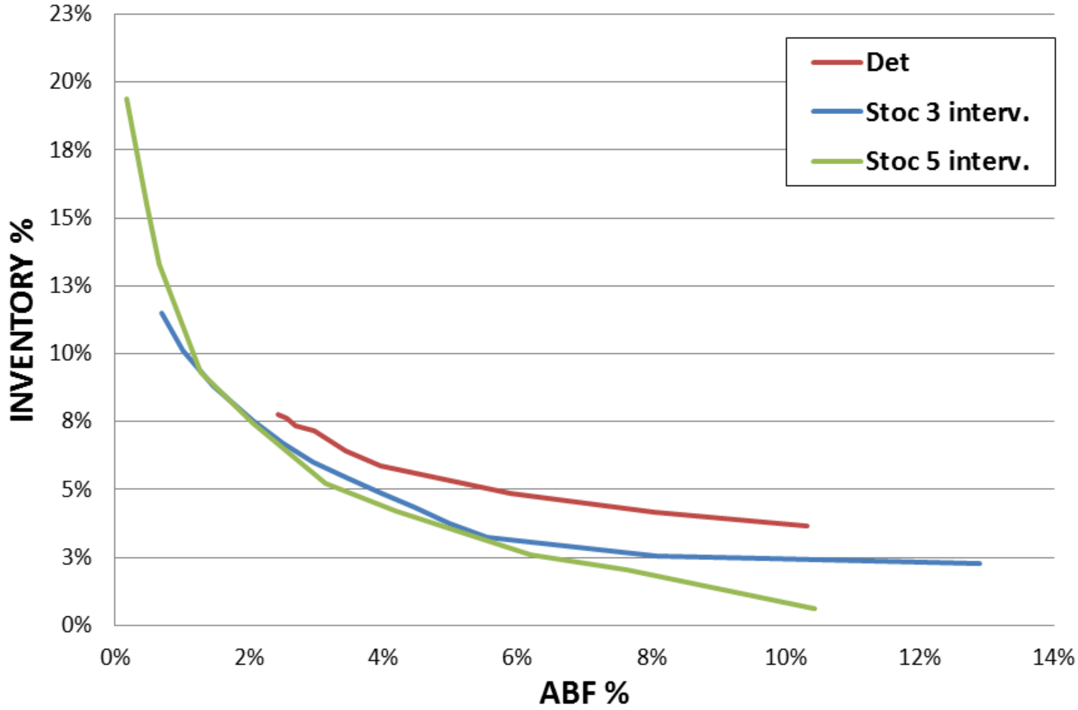


Figura 5.6: Demand lower than capacity: inventory

Consider the figure (5.7):

- x-axis → ABF [%] (Actual Backorder Factor), defined as before.
- y-axis → PRODUCTION [%], defined as:

$$\text{PRODUCTION [\%]} = \frac{\text{Looms running in the 4 weeks}}{\text{4 weeks} \cdot \text{Number of looms}}$$

As we can see, the number of looms that run in average is lower using a stochastic approach.

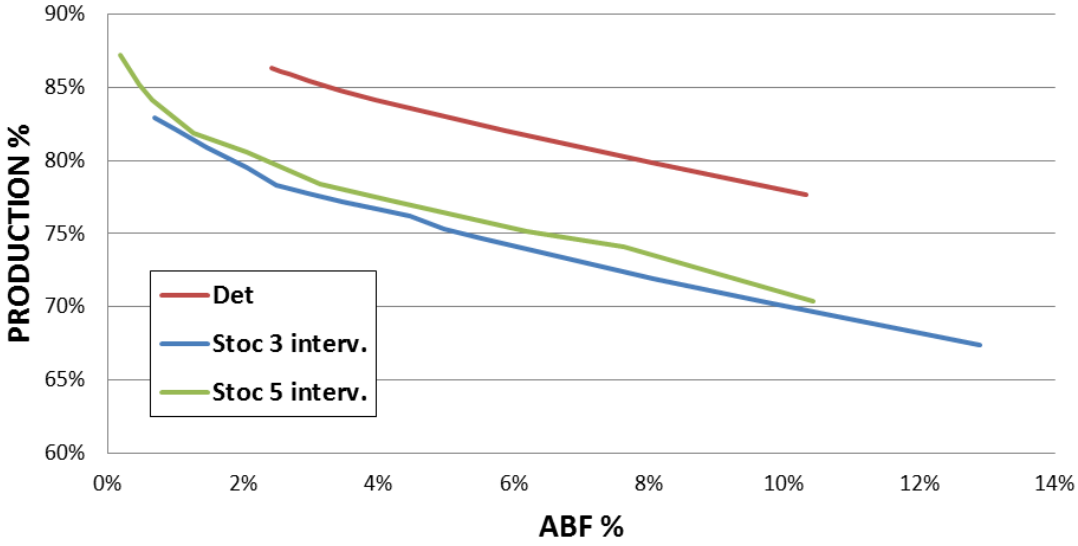


Figura 5.7: Demand lower than capacity: production

Average demand forecast equal to capacity [100%]

We report the same plots in this condition. We can get the same information of the first instance, where the average demand forecast was lower than the capacity.

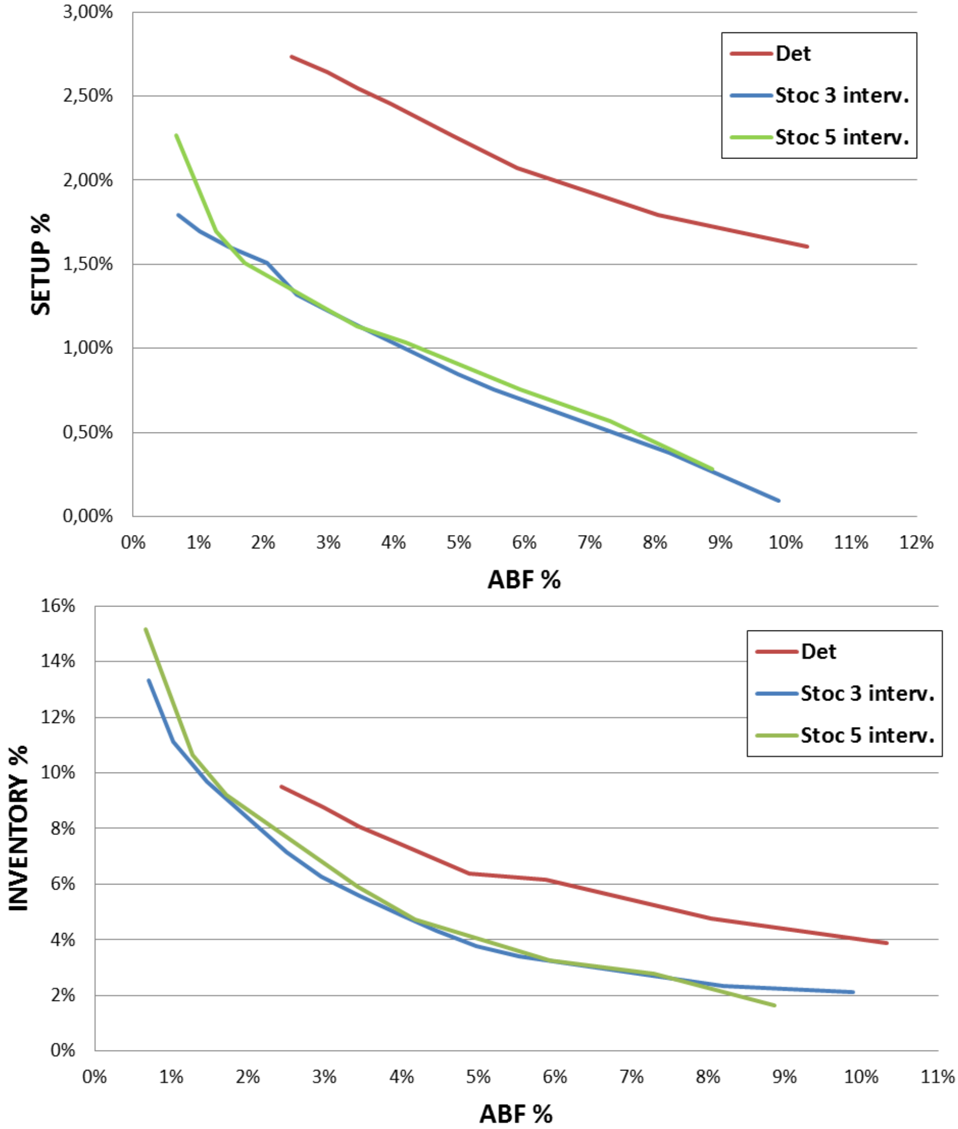


Figura 5.8: Demand equal to capacity: setup and inventory

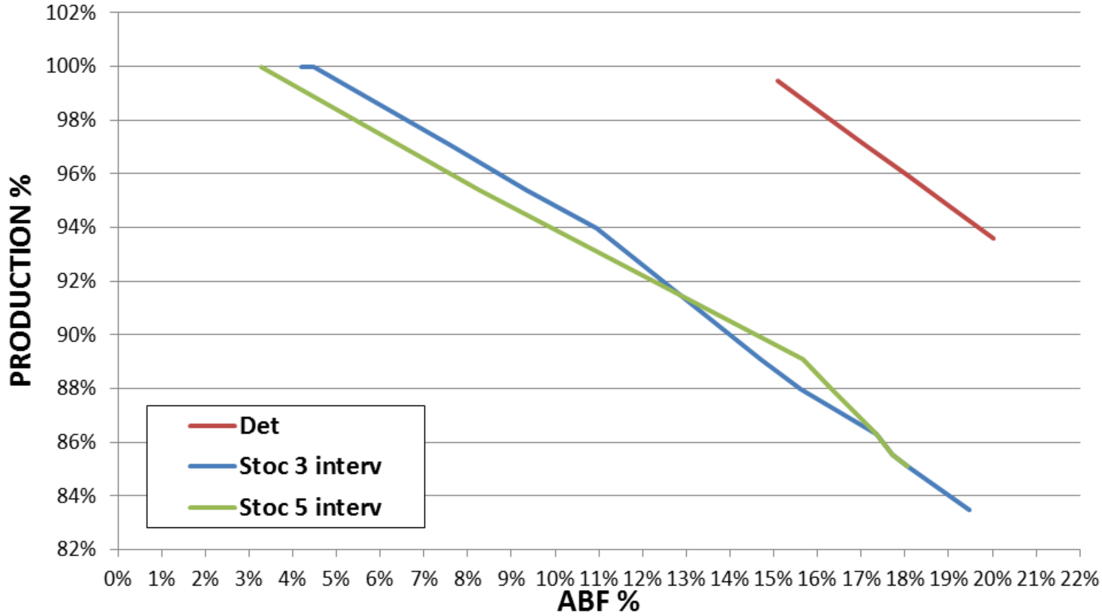
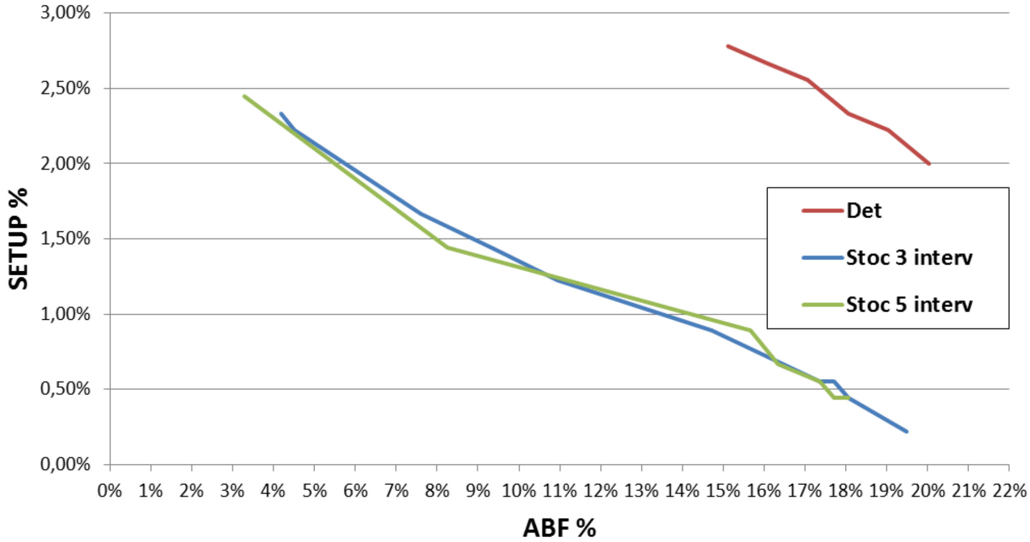


Figura 5.9: Demand equal to capacity: inventory and production

Average demand forecast higher than capacity [115%]

We also report the plots related to the last instance, making the same conclusions as before.



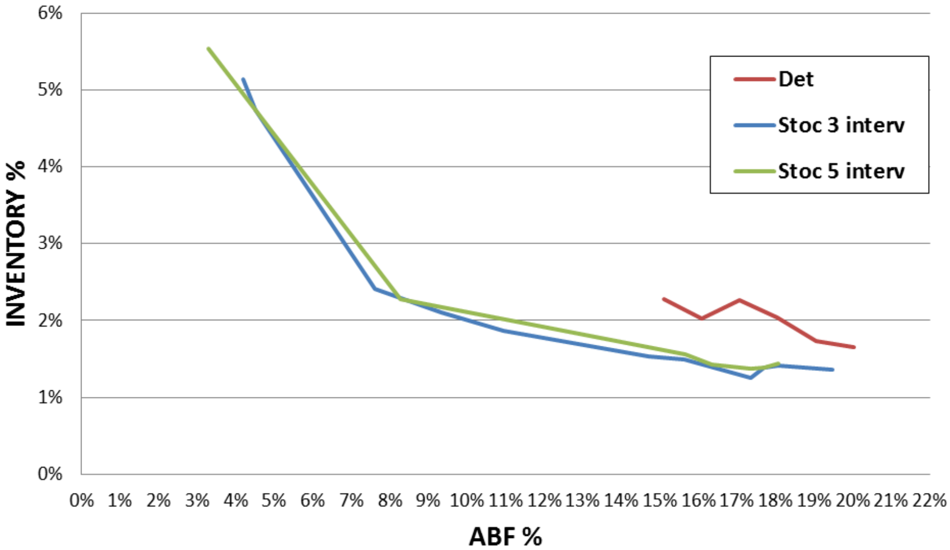


Figura 5.10: Demand higher than capacity: setup and inventory

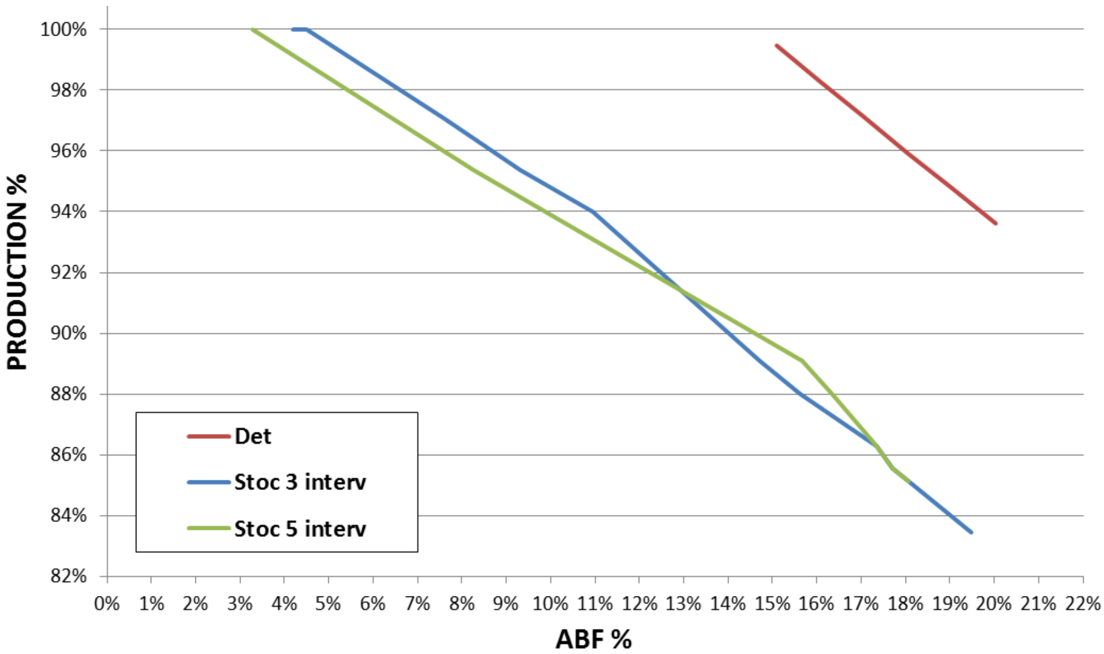


Figura 5.11: Demand higher than capacity: production

In conclusion, we can say that:

- stochastic models provide better results with respect to the deterministic model in terms of objective function (cost reduction) and performance (service level);
- there is no a big difference of results between the 3 intervals and 5 intervals stochastic models;
- the advantages obtained from stochastic models are obtainable in every condition of demand with respect to capacity;
- as we can see in (5.12) and (5.13), when we consider a demand over capacity condition the deterministic model gives very bad results comparing to the other two conditions, while the stochastic model doesn't lose so much in terms of quality of results.

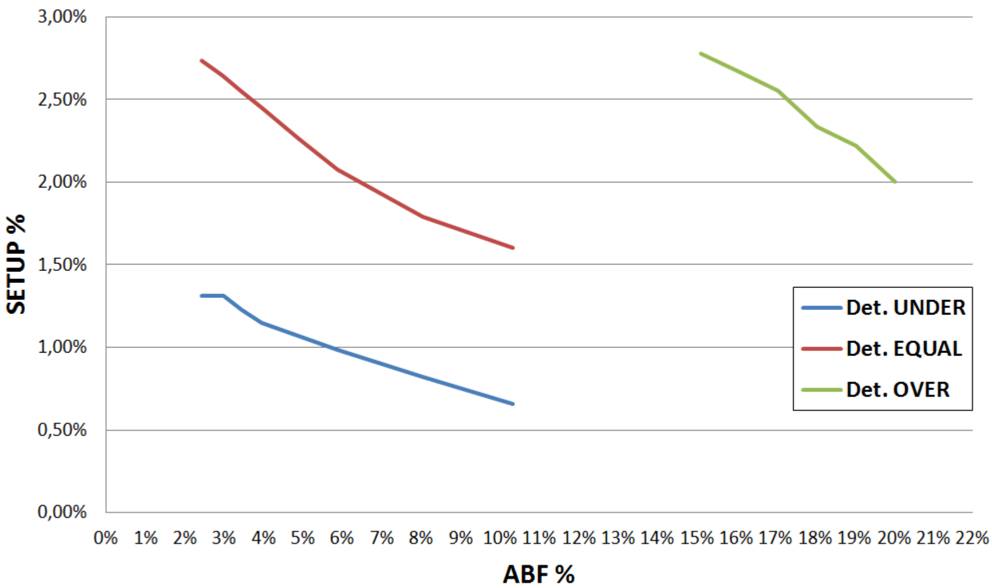


Figura 5.12: Comparison among different conditions: DET

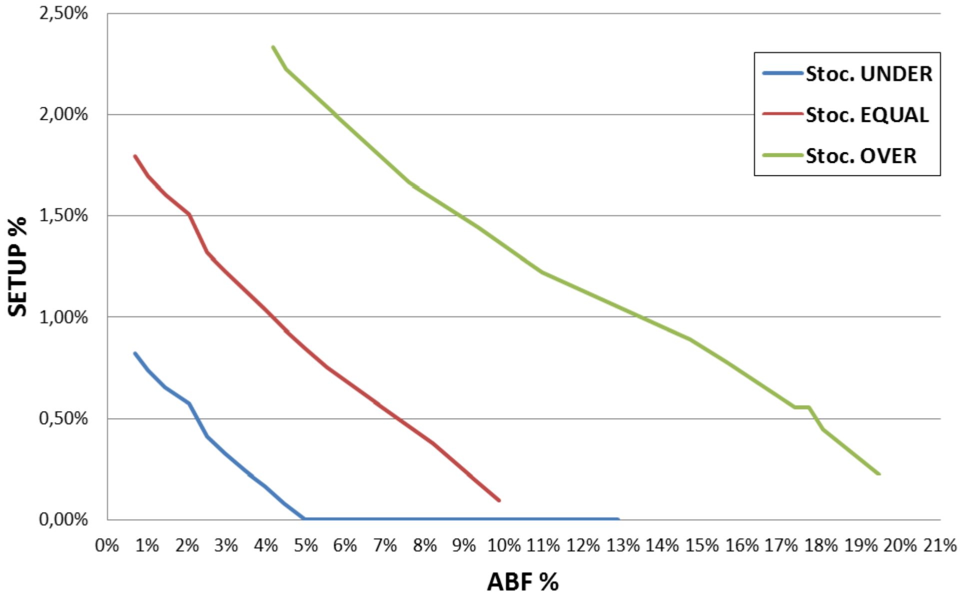


Figura 5.13: Comparison among different conditions: STOC

Capitolo 6

Computational experiments

In this chapter we are going to analyze in detail experiments in a dynamic setting, followed by a sensitivity analysis. Indeed, in the first part of the chapter an introduction to the dynamic setting is presented together with the results of a particular case of study. Later, in the second part, we analyze what happens if we change some of the factors involved. In particular, we show a sensitivity analysis focused on three parameters:

- Capacity of the manufacturing plant (with demand fixed);
- Shift factor (ratio between setup and inventory costs);
- Uncertainty level of the demand.

6.1 Dynamic setting experiments

6.1.1 Description of the experiments

With the static experiments we wanted to get some general ideas about how the implemented models perform. Now we want to compare the deterministic and the stochastic model (from now we consider only the three intervals model) using a Rolling Horizon approach, to prove that the stochastic model is a good approach to follow in order to solve the problem of the reactive modification of production schedules.

We consider one year of expected demand pattern for each SKU, the same ones described in the previous chapter. We solve both the models for a period of one year, approximated to 52 weeks. Indeed, what we do is to solve both models as we did in the static setting (with a four weeks horizon), but now with the application of the rolling horizon procedure. At the end of each week we compute the actual inventory and backorder values using the actual demand realized in that week (equal for the two models) and we solve again the models for the following four weeks starting from these values. We repeat this procedure as long as we complete the year.
Then, for each model, we compute the total costs on the basis of the storyline of realized changeovers and actual inventory values.
Every experiment we show is solved for one year 100 times for each set of conditions we want to investigate.
Forecasts don't change year by year, while every year has its actual demand history. In particular, for each year a demand storyline for each product is defined based on the uncertainty that characterizes the product. We generate 100 actual demand sets and we store these in a database, that we use as a simulation environment in order to compare the two models considering the same actual demand history. Every set contains the demand values for each product, for each week within the year.

6.1.2 Additional input data for the dynamic experiments

All the data considered as an input for the experiments have been explained in detail in paragraph (5.1). For the dynamic experiments we have to set these particular parameters:

- L = 240. We decide to set the number of looms (capacity of the plant) equal to 90% the expected pick of demand; in this way, the total number of weeks in which demand is higher than capacity is equal to 18. In figure (6.1) there is a representation of the capacity chosen, compared to the total expected demand.

- BF = 0, meaning that we ask the models to provide the best results in terms of service performance.

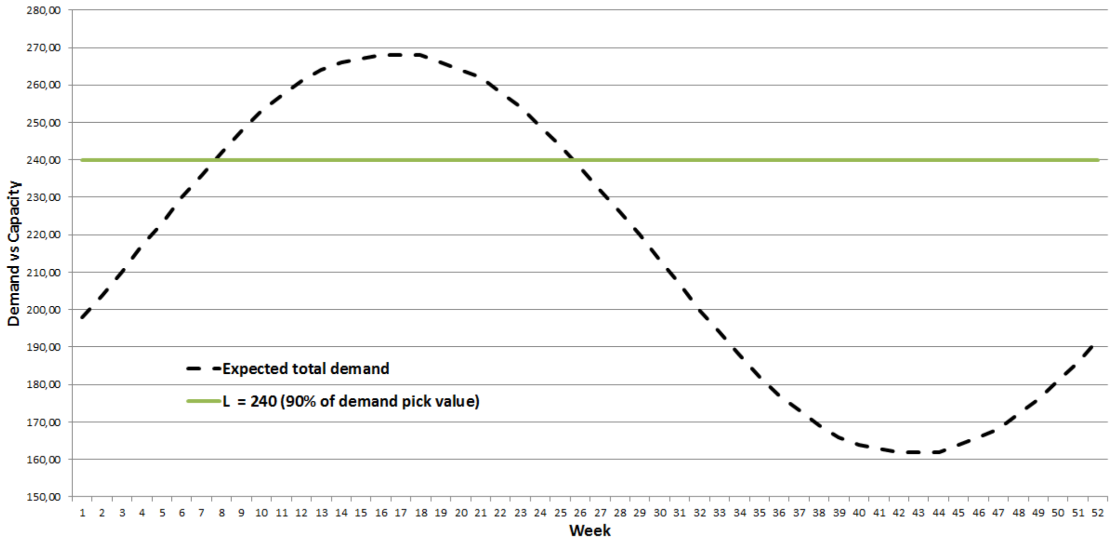


Figura 6.1: $L = 240$

6.1.3 Results

For every simulated year we get the results from both models in terms of:

- Total costs (changeover costs and inventory carrying costs);
- Number of changeovers;
- Backorder amount (sum of backorder values registered every week within the year) ;
- Inventory amount (sum of inventory values registered every week within the year);
- Number of looms running per week (utilization of the machines)

Before showing the results, we need to define some important indexes. We are going to consider the behaviors of both the stochastic and the deterministic models on the basis of service and cost performance. The cost performance is represented by the SETUP [%] value because it is the main cost element in our scheduling problem. On the other hand, service performance is described by the AVERAGE BACKORDER [%], that is the weekly average backorder (actual).

$$\text{SETUP [\%]} = \frac{\text{Total number of looms with a scheduled changeover in the year}}{\text{Number of looms} \cdot 52 \text{ weeks}}$$

$$\text{AVERAGE BACKORDER [\%]} = \frac{\text{Total backorder amount in the year}}{\text{Total demand of the year}}$$

Other important indexes we define are:

$$\text{d [Weekly average demand]} = \frac{\text{Total demand of the year}}{52 \text{ weeks}}$$

$$\text{IR [Inventory Rotation]} = \frac{d}{\text{Weekly average inventory}}$$

$$\text{UTILIZATION [\%]} = \frac{\text{Looms running in the year}}{\text{Number of looms} \cdot 52 \text{ weeks}}$$

We start the analysis of the results from figure (6.2). Every point represents the sum of the total costs for one year. Since we performed 100 years, in the chart we have 100 values for each model. As the weekly average demand increases (the values of demand, as we said before, are random and the same for both the models), also the total costs increase, because because the system work under stress and more changeovers are needed to follow the demand. We can also observe that total costs are always lower using the stochastic model.

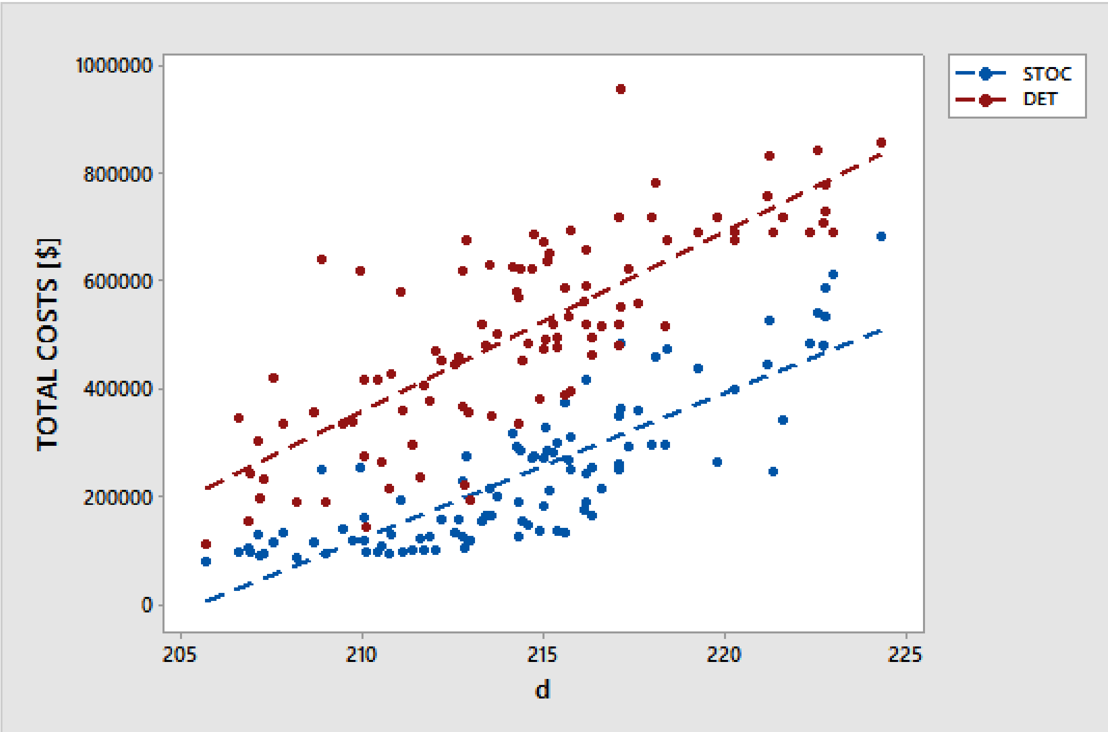


Figura 6.2: TOTAL COSTS [$] VS d

Consider the service performance index, figure (6.3): as the weekly average demand increases also the actual weekly average backorder amount increases, because there aren't enough looms to follow the demand properly. From this graph we can clearly see that the backorder amount values are considerably high: this is due to the fact that we assume that it's not available a forecast of the demand for a horizon longer than 4 weeks. This causes the so called myopia of the system, meaning that the planner can't organize the production taking into account medium and long-term forecasts.

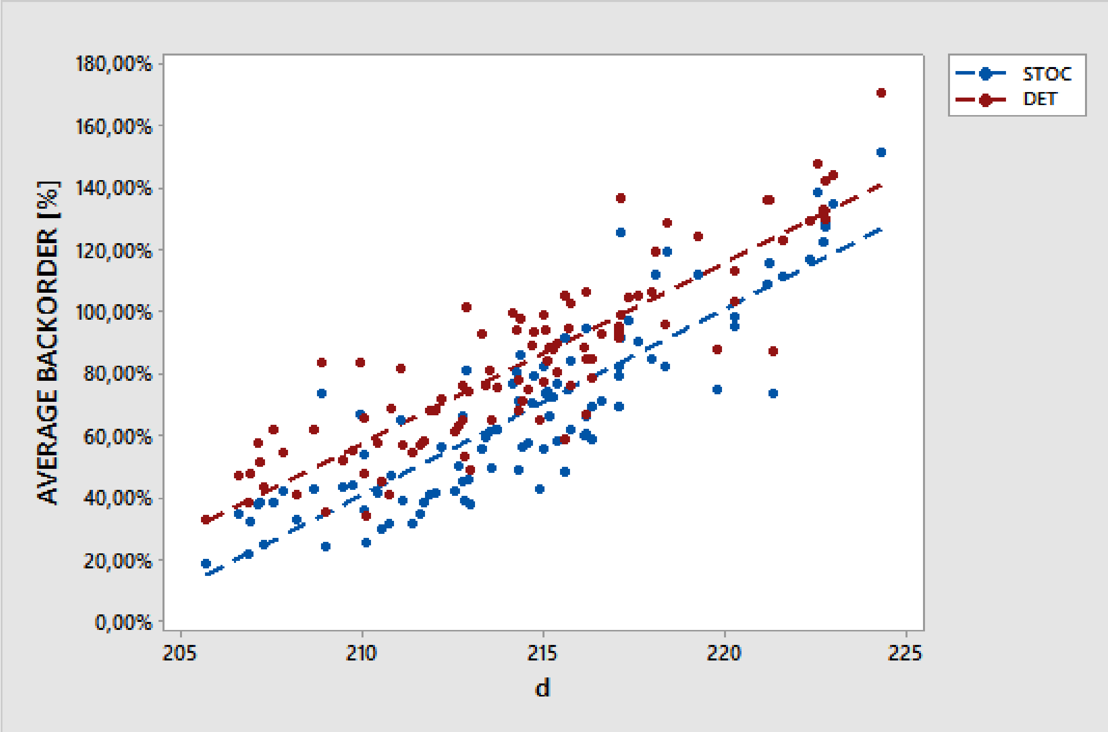


Figura 6.3: AVERAGE BACKORDER [%]VS d

In figure (6.4) the total costs are represented with respect to the AVERAGE BACKORDER [%], as a summary.

It seems, according to (6.2), that the stochastic model is better from the cost performance point of view. We perform a two-sample t-test with unequal variances, where the two samples are the total costs values of the deterministic and the stochastic models. The difference of the means is significative ($\alpha = 0.05$) with a p-value < 0.0001.

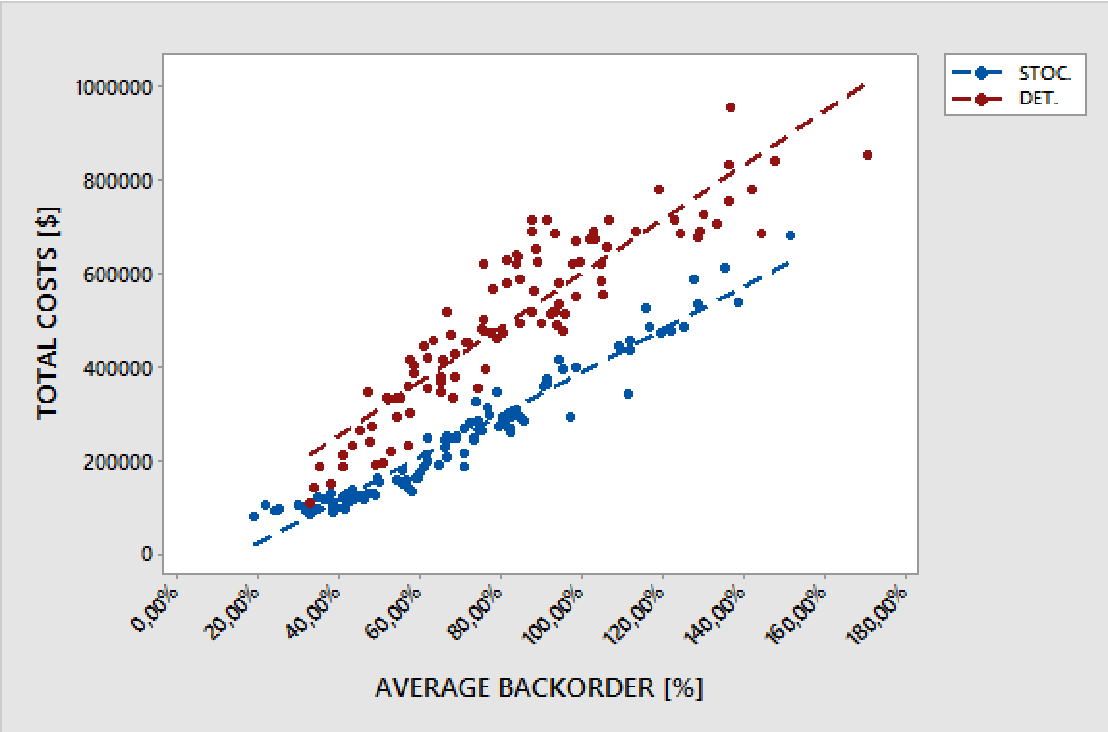


Figura 6.4: Total costs

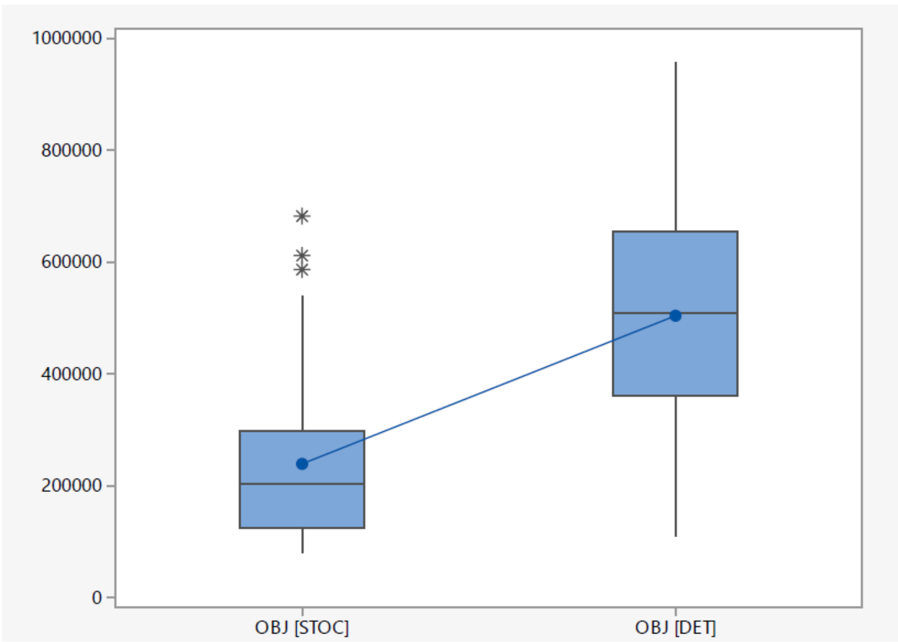


Figura 6.5: Box Plot of total costs

Method

μ_1: mean of OBJ [STOC]
μ_2: mean of OBJ [DET]
Difference: $\mu_1 - \mu_2$

Equal variances are not assumed for this analysis.

Descriptive Statistics

Sample	N	Mean	StDev	SE Mean
OBJ [STOC]	100	239426	139867	13987
OBJ [DET]	100	504143	183866	18387

Estimation for Difference

Difference	95% CI for Difference
-264718	(-310296; -219139)

Test

Null hypothesis H_0: $\mu_1 - \mu_2 = 0$
Alternative hypothesis H_1: $\mu_1 - \mu_2 \neq 0$

T-Value	DF	P-Value
-11,46	184	$<0{,}0001$

Figura 6.6: Two sample t-test with unequal variances: difference between total costs

We report in figure (6.7) a similar plot with respect to (6.4), but considering only changeovers on the y-axis. We see that the number of changeovers [%] suggested by a stochastic model is lower than the ones suggested by the deterministic model. Plots (6.7) and (6.4) are very similar: this is due to the fact that setup costs represent the main term of the total costs. A t-test is performed also in this case, obtaining a p-value < 0.0001.

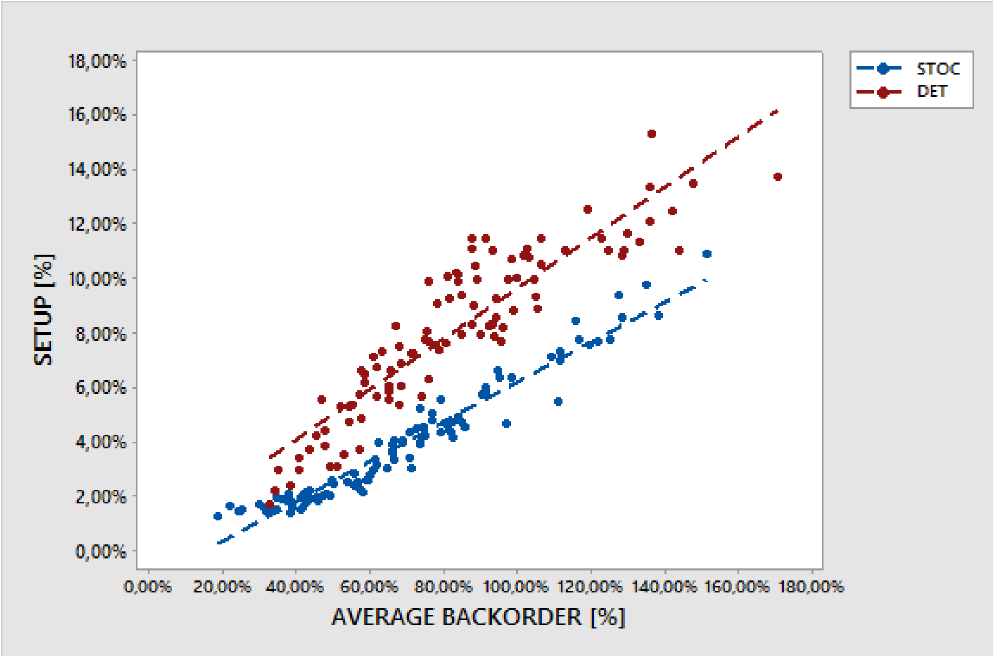
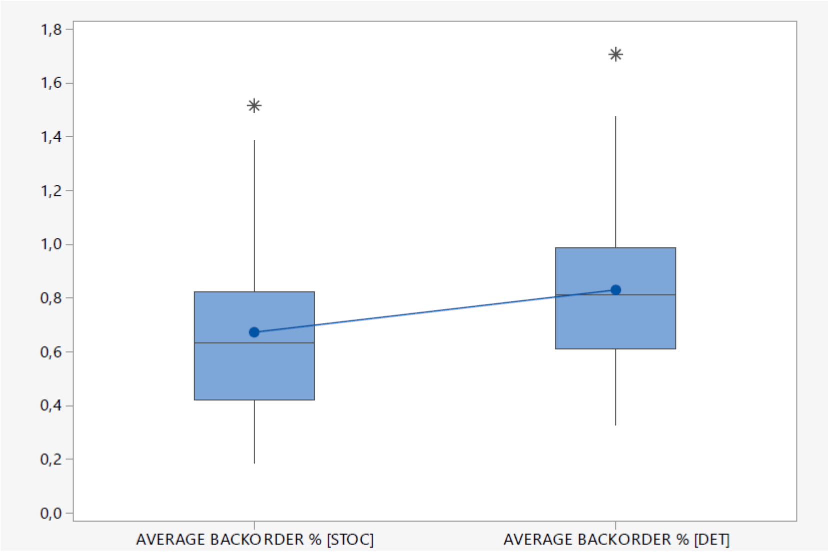


Figura 6.7: SETUP [%]

It also seems, according to (6.3), that the service performance given by the stochastic model is better than the deterministic one. In order to ensure statistical evidence to that, we perform a t-test in figure (6.8). The results shows that there is statistical evidence between the means.

The last results are ensured by two t-test we report in (6.9) and (6.10). We observe that there is not statistical evidence of the difference between the two models as regards the inventory level (represented by the Inventory Rotation parameter) and machines utilization.



Method

μ_1: mean of AVERAGE BACKORDER % [STOC]
μ_2: mean of AVERAGE BACKORDER % [DET]
Difference: $\mu_1 - \mu_2$

Equal variances are not assumed for this analysis.

Descriptive Statistics

Sample	N	Mean	StDev	SE Mean
AVERAGE BACKORDER % [STOC]	100	0,67358	0,29885	0,02989
AVERAGE BACKORDER % [DET]	100	0,83104	0,28954	0,02895

Estimation for Difference

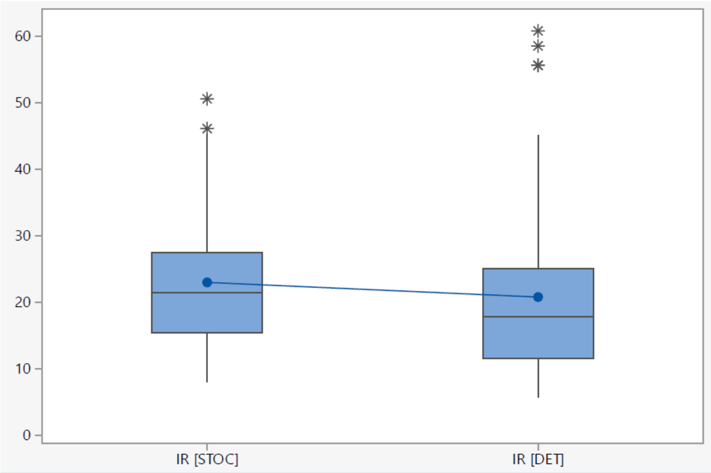
Difference	95% CI for Difference
-0,15746	(-0,23952; -0,07540)

Test

Null hypothesis H_0: $\mu_1 - \mu_2 = 0$
Alternative hypothesis H_1: $\mu_1 - \mu_2 \neq 0$

T-Value	DF	P-Value
-3,78	197	0,0002

Figura 6.8: Box Plot and t-test: service performance



Method

μ_1: mean of IR [STOC]
μ_2: mean of IR [DET]
Difference: $\mu_1 - \mu_2$

Equal variances are not assumed for this analysis.

Descriptive Statistics

Sample	N	Mean	StDev	SE Mean
IR [STOC]	100	23,0157	9,8223	0,9822
IR [DET]	100	20,812	12,046	1,205

Estimation for Difference

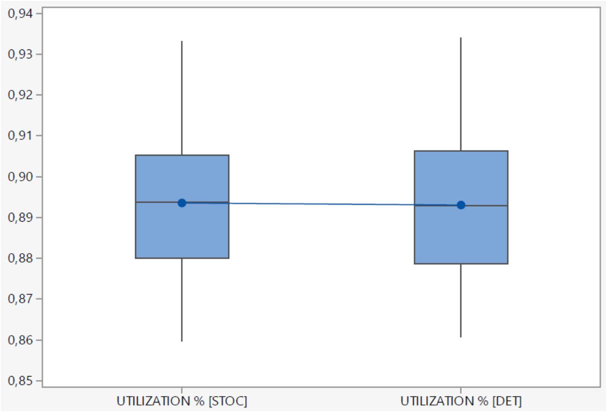
Difference	95% CI for Difference
2,203	(-0,862; 5,269)

Test

Null hypothesis	$H_0: \mu_1 - \mu_2 = 0$
Alternative hypothesis	$H_1: \mu_1 - \mu_2 \neq 0$

T-Value	DF	P-Value
1,42	190	0,1579

Figura 6.9: Box Plot and t-test: Inventory Rotation



Method

μ_1: mean of UTILIZATION % [STOC]
μ_2: mean of UTILIZATION % [DET]
Difference: $\mu_1 - \mu_2$

Equal variances are not assumed for this analysis.

Descriptive Statistics

Sample	N	Mean	StDev	SE Mean
UTILIZATION % [STOC]	100	0,893538	0,017844	0,001784
UTILIZATION % [DET]	100	0,893035	0,017864	0,001786

Estimation for Difference

Difference	95% CI for Difference
0,000503	(-0,004477; 0,005482)

Test

Null hypothesis H_0: $\mu_1 - \mu_2 = 0$
Alternative hypothesis H_1: $\mu_1 - \mu_2 \neq 0$

T-Value	DF	P-Value
0,20	197	0,8424

Figura 6.10: Box Plot and t-test: machines utilization

6.2 Sensitivity analysis

In this part of the chapter we are going to see how the results change on varying these three different factors:

- Capacity of the manufacturing plant (with demand fixed);
- Shift Factor (ratio between setup and inventory costs);
- Uncertainty level of the demand.

We decide to show the results in terms of SETUP [%] (defined in 6.1.3) and ΔSETUP [%], to focus the attention on the topic of the work. We define ΔSETUP [%] as:

$$\Delta\text{SETUP [\%]} = \frac{\text{SETUP[\%]}_{\text{DET}} - \text{SETUP[\%]}_{\text{STOC}}}{\text{SETUP[\%]}_{\text{DET}}}$$

It represents the reduction percentage of the number of changeovers obtained using a stochastic model with respect to a deterministic one.

6.2.1 Capacity of the manufacturing plant

In this section we analyze the results of the same experiment studied in 6.1.3 modifying only the capacity of the plant. Indeed, the demand is kept constant while the number of looms L changes. we are going to consider 6 different cases of capacity ($L = 210$, $L = 225$, $L = 240$, $L = 250$, $L = 268$, $L = 280$). All the other input data remain the same.

In figure (6.11) we show the different cases considered in the experiments. For example, the instance $L = 210$ is not so realistic, because the number of looms is too much small. This means that for more than 6 months in the year the expected demand is higher than capacity. The other instance $L = 280$ is not realistic too, because the number of looms is too high. In a real environment it's difficult to find a capacity that is equal to the maximum expected demand. However, it's interesting to study the results of the models in these instances to get an idea of the conditions in which the models (especially the stochastic model) perform better.

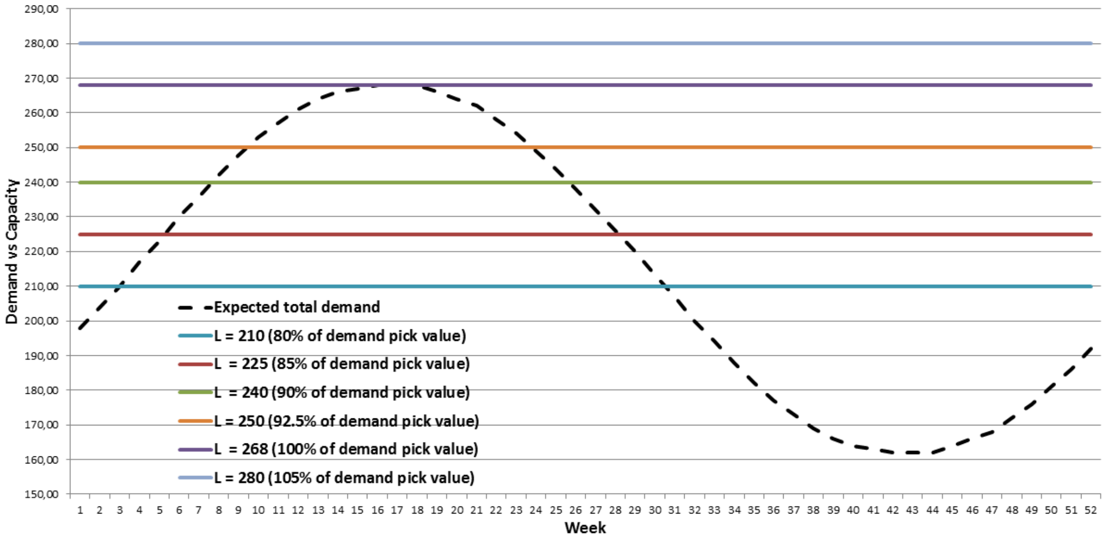


Figura 6.11: Different number of looms

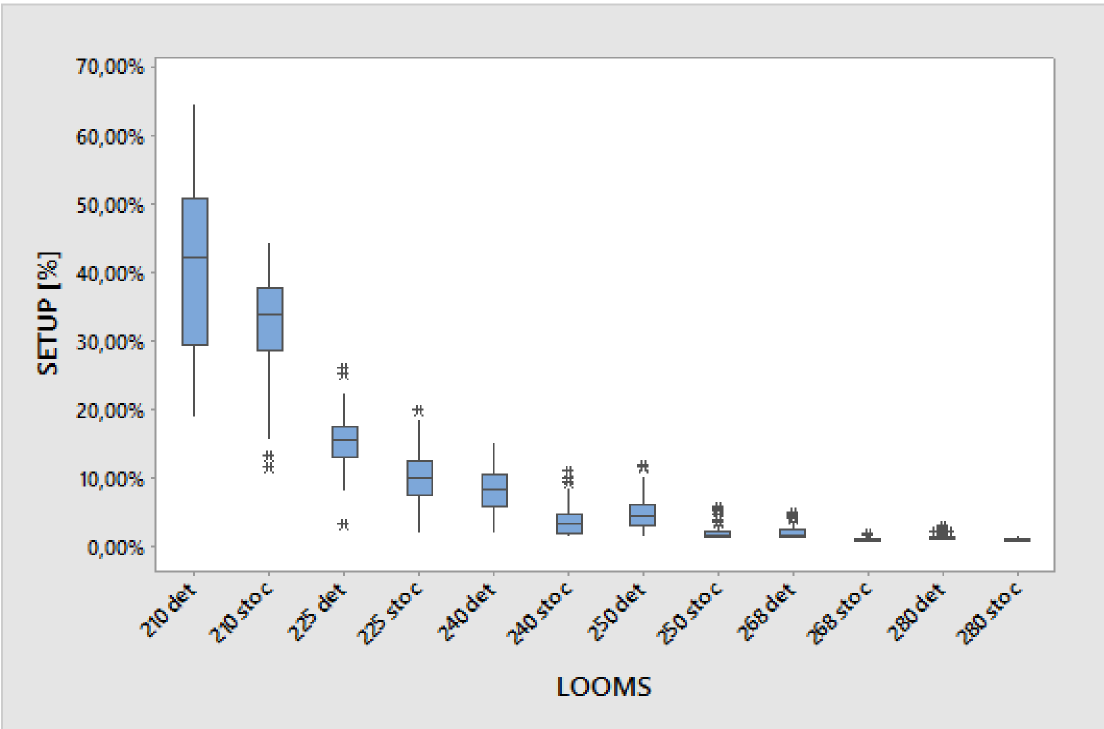


Figura 6.12: [Looms] SETUP [%]

In all the instances there is statistical evidence of the difference between the stochastic and the deterministic models.

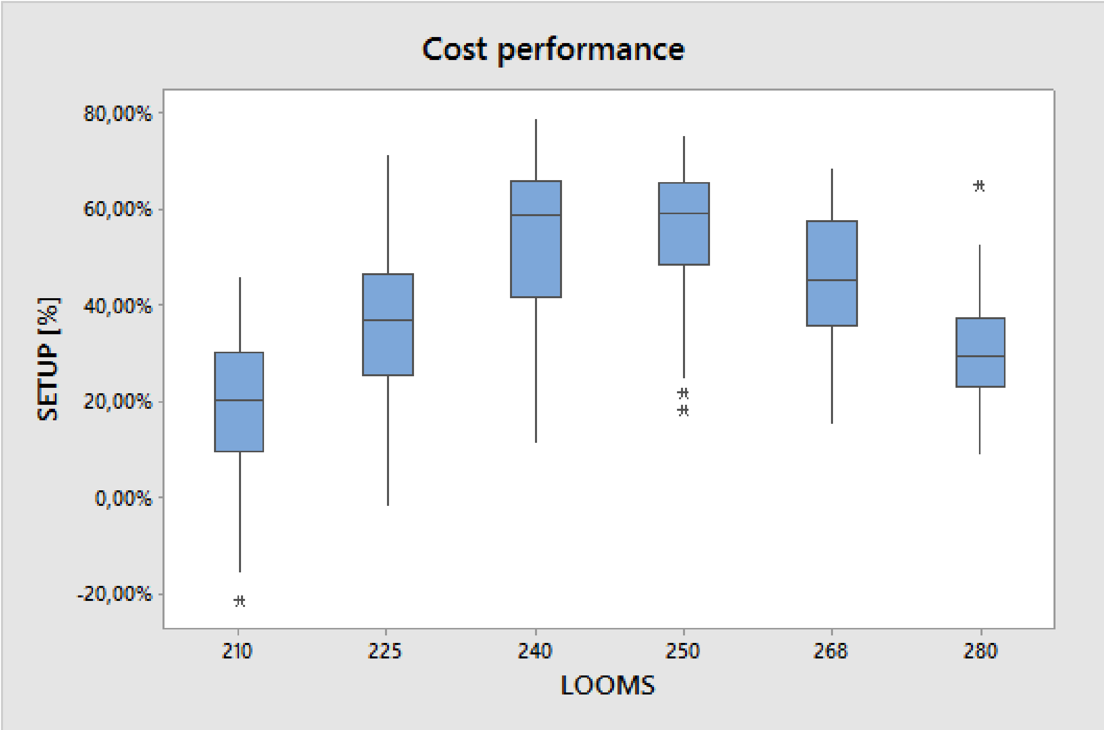


Figura 6.13: [Looms] ΔSETUP [%]

Considering the advantage of using a stochastic model, we see in figure (6.13) that it's possible to find an optimal value for a certain number of looms. It's interesting to note that either with a very large or very small number of looms with respect to demand the advantage of using a stochastic model decreases.
When the number of looms is too small, the system works in a stressed condition in which both the models suggest a lot of changeovers. The advantage offered by the stochastic model can't be so much.
On the other hand, when we have too many looms, the number of changeovers is very small also for the deterministic model. In this way the stochastic model can't perform so much better than the deterministic one.

6.2.2 Shift factor

The shift factor is an interesting ratio that we are going to study. It's defined as:

$$\text{SF} = \frac{\text{Setup cost}}{\text{Inventory carrying cost}}$$

In a similar way as we did in the previous section, we analyze the results of experiments with same data of section 6.1.3 on varying only the shift factor value. To do that, we change the setup cost keeping fixed the inventory cost. The three cases we are going to analyze are:

Setup cost	**Shift Factor**
500 $	866
400 $	693
300 $	520

We decided to decrease the changeover costs because we started our dissertation with a high setup cost. Thus, we want to study what happens in the comparison between the models if the changeovers were cheaper than the starting condition. In (6.14) we can see that for both the models the number of changeovers doesn't change very much on varying the shift factor, and the gap between the two models is confirmed among the three cases, (6.15). Of course, the results don't change because all the considered shift factors are very high. The objective function gives more importance to the setup reduction instead of the inventory one.

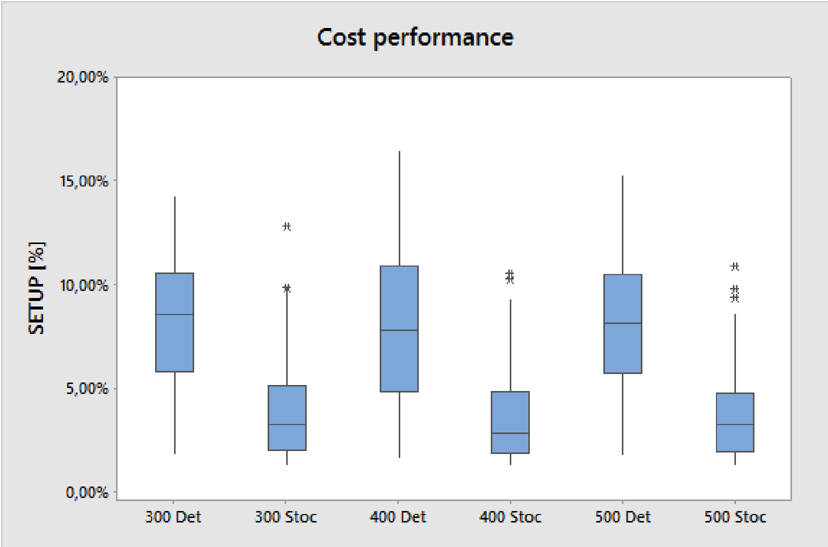


Figura 6.14: [Shift Factor] SETUP [%]

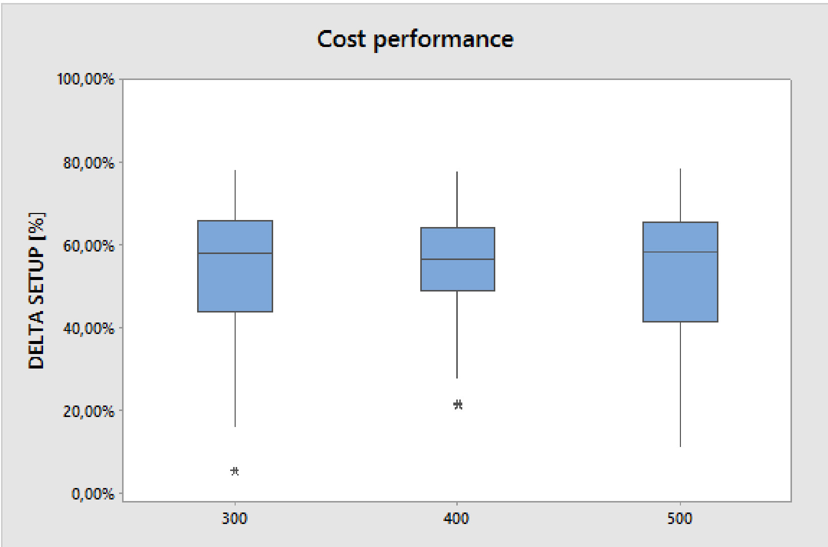


Figura 6.15: [Shift Factor] ΔSETUP [%]

6.2.3 Uncertainty level of demand

In this paragraph we study the cost performance obtained considering different set of CVs for the SKUs. In the following table we report the values used for the coefficients of variation.

SKU	$\mathbf{CV}_{\text{HIGH}}$	$\mathbf{CV}_{\text{MEDIUM}}$	$\mathbf{CV}_{\text{LOW}}$
SKU_1	0.1	0.05	0.005
SKU_2	0.2	0.1	0.01
SKU_3	0.4	0.2	0.02
SKU_4	0.3	0.15	0.015

The behavior of the stochastic model doesn't change on varying the uncertainty level, while the stochastic one lose performances decreasing the uncertainty grade (6.16). In figure (6.17) we can see that, decreasing the uncertainty of demand ($CV \downarrow$) the advantage in using a stochastic model decreases as well. This means that a stochastic approach is more useful the higher is the uncertainty of demand.

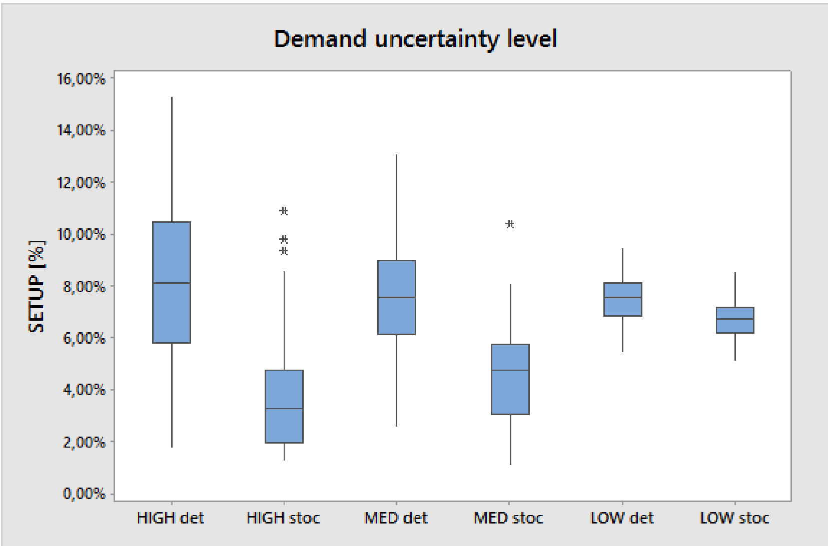


Figura 6.16: [CV] SETUP [%]

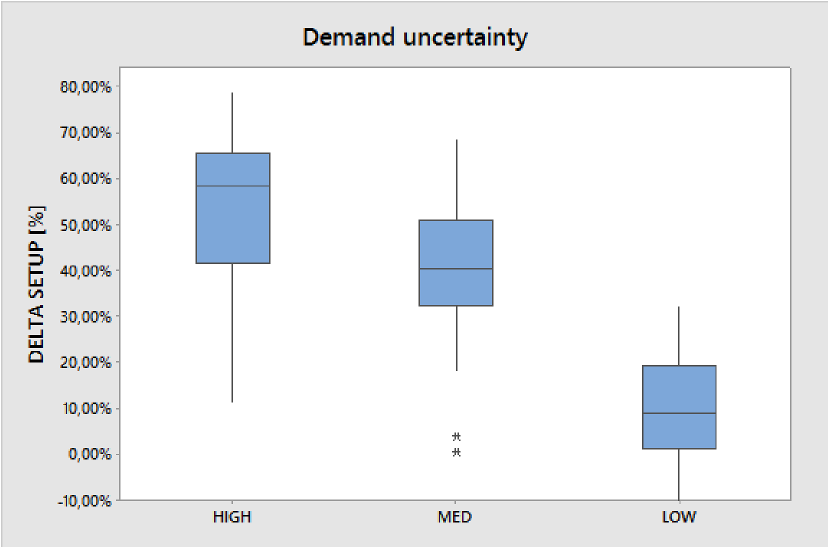


Figura 6.17: [CV] ΔSETUP [%]

Capitolo 7

Conclusions

The aim of the Thesis was to find a proactive approach for a lot-sizing and scheduling problem under uncertainty of demand.
The objective of the scheduling was the optimization of total costs, including setup costs and inventory carrying costs. Backorders are allowed, but they are not considered in the objective function. We assume it's not possible to assign a cost to backorder, so it is considered in terms of quantity as a constraint. Another hypothesis was that setup times were negligible.
We developed a deterministic model and we used together with a rolling horizon approach, the most used method to follow the evolution of an uncertain demand over time. According to the literature, this model presents the problem we wanted to study: the reactive modification of schedules, that leads into a large number of changeovers.
We developed a stochastic approximated model based on the expected value of the wait and see variables (see chapter 3), since the complexity of the problem and so high computational times, didn't let us using the complete scenario tree.
We considered a textile manufacturing scheduling problem as environment in which we want to test the models, because of the high uncertainty of the demand in that field. Another feature of this field is the high setup cost with respect to the inventory carrying cost.
Before performing the computational experiments related to the main topic of the research we considered a set of preliminary experiments. With these first experi-

ments we wanted to get a general idea of how the developed models worked.
We performed the so called static setting experiments, in which we compared the deterministic and the stochastic models generating the schedules for the same period of four weeks and observing the difference in terms of setup, number of looms running (utilization), inventory and backorder amount. In particular, we generated a database of actual demand values for each SKU and for each week and we observed how much inventory and backorder they produced according to the schedules suggested (without the possibility to change them when new information about the actual demand was available). As we can see in chapter 5, a stochastic approach provided a reduced number of changeovers (and a significative reduction of the costs in the objective function). Therefore, we saw that also performances in terms of service level increased and that inventory levels and number of looms running (on equal actual backorder amounts) were respectively lower and a little bit lower than using a deterministic approach. We tested this differences under different conditions (demand higher than capacity, demand equal to capacity and demand lower than capacity) and the same conclusions can be drown in every instance.
Thus, in experiments without the application of the rolling horizon procedure, we discovered that the stochastic models provide better results under all the performance indexes considered. Therefore, we saw that the performance of the stochastic models doesn't change very much as we increase the number of outcomes (for each product, for each period) from three to five.
However, the main objective of the work was to see in a rolling horizon environment ("dynamic setting") if a stochastic approach could be useful to reduce schedule nervousness.
From the results in a particular case ($L = 240$, high CV, BF = 0, SF = 866) we can conclude that there are good results in terms of total costs: there is statistical evidence of the (positive) difference between total costs obtained from the deterministic model and the stochastic model. There is also statistical evidence that the service performance is better with a stochastic approach. It's interesting to note that, for this instance, both the inventory level and machines utilization are the same for both the models. This means that machines are used in a better way.
After that, we wanted to see if, varying the values of some factors, the results in

terms of setup reduction were confirmed. We decided to focus our attention only on this performance index because this is the main problem related to the case of study, due to high setup costs.
We performed what we call a "sensitivity analysis" on three factors:

1. number of looms (capacity);

2. shift factor (setup and inventory carrying costs ratio);

3. coefficient of variations (uncertainty level of demand).

Varying the number of looms we saw that the advantage in terms of setup reduction is obtainable in each condition, but this advantage is lower with either a very small or a very large number of looms (respectively a condition in which capacity is very low or very high with respect to demand). Thus, there is a particular capacity condition in which this advantage is optimal.
The second factor we wanted to study is the shift factor: in the environment considered for the experiments, shift factors are very high. We observed, with the three shift factors used, that the results remained almost the same, as regards advantage in terms of setup reduction.
It's interesting to report the results on varying the uncertainty level of demand: the more uncertain is the demand the greater advantage can be obtained using a stochastic model. Even if the uncertainty of the context is very low, the stochastic model performs better than the deterministic one. If the context is not characterized by uncertainty of demand, the two models perform in the same way.

Limits and suggestions for a future work

In this section we write the main limits of this work, and suggestions for a future work:

- As we said in chapter 4, we chose a simple set of SKUs and we assigned to them an expected demand function. Results are obtained with these demand functions and solving the models 100 times considering a single year of demand pattern for each SKU. It should be interesting to solve more times the models as we did, but changing randomly the expected demand functions year by year, in order to assess the advantages of a stochastic model also in different products mix with different parameters.

- Other work can be done by changing the input data regarding costs, to study different environments (we studied only the instance in which setup costs are several times higher than inventory carrying costs). Since our model also includes inventory costs in the objective function, it should be useful to solve problems that are different from the case we studied.
- Therefore, it should be interesting to compare the deterministic model solved every week to a stochastic model solved not every week as we have done in this dissertation, but every two or more weeks. This strategy can be useful to improve the solution of the problem of the research.
- It's interesting to quantify the difference between the model suggested in this dissertation and a "complete" stochastic model, to verify how much do we loose in the approximation.
- We created a model that reduces the computational issue related to this multi-stage stochastic scheduling problem. Some of the features introduced by our model could be applied to other multi-stage stochastic problems.

Bibliography

[1] Graves S.C., *A Review of Production Scheduling*, Operations Research, 1981.

[2] Allahverdi A., Gupta J.N.D., Aldowaisan T., *A review of scheduling research involving setup considerations*, OMEGA The International Journal of Management Sciences, 1999

[3] Allahverdi A., Ng C.T., Cheng T.C.E., Kovalyov M.Y., *A survey of scheduling problems with setup times or costs*, European Journal of Operational Research, 2006.

[4] Ovacik I.M., Uzsoy R., *Rolling horizon procedures for a single machine dynamic scheduling problem with sequence-dependent setup times*, Annals of Operations Research, 1994.

[5] Chand S., Traub R., Uzsoy R., *Rolling horizon procedures for the single machine deterministic total completion time scheduling problem with release dates*, Annals of Operations Research, 1997.

[6] Kensuke I., Jun I., Takayuki S., Susumu M., *Stochastic Programming model for discrete Lotsizing and Scheduling Problem on Parallel Machines*, American Journal of Operations Research

[7] Tiacci L., Saetta S., *Demand forecasting, lotsizing and scheduling on a rolling horizon basis*, International Journal of Production Economics

[8] Sahinidis N.V., *Optimization under uncertainty: state-of-the-art and opportunities*, Proceedings Foundations of Computer-Aided Process Operations, 2003.

[9] Blackburn J.D., Kropp D.H., Millen R.A., *MRP system nervousness: causes and cures*, Engineering Costs and Production Economics, 1985.

[10] Kanakamedala K.B., Reklaitis G.V., Venkatasubramanian V., *Reactive Schedule Modification in Multipurpose Batch Chemical Plants*, Ind. Eng. Chem. Res, 1994.

[11] Jain A.K., Elmaraghy A., *Reactive Schedule Modification in Multipurpose Batch Chemical Plants*, International Journal of Production Research, 1997.

[12] Vin J.P., Ierapetritou M.G., *A new approach for efficient rescheduling of multiproduct batch plants*, Ind Eng Chem Res, 2000.

[13] Mendez C., Cerda J.C., *An MILP framework for reactive scheduling of resource-constrained multistage batch facilities*, in Proceedings of the Conference on Foundations of Computer Aided Process Operations, 2003.

[14] Birge J.R., Louveaux F., *Introduction to Stochastic Programming*, Springer, 2011.

[15] Kall P., Wallace S.W., *Stochastic Programming*, John Wiley and sons, 1994.

[16] Gassmann H.I., Ireland A.M., *Scenario formulation in an algebraic modelling language*, Annals of Operations Research, 1996.

[17] Dupoçova J., Consigli G., Wallace S.W., *Scenarios for multistage stochastic programs*, Baltzer Journals, 2000.

[18] Kaut M., Wallace S.W., *Evaluation of scenario-generation methods for stochastic programming*, 2003.

[19] Karabuk S., *Production planning under uncertainty in textile manufacturing*, Operational Research Society, 2007.

[20] Mitra S., *A white paper on Scenario Generation for stochastic programming*, 2006

[21] Dupoçova J., Growe-Kuska N., Romisch W., *Scenario reduction in stochastic programming*, Mathematical Programming, 2003.

[22] Romisch W., Heitsch H., *Scenario reduction algorithms in stochastic programming*, Computational Optimization and Applications, 2003.

[23] Higle J.L., Sen S., *Stochastic decomposition: An algorithm for two-stage linear programs with recourse*, Mathematics of Operations Research,1991.

[24] Dantzig G.B., Infanger G., *Large-scale stochastic linear programs-importance sampling and Benders decomposition*, in Computational and applied mathematics, 1992.

[25] Infanger G., *Planning under Uncertainty: Solving Large-Scale Stochastic Linear Programs*, Boyd and Fraser, Danvers, 1994.

[26] Infanger G., *Monte Carlo (importance) sampling within a Benders decomposition algorithm for stochastic linear programs*, Ann. Oper. Res., 1992.

[27] Ermoliev Y.M., *Methods of Stochastic Programming*, Nauka, Moscow, 1976.

[28] Ermoliev Y.M., Gaivoronski A.A., *Stochastic quasigradient methods for optimization of discrete event systems*, Ann. Oper. Res., 1992.

[29] Casey M., *The scenario generation algorithm for multistage stochastic linear programming.* Available at http://www.math.ups.edu/~mcasey/, 2002.

[30] Dempster M. A. H., Thompson R. T., *EVPI-based importance sampling solution procedures for multistage stochastic linear programmes on parallel MIMD architectures*, Annals of Operations Research,1996.

[31] Balasubramanian, Grossmann, *Approximation to Multistage Stochastic Optimization in Multiperiod Batch Plant Scheduling under Demand Uncertainty*, 2004.

[32] Chen L., Homem-de-Mello T., *Re-Solving Stochastic Programming Models for Airline Revenue Management*, Annals of Operations Research, 2009.

[33] Cizkova J., *Value of information in stochastic programming*, 2006.

[34] Wagner J.M., Berman O., *Models for planning capacity expansion of convenience stores under uncertain demand and the value of information*, Annals of Operations Research, 1995.

[35] Baker K.R., *Minimizing earliness and tardiness costs in stochastic scheduling*, European Journal of Operational Research, 2014.

[36] Thomassey S., *Sales forecasts in clothing industry: The key success factor of the supply chain management*, International Journal of Production Economics, 2010.

[37] Christie R., Wu S.D., *Semiconductor Capacity Planning: Stochastic Modeling and Computational studies*, IIE Transactions, 2002.

[38] Karabuk S., Wu S.D., *Coordinating strategic capacity planning in the semiconductor industry*, Operations Research, 2003.

[39] Rockafellar R.T., Roger J., Wets B., *Scenarios and Policy Aggregation in Optimization under Uncertainty*, Mathematics of Operations Research, 1991.

[40] Thompson S.D., Watanabe D.T., Davis W.J., *A comparative study of aggregate production planning strategies under conditions of uncertainty and cyclic product demands*, International Journal of Production Research, 1993.

[41] Vargas V., Metters R., *A master production scheduling procedure for stochastic demand and rolling planning horizons*, International Journal of Production Economics, 2011.

[42] Raiffa H., Schlaifer R., *Applied Statistical Decision Theory*, Harvard University, Boston, MA, 1961.

[43] Madansky A., *Inequalities for stochastic linear programming problems*, Management Science 6, 1960.

[44] Choi T.M., *Pre-season stocking and pricing decisions for fashion retailers with multiple information updating.*, International Journal of Production Economics, 2007.

[45] Sen A., *The US fashion industry – a supply chain review*, International Journal of Production Economics, 2008.

[46] Vaagen H., Wallace S.W., *Product variety arising from hedging in the fashion supply chains*, International Journal of Production Economics, 2008.

[47] Brown R.G., *Smoothing forecasting and prediction of discrete time series.*, Prentice Hall, Englewood Cliffs, 1959.

[48] Winters P.R., *Forecasting sales by exponential weighed moving averages.*, Management Science, 1960.

[49] Box G.E.P., Jenkins G.M., *Time series analysis forecasting and control.*, Prentice Hall, Englewood Cliffs, 1969.

[50] Papalexopoulos A.D., Hesterberg T.C., *A regression-based approach to short-term system load forecasting*, IEEE Trans Power Syst, 1990.

[51] Kuo R.J., Xue K.C., *Fuzzy neural networks with application to sales forecasting*, Fuzzy Sets Syst, 1999.

[52] Yoo H., Pimmel R.L. *Short-term load forecasting using a self-supervised adaptive neural network*, IEEE Trans Power Syst, 1999.

[53] Dallari F., Milanato D., *Metodologie di Sales Forecasting nel fashion-retail*, Logistica (settembre 2011), 2011.

[54] Ching-Chin C., Ieng AIK, Ling-Ling W., Ling-Chieh K., *Designing a decision-support system for new product sales forecasting*, Expert Syst Appl, 2010.

[55] Thomassey S., Fiordaliso A., *A hybrid sales forecasting system based on clustering and decision trees*, Decis Support Syst, 2006.

[56] Thomassey S., Happiette M., *A neural clustering and classification system for sales forecasting of new apparel items*, Appl Soft, 2007.

List of Figures