POLITECNICO
MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE E DELL'INFORMAZIONE

Laurea Magistrale in Ingegneria Meccanica

EVALUATING THE SELECTIVE LASER MELTING APPLICATION ON INVENTORY MANAGEMENT: detailed production process modelling and inventory policies analysis

Supervisor:
Professor Andrea Matta

Co-supervisor:
Professor Barbara Previtali

Candidate:
Beatrice Marisa Montano

Student number:
921018

Academic Year 2019/2020

A nonna Fiò,

che non ha mai smesso di accompagnarmi

nel mio percorso di crescita.

Somehow, I can't believe that there are any heights that can't be scaled by a man
who knows the secrets of making dreams come true.
This special secret, it seems to me, can be summarized in four Cs.
They are curiosity, confidence, courage, and constancy,
and the greatest of all is confidence.
When you believe in a thing, believe in it all the way,
implicitly and unquestionable.

Walt Disney

RINGRAZIAMENTI

Sono ormai giunta al termine di questo mio percorso universitario e formativo, che mi ha permesso di acquisire solide competenze tecniche e specialistiche da sempre ricercate, consentendo di realizzare il grande desiderio di diventare Ingegnere che portavo con me sin da bambina. A questo proposito, desidero ringraziare il Professor Andrea Matta, che mi ha costantemente guidato nello sviluppo del mio progetto di tesi, permettendomi di accettare la sfida di analizzare un ambito ancora poco studiato nella letteratura, così come la Professoressa Barbara Previtali e Francesco, che mi hanno pazientemente fornito conoscenze di dettaglio sui processi Additive, indispensabili per il successo del mio lavoro. Un grazie particolare anche alle Professoresse Erica Pastore e Arianna Alfieri, la cui guida è stata fondamentale per il pieno discernimento dei processi di gestione delle scorte. Ed infine grazie a tutti i Professori del Politecnico di Milano, che con i loro insegnamenti mi hanno trasmesso la passione verso il mondo tecnico-scientifico, formando un pensiero critico necessario a ricoprire consapevolmente il ruolo di Ingegnere.

Il conseguimento di questo importante traguardo è sicuramente frutto del costante supporto dei miei genitori Monica e Piero, che non hanno mai smesso di credere in me, incoraggiandomi nell'accettare e saggiamente superare ogni sfida che si è presentata, grande o piccola che fosse, gioendo insieme a me di ogni vittoria raggiunta. Grazie a Filippo, che mi ha sempre donato le parole, gli abbracci o semplicemente gli sguardi che cercavo, condividendo e valorizzando ogni singolo giorno di questo viaggio. Grazie anche a tutta la mia famiglia, ai miei zii, cugini e cuginette, i cui sorrisi e semplici ma attenti e preziosi gesti hanno sempre spontaneamente accompagnato questo percorso. Ed infine soprattutto grazie ai miei nonni, ed in particolare a nonna Fiò, che mi sono sempre stati accanto, indipendentemente da luoghi e circostanze, e non lasciando mai cadere quella promessa del *pensaghi no* mantenuta in ogni occasione. Devo a tutti loro molto più di quanto io immagini, e un grazie non potrà mai bastare.

Sicuramente affrontare questi cinque e più anni non sarebbe stato lo stesso senza le mie amiche e i miei amici, che da tempi immemori sono compagni di risate, viaggi ed avventure impareggiabili, e che hanno saputo come rallegrare e alleggerire questo percorso. Grazie anche a tutti i miei compagni del Poli, che hanno reso le giornate in aula uniche nel loro genere e senza i quali raggiungere questo obiettivo non avrebbe avuto lo stesso sapore. Infine, grazie a tutte le persone che ho conosciuto durante questa esperienza, ed in particolare ai miei compagni di Erasmus e coinquilini dello Studax, con i quali ho trascorso un'esperienza indimenticabile che sempre mi porterò nel cuore.

Grazie ad ognuno di voi, per aver reso possibile il compimento di questo percorso, offrendomi la possibilità di crescere, maturare e formarmi, senza mai farmi smettere di sorridere ed essere pienamente soddisfatta della persona che sono.

Beatrice

ABSTRACT

In the recent years, a growing interest on Additive Manufacturing (AM) emerged in scientific literature. The technology flexibility and agility, that allow production without specific tools needs, promoted the AM application in industrial scenarios, suggesting the technology to be a disruptive solution to respond quickly to unstable and unpredictable demands, possibly revolutionizing inventory management and Supply Chain (SC). Despite this, very limited research has been devoted to quantitatively analyse how AM can effectively fulfil demands and eventually reduce inventory levels. In addition, it has been noted a literature gap between the research focused on the AM production efficiency, that selects a batch production for optimizing at best the machine capacity, and the literature that investigates the AM application on inventory management and SC, that promotes the technique for a Make To Order (MTO) or piece-by-piece production. Furthermore, very often, the latter provides a generic description of the AM process, asserting many assumptions as infinite machine capacity, negligible set up and cool down times, not contemplated post processing operations and constant or generic distributed lead times. This work has therefore the aim to develop an accurate AM process description, critically removing the literature hypotheses and focusing in particular on Selective Laser Melting (SLM), to provide an exhaustive evaluation on the AM impact on the inventory management, considering a single-item, single-location scenario.

Markov Chains with different levels of SLM modelling details have been developed to estimate the resupply time. It has been obtained that an accurate SLM process description can lead to more precise lead time evaluation, sensibly reducing the estimation error. Simulation models have been created to assess how AM, used as production technique to replenish the stock, can influence the inventory management. With the objective to fulfil the literature gap, the mentioned production strategies have been studied applying two different continuous review inventory policies: the (S-1,S) for MTO, and the (r,Q) for batch production. To strengthen the validity of the models, a case of study provided by the Italian Company Fubri has been examined. The analysis pointed out that a detailed SLM process description, considering all the time phases that the technique requires, can improve the selection of the optimal inventory parameters, reducing the total annual inventory cost and increasing the service level with respect to a more generic modelling. Furthermore, it has been shown that the (r,Q) inventory policy allows a reduction of the unitary resupply time and finally of the total annual inventory cost compared to the (S-1,S) one. This noteworthy outcome is the consequence of an improved optimization of the SLM machine capacity, a proper allocation of the production fixed times and a subsequent reduction of the total annual order costs.

SOMMARIO

Negli ultimi anni, la letteratura scientifica ha dimostrato un crescente interesse per l'Additive Manufacturing (AM). La flessibilità e l'agilità di questa tecnologia hanno promosso la sua applicazione in diversi scenari industriali. L'AM è considerata infatti essere una soluzione dirompente per rispondere rapidamente a domande instabili e difficilmente prevedibili, e che possa offrire l'opportunità di rivoluzionare la gestione delle scorte e la Supply Chain (SC). Ciononostante, sono ancora limitati gli studi dedicati all'analisi quantitativa dell'efficacia e prontezza dell'AM in risposta al mercato, e di come questa tecnologia possa eventualmente ridurre i livelli di magazzino. Inoltre, è stato notato un divario bibliografico tra ricerche sull'efficienza produttiva dell'AM, che promuovono una produzione per lotti per ottimizzare al meglio la capacità di macchina, e gli studi sull'applicazione dell'AM nella gestione delle scorte e della SC, che incoraggiano una produzione Make To Order (MTO) o pezzo per pezzo. Si è notato, in aggiunta, come questi ultimi forniscano una descrizione generica del processo AM, portando ipotesi come capacità infinita di macchina, tempi di set up e raffreddamento trascurabili, operazioni di post-processing non considerate e lead times costanti o genericamente distribuiti. Questo lavoro si pone quindi come scopo lo sviluppo di un'accurata descrizione del processo AM, rimuovendo le ipotesi della letteratura e concentrandosi in particolare sulla fusione laser selettiva (SLM), per fornire una valutazione esaustiva dell'impatto AM sulla gestione dei magazzini in uno scenario mono-prodotto e a singola location.

Per rispondere a questo obiettivo, sono state sviluppate Catene di Markov con diversi livelli di dettaglio nella modellazione SLM per stimare il tempo di rifornimento delle scorte. Lo studio ha rivelato che un'accurata descrizione del processo SLM porta a una valutazione più precisa del lead time, riducendone sensibilmente l'errore di stima. Sono stati quindi creati modelli simulativi per valutare l'impatto dell'AM come tecnica produttiva per rifornire i magazzini. Considerando lo scopo di indagare il divario menzionato in letteratura, sono state suggerite due politiche di gestione delle scorte: (S-1,S) per una produzione MTO e (r,Q) per una produzione per lotti, entrambe a revisione continua. È stato possibile rafforzare la validità dei modelli sviluppati esaminando un caso di studio fornito dall'azienda Fubri. L'analisi ha sottolineato che una descrizione dettagliata del processo SLM, considerando tutte le fasi che la tecnica richiede, può migliorare la selezione dei parametri ottimali delle politiche di gestione delle scorte, riducendo il costo totale annuo di magazzino e potenziando il livello di servizio rispetto a una modellazione più generica. Inoltre, la scelta della politica (r,Q) consente una riduzione dei tempi unitari di rifornimento ed infine del costo totale annuo di magazzino rispetto a quelli (S-1,S). Questo importante risultato è la conseguenza di un'ottimizzazione della capacità della macchina SLM, di una corretta allocazione dei tempi fissi di produzione e di una risultante riduzione dei costi annui d'ordine.

Table of content

List of Abbreviations

3D: Three Dimensional

AM: Additive Manufacturing

B: Backorders

CAD: Computer Aided Design

CAM: Computer Aided Manufacturing

CI: Confidence Interval

CM: Conventional Manufacturing

CNC: Computerized Numerical Control

CV: Coefficient of Variation

DLP: Digital Light Projection

DMLM: Direct Metal Laser Melting

DSP: Digital Spare Parts

EDM: Electric Discharge Machining

EMB: Electron Beam Melting

EOQ: Economic Order Quantity

Er: Erlang

EX: Exponential

FCFS: First Come First Served

FDM: Fused Deposition Modelling

HE: HypoExponential

JIT: Just In Time

KPI: Key Performance Indicator

LBM: Laser Beam Melting

LOM: Laminated Object Manufacturing

MC: Markov Chain

MRO: Maintenance, Repair and Operations

MTO: Make To Order

O: Orders

OH: On-hand

PBF: Powder Bed Fusion

PH: Phase Type

QR code: Quick Response code

R.T.: Resupply Time

RBV: Reduced Build Volume

SC: Supply Chain

SLA: StereoLithography Apparatus

SLM: Selective Laser Melting

SLS: Selective Laser Sintering

STL: STereoLithography or Standard Triangulation Language

TM: Traditional Manufacturing

TT: Thermal Treatment

U.R.T: Unitary Resupply Time

UV: UltraViolet

List of Figures

List of Tables

1 Introduction

1.1 Industrial background: Additive Manufacturing technology

Additive Manufacturing (AM) is a technology that allows the production of physical objects adding up material layer-by-layer, line-by-line or surface-by-surface. This technology diffusion was possible thanks to the invention of computers, as well as of software for Computer Aided Design (CAD) and Manufacturing (CAM) and numerical control machines. In fact, AM process starts from the creation of a three-dimensional model of the desired product using CAD software, considering all the details and final dimensions required. It is also possible to generate the model by computer tomography or magnetic resonance imaging that directly scan an already existing object to produce its digital model. Next, the 3D CAD file is converted into the stereolithography format file (STL), which is a discretization of the object surface into triangles, creating in this way a document that can be processed directly by the AM printing machines. Software dedicated to AM printing slices the digital model creating the cross-sections that the manufacturing process would build up, with the possibility to set up the scanning mode and scanning patterns. The model is subsequently setup on the AM machine, taking care of part locations and orientations. It is important to underline that AM technology often requires the introduction of support structures, which sustain the overhanging parts of the object to be built, avoiding them to undergo excessive stress, and allowing a better distribution of the thermal load. After these modelling phase, the AM machine is setup: material is loaded, process parameters initialized and often AM machine is preheated, especially for those processes that can cause consistent thermal gradients, as the ones that require the fusion of metal powder (e.g. Selective Laser Melting). Once the setup is completed, the AM machine proceeds printing the object following the sequence of layers imposed by the digital file. After production, the printed parts are removed from the machine, cleaned and submitted to post processing treatments, that depend specifically on the AM technique used. A scheme of the described process flow can be observed in Figure 1.1.

The Additive Manufacturing technology dates back to 1986 with the invention of the Stereolithography Apparatus (SLA) by Charles W. Hull, a technique that is based on the photopolymerization: a photosensitive polymer resin is scanned and cured layer-bby-layer by means of an UV laser. Starting from this first patent, a series of different AM technologies were born, such as MIT's 3D printing process in 1989 and Laser Beam Melting (LBM) processes in the early 1990s, followed by the successful commercialization of process technologies including Fused Deposition Modelling

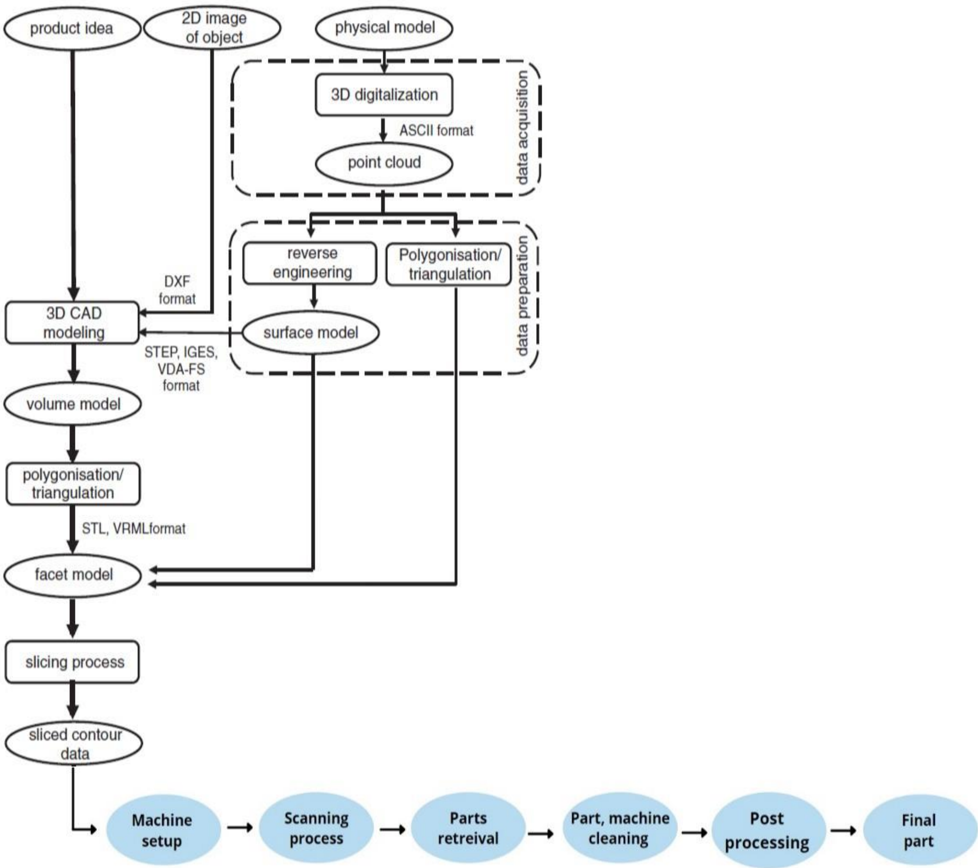


Figure 1.1 Additive Manufacturing digital and physical process flow (Thompson et al., 2016).

(FDM), solid ground curing, and laminated object manufacturing in 1991, and laser sintering in 1992 (Thompson *et al.*, 2016). This important growth in the AM technology is firstly due to improvement in geometric modelling capabilities and to the development of new and more performing CAD software, as well as the diffusion of computers and CNC machines. Consistent progresses in the technologies were gained in 1990s and 2000s with commercialization of new processes as Electron Beam Melting (EBM) and the improvement of the already existing ones, concurrently with the introduction of software specifically designed for AM printing, as the Materialise's Magic, which is still used. These improvements and the potential that the AM technology was showing lead to its application to the production of specific tools and prototypes. Therefore, terms like Rapid Tooling', 'Rapid manufacturing' and 'Rapid Manufacturing' were born. The ability to produce a great variety of features with the freedom in creating exactly the desired shape and the lack of need of specific tools that AM allowed, stimulated researchers to use this technique to make almost just in time a 3D model of the part to produce: this permitted the 3D visualization and allowed the testing of the object in terms of shape and design before giving it to production. An important application of rapid prototyping is found in the biomedical field: in case of complex operations, surgeons often take advantage of a

3D model of the body part that needs medical treatment to visualize it and define a better intervention strategy (Banoriya, Purohit and Dwivedi, 2015).

The growing interest in the Additive technology and the evidence of its potential lead to the development of more performing technologies, investments on improving quality and reliability of the parts produced and finally, thanks to the expiring of the 1980s patents, the process diffusion in the market competition. This spurred the innovation, increasing the demand encouraging the AM technology diffusion and application.

Nowadays Additive Manufacturing is used in a great variety of fields: aerospace, automotive, military, medical and dental industries are just a bunch of the areas where the technology finds its applications. Giving some examples, Boeing equipped the 777X jet engine with more than 300 parts produced by GE Aviation (GE Additive, 2020); NASA made possible to even produce 3D printed part on the International Space paving the way to future long-term space expeditions (NASA, 2020) and launched a project for creating a rocket combustion chamber thanks to a combination of 3D-printing techniques, revealing its success (Ridinger, 2018); many examples of customized medical equipment as hearing aids, dental crowns and implants, patient-specific prostheses with efficient anatomical alignment can be found in the market (Thompson *et al.*, 2016) as well as applications of the technology in aesthetic and fashion industry, as 3D printed dresses saw in Iris van Herpen's Haute Couture show, 'Voltage' (Materialise, 2013) or the Adidas Futurecraft sport shoes made with an innovative Additive Manufacturing technique (Materialise, 2015).

The latest frontier in the diffusion in Additive Manufacturing consists in end users' home. In fact, the increasing interest on the technology capabilities allowed a reduction of machine investment costs, making the process affordable even for common consumers, creating a completely new scenario of personal manufacturing, where end users are able to directly produce parts for their needs, opening a new future operation perspective (Ryan *et al.*, 2017).

1.2 Additive Manufacturing as production process: technological advantages and challenges

The evolution of industries depends on innovative and cutting-edge research activities associated with manufacturing processes, materials, and product design. In addition to the customary demands of low price and best quality, the market competition in current production industries requires products that are intricate, possess shorter life cycles, exhibit reduced delivery times and involve customization. In fact, the current breed of products is very complicated and challenging to design (Abdulhameed *et al.*, 2019). Accordingly, there is a strong incentive toward the design, development, and implementation of new and ingenious manufacturing processes.

These market and industrial requests can be satisfied by what the experts describe as a disruptive technology in the industry 4.0: Additive Manufacturing. The great potential that this technology offers had led to an increasing interest in its study. This can be appreciated by looking at the results of Ryan *et al.*, (2017) work. The authors carried out a systematic literature review to appreciate the AM attention on research in terms of application of the technology in different production scenarios, starting from the technology birth until 2016. From Figure 1.2 it is possible to appreciate an increasing interest in AM application, especially since the old technology patents expired, making AM more accessible.

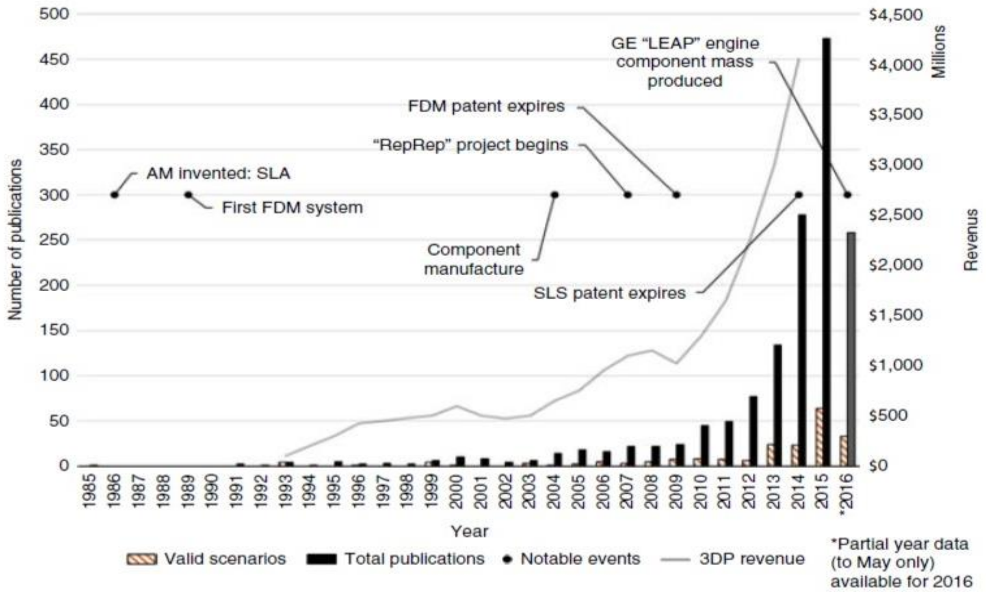


Figure 1.2: Number of publications regarding AM applications in production scenario per year (Ryan *et al.*, 2017)

In addition, a sensible industrial diffusion of AM processes is taking place nowadays: according to the annual worldwide report on the progress of AM technologies produced by Wohlers Associates (Wohlers Report, 2012), the global revenues from AM production and associated services grew from $2.25 billion in 2012 to more than $6 billion in 2016 and the forecast is to reach $21 billion by 2020. In 2012, just 28% of AM components were functional and therefore used in a real industrial context, while in 2016 this percentage went up to almost 34% (Gisario *et al.*, 2019). Furthermore, companies willing to invest in this technology keep increasing. Indeed, Marchese, Crane and Haley, (2015) state that 24% of manufacturing firms use AM in some form, a percentage that rises to 50% among Supply Chain leaders and affirm that this percentage is going to grow in the subsequent years.

This growing interest in AM technology and its application in industrial production can be justified by looking at the benefit that this technology brings with respect to traditional manufacturing. One of the most mentioned advantages is the *design freedom and flexibility* that AM allows without the request of specific tools. For example, AM has been used to create complex internal pathways for acoustic damping devices, optimized fluid channels and improved conformal cooling. Conformal cooling channels follow the external geometry to provide more effective and consistent heat transfer: complex geometry and internal features that conventional manufacturing techniques like milling have difficulties to obtain, are realized thanks to AM technique, leading to more uniform temperature distribution and improved quality (Figure 1.3) (Thompson *et al.*, 2016).

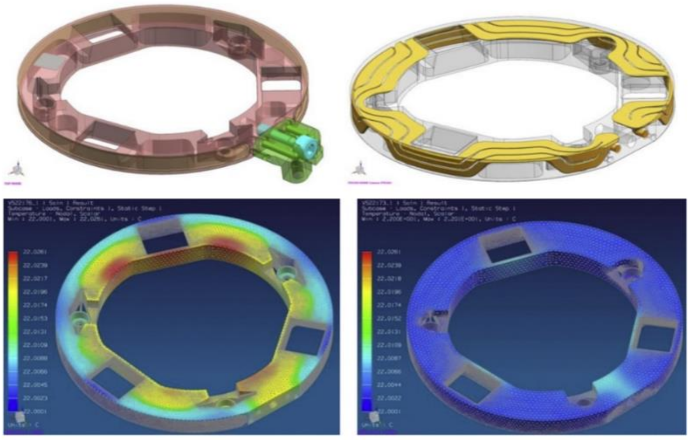


Figure 1.3: Thermal conditioning ring with milled colling channels enclosed by a welded cover (left) and with additively manufactured conformal cooling channels (right). Temperature plots from finite elements model in the corresponding components (Thompson et al., 2016).

The design flexibility that AM provides allows a great degree of *customization* in the final product. An example is offered by the biomedical field, where AM is used to produce patient-specific models to facilitate surgical planning to improve accuracy and efficiency, as well as a wide variety of personalized and tailored products including hearing aids dental crowns, implants, customized prostheses (Figure 1.4).

Figure 1.4: Titanium implants for the skull (left) and pelvis (right) produced using an EOSINT M 280. (Thompson et al., 2016).

An additional advantage can be referred to *topology optimization*. It is a numerical approach that identifies where material should be placed in a given domain to achieve a desired functionality (e.g. stiffness) for a given set of loads and constraints, while optimizing for properties such as minimal material usage/weight or uniform stress distribution (Thompson *et al.*, 2016). For example, AM lattices can be used to produce high stiffness low weight structures really used in the aerospace field. In particular, AM implementation can reduce the critical "buy-to-fly" ratio, which represents in aeronautical field the amount of raw material used and the final component weight. Generally, this term is quite high because of hard manufacturing materials as titanium alloys that the sector requires. An example is provided by Lockheed: they have a buy-to-fly ratio of 33:1 traditionally machining a bleed air-leak detect (BALD) bracket using subtractive manufacturing. This ratio was brought down to nearly 1:1, reducing bracket manufacturing costs by over 50 percent thanks to the introduction of AM (Marchese, Crane and Haley, 2015). In addition, because of the lighter structures that can be designed, it is possible to obtain a *weight reduction* of 70% (Ghadge *et al.*, 2018), leading to tangible benefits as fuel saving during utilization (Togwe, Eveleigh and Tanju, 2019). Finally, this advantage has implications even in *supply chain* in inventory management, transportation, warehousing, and purchasing: lower order quantities mean less transportation, and lower space requirements for raw materials (Kunovjanek and Reiner, 2020).

Finally, AM can *simplify process flow*. In fact, parts that are made by different sub-components, can be produced as just one piece. This reduces the time requested for the assembly process, the need of welding that can be critical for mechanical properties and can make the production process as well as the supply chain leaner. At the same time, manufacturers do not need to store and maintain different moulds to produce different components. Marchese, Crane and Haley (2015) report a case of a manufacturer that made use of die casting and then collected the parts for the final part assembly: they estimated an investment in die cast tooling accounted for 91 to 99% of total part cost, depending on production volume. Therefore, AM can lead to a cost reduction, being competitive for low-volume production.

Despite all the advantages reported, AM is still a quite new technique and possible limits related to the need of additional research can be found. In particular, the last-mentioned advantage of assembly reduction can also show drawbacks. In fact, consolidating all the sub-components in just one part would require replacing the whole total component in case of failure, while the assembled one could have just one of its parts replaced. For sure having lots of component means keeping high inventory level, so strategies to balance both of these elements shall be discussed (Gisario *et al.*, 2019). Furthermore, complains are related to the final part *quality and mechanical properties*. In fact, the components produced with AM generally shows anisotropic properties strictly related to the build direction.

One common method to address these anisotropies is to modify the part or assembly orientation to minimize their impact. Other options include finishing operations after each layer (Thompson *et al.*, 2016).

Additionally, the layers made by AM are rarely, if ever, seamless. This leads to an object roughness strictly connected to the length scale and layer thickness used in discretization. For this reason, AM parts often require finishing operations. Another peculiarity of AM process is the need of *support structures*. Overhanging structures that are not sustained during production can collapse, having as consequence the total part production loss. Therefore, changes in build part orientation or the introduction of support structure are requested (Figure 1.5).

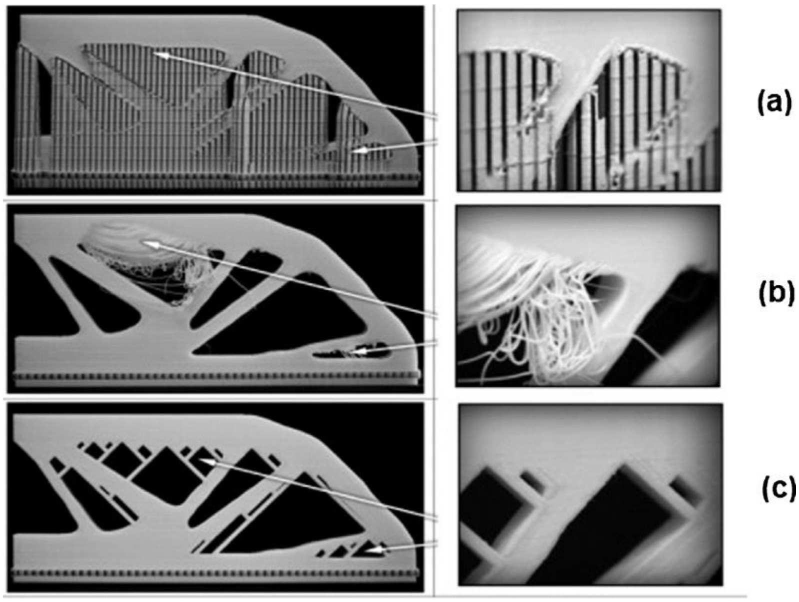


Figure 1.5: Closeup of build support strategies, (a) successful build with support, (b) failed to build with no support, (c) successful build with self-supporting structure (Thompson et al., 2016)

Furthermore, support structures help to improve the thermal distortion caused by the thermal gradients generated during the manufacturing process (Jiang, Xu and Stringer, 2018). These additional structures have to be subsequently removed in post processing operations (Gisario *et al.*, 2019).

Finally, it should be underlined that, in case of shift in production from conventional manufacturing (CM) to additive manufacturing, it is required to *test and certify* the final part mechanical and thermal properties, especially if it is a critical component. This is associated to cost spent in test phase, as well as time to be waited before the component launch on the market. Additionally, the transition would imply the purchase of industrial AM machines, that can be very expensive.

To sum up, a transition to AM technology potentially comes with several challenges. For sure the capability and knowledge to produce properly with this technique should be acquired, and still barriers as “limited variety of materials,” “difficulties regarding the development of new materials” and “insufficient quality of (metal) parts” are preventing AM diffusion (Durach, Kurpjuweit and Wagner, 2017). Despite this, experts are confident that most of these barriers can be overcome in the near future thanks to technological improvement and a decrease in AM production cost is strongly expected. For this reason, the already appreciated trend in AM growing is supposed to be further enhanced, leading to completely exploit all the technological advantages that this technique can bring, both in production but even in inventory management and eventually on the whole supply chain.

Benefits and Opportunities	**Limitations**
Flexibility in design and operations	Anisotropic mechanical properties
Customization	Not high level in roughness
Topology optimization	Tests and certifications requested
Weight reduction and less wasted material	Need of support structure
Assembly reduction	Limited raw materials
No need of tools or moulds	High machine costs
Reduction of lead time and inventory	

Table 1.1: Benefits and Limitation of AM technology.

1.3 Problem introduction: Additive Manufacturing and Supply Chains

It has been underlined that AM can bring numerous advantages in terms of design flexibility, fast response to changing requirements and to customers specific requests. The number of scientific researches on AM is proliferating, especially after the expiring of hold patents, leading to a wider technology diffusion. Despite this, most of them focus on technological aspects, without providing insights on a more managerial point of view, as the AM application in production and its impact on the supply chain.

According to Costabile *et al.* (2016) more than the 50% of the AM literature focuses on 'AM overview' and 'AM technologies' and around the 30% is about AM materials and technological parameters selection to obtain target quality and mechanical properties on the final produced part. Only the 10% of the AM research investigates the technology impact on supply chain cost structures and performance. The scarcity of literature regarding this last topic would prevent a more conspicuous diffusion of AM technology in production, considering possible transitions from Conventional Manufacturing (CM) to AM. In fact, it should be considered that manufacturers always struggle in balancing high capital and production costs, while on the other side they lack of incentives robust enough to evaluate AM production and its impact on supply chain. Furthermore, the AM is a quite new technology: born in the '80, it starts having real production diffusion just in the recent years. This is due to the constant technological evolution that allows manufacturing improvement and costs reduction. Anyway, this technological progress has as drawback that the results reported in literature may be not so up to date and get old really rapidly, leading to misguided conclusions. Therefore, continuous research and improvement are needed to further investigate the AM potential and quantitatively appreciate the impact of the technology. For this reason, the aim of this work is to provide a further insight in AM application in supply chain and production, with particular focus on inventory management. Regarding these areas, three main topics of interest are identified in literature:

- AM production efficiency
- AM impact on single-location inventory
- AM impact on the whole supply chain

Literature reports that AM can bring numerous advantages on the inventory management and eventually on the supply chain, as annual inventory cost reduction, lowered lead time and decrease in transportation, assessing the technology potential over the traditional manufacturing. Despite this, it has been noted that the analysis performed in the listed topics of interest is based on different assumptions. In fact, very often the technological choices suggested for an efficient production are different from the ones applied in inventory and supply chain management. One of the most relevant difference is the production strategy. Indeed, researches regarding AM impact on inventory and supply chain often select the AM technique for on demand or make to order production. This means that the stock held is equal to zero or to a minimum safety stock and manufacturers would wait until demands arrive to start production. On the contrary, studies focused just on AM production reveal that building more pieces in one job would optimize the SLM machine capacity, suggesting that collecting a proper number of demands before starting production would dramatically reduce the AM production time and costs.

Furthermore, literature centred on AM production describes the AM technique in a detailed way, focusing on the different time steps that the technology requires. Indeed, these publications consider the set up and cool down times that the machine requests, but also post processing treatments that the manufactured parts have to undergo, and finally the limited machine capacity that characterize this process. These types of details are frequently neglected in a supply chain perspective, where the AM lead times are considered constant or described by comprehensive statistical distributions, the post processing is not always remarqued and the machine capacity is often infinite.

It has been therefore noted a gap between AM production's field of research and the application of AM in production when the aim is instead the evaluation of the impact on supply chain and inventory management. In this last area, in fact, lots of assumptions are made in describing AM production. These assumptions, that are mostly removed in a pure production contest, may therefore have an important impact on the final choices and results of inventory and supply chain management.

1.4 Aim of the work

In the previous section, it has been underlined a literature gap in term of hypothesises and modelling details when the research focus regards mainly the AM production, and when instead the technology is employed in a wider and managerial context, as the inventory and supply chain ones. In particular, to deal with the first topic, an accurate description of the AM production process is provided, while for the second one a more comprehensive perspective is given, often making assumptions in order to simplify the modelling of a wider scenario. Having noted this discrepancy, it is considered interesting to further investigate if a detailed technology description could also have repercussions on decisions at inventory and supply chain management level. In particular, this topic is of specific interest because AM is often acclaimed in literature as a favourable technology that can bring numerous benefits on inventory and supply chain management, as stock reduction, annual cost decrease, lead time lowering and transportation requirements diminution with respect to conventional manufacturing.

The technology analysed to answer this question is the Selective Laser Melting (SLM), defined one of the most promising and diffused AM technique, able to produce parts effectively used in different industrial field thanks to the obtained final mechanical properties and to the possibility of using different raw materials, one above all, metal powder. In the modelling, all the process stages would be considered, starting from the different time steps that the technique requires, but also post processing operations and limited machine capacity. A specific analysis on the SLM production time modelling is performed, evaluating both detailed and comprehensive solutions to represent it.

AM would be then applied to resupply single-items, single-location inventory in two different scenarios. The first one considers the (S-1,S) inventory policy. This choice, that is often selected in literature, would represent the "on demand" or "Make to Order" (MTO) production scenario: in fact, when using the (S-1,S) model with unitary demand size, an order of unitary size is issued every time a demand arrives to replace the stock. On the contrary, the second scenario would model the lot size production, always preferred in case of efficient AM production. In this case, the (r,Q) inventory policy is chosen: an order of batch size Q is issued every time the inventory level reaches the reorder point *r*.
In order to compare these two different scenarios, a proper AM inventory cost model is developed, tailored for the SLM technique.

In conclusion, this works sets as a goal a systematic investigation of the SLM production process, aiming to establish a framework that could capture the impact of AM system on managing the inventory comparing different inventory policies, and providing an exhaustive and quantitative evaluation of SLM application for a single item, single location scenario.

2 State of the art

2.1 Additive Manufacturing processes

Nowadays, several Additive Manufacturing technologies are available, with differences in the raw materials used and in the specific processes followed, but mostly sharing the same principle for productions: layers of materials are built once per time up to the completion of the final part, rather than subtracting materials as conventional manufacturing does.

To give a general overview on the Additive Manufacturing processes, a distinction is based on the raw material used. In fact, it is possible to find:

- Liquid based AM
- Solid based AM
- Powder based AM

For the first category, the most known processes are:

- Stereolithography (SLA), which is the first created AM technology and it is based on curing photopolymer liquid resins line by line thanks to an ultraviolet (UV) laser.
- Polyjet, which is again based on photopolymerization but in this case the raw material is placed using more precise jetting laser heads, following a technique really similar to normal inkjet printing method, leading to an evolution in term of accuracy and quality.
- Digital Light Projection (DLP): in this case, the resin is cured surface by surface, thanks to the projection of the entire cross-section of the final part by means of UV light.

The general applications for liquid-based AM techniques are prototypes for design analysis and functional testing, prototype tooling and models for conceptual visualization with a competitive build time and cost. Because of the raw materials generally poor mechanical properties, the parts produced are rarely used as final commercial components.

The solid based AM technologies are quite different one from another in terms of processes, but all have the basic common feature of using solid materials to create the final part. The most diffused ones are:

- Fused Deposition Modelling (FDM): this technique is based on the extrusion of a heated solid filament of a thermoplastic material, which is deposited line by line in a semi-liquid state. This technology is one of the cheapest and this characteristic has allowed its diffusion even at amateurs' home. Even thought, the final part quality is quite poor.

- Solidscape benchtop system: it is an evolution of FDM which still uses heated thermoplastic material for creating the final part but, in this case, deposited thanks to two ink-jet type print heads, one for the final part's material and one for the support's material, generally made by wax, which is easier to remove.
- Laminated Object Manufacturing (LOM): the raw materials are paper or plastic sheets that are cut by a laser or a knife and glued together layer by layer.

It is clear that, because of the limited solid materials that can be used and because of the poor accuracy obtained, these types of processes find their applications in prototypes and rapid tooling, but they have the advantages to be very simple, cheap and versatile, with still a great margin for improvement, being the Solidscape benchtop system an example.

The last family of AM technology is the one that uses powder as raw material. The most employed commercially are:

- Inkjet printing binder, better known as 3D printing: this technology works creating layers of materials, mainly plaster (the most used is Gypsum), that are glued together by the deposition of a binder by the printer. This is a really simple and versatile technique, that allows high speeds and it is often used for producing footwear, packaging and medical equipment. Anyway, the surface finish is not so high and, because the powder is just glued together, the final object can be weak.
- Electron Beam Melting (EBM): in this process, a high power electron beam is exploited to selectively melt the powder under vacuum conditions. The process guarantees superior productivity in comparison with Selective Laser Melting or Sintering, but with lower quality in terms of surface roughness and accuracy. Commercially, machine cost and the energy power consumption are high. The main application field is the biomedical one (e.g. prosthesis manufacturing).
- Selective Laser Sintering (SLS): this technology follows a workflow really similar to the inkjet printing binder, but here the raw materials, which are generally polymers, are sintered by means of a laser beam. Indeed, the temperature increment generated by the laser-material interaction is sufficient to induce grain bonding just below the melting point. This technique leads to sufficiently good final parts that require little postprocessing. Anyway, the surface finishing is not so high and the polymeric materials find just a limited variety of applications in the industrial environment.

- Selective Laser Melting (SLM) or Direct Metal Laser Melting (DMLM): the process is really similar to SLS but here the metallic powder is selectively melt by a laser beam. Hence, this procedure allows to obtain higher density components. Considering its capability in terms of final part quality and raw materials that can be used, this technology finds application in multiple industrial fields, as aerospace and automotive.

Among the described technologies, SLM is considered one of the most promising processes that can lead the future AM market (Durach, Kurpjuweit and Wagner, 2017). In fact, it can operate with metals but also ceramics and composites, allowing the production of parts with critical applications, manufacturing tools and spare parts, providing high accuracy, complex geometries (like tools with undercuts and channels for conformal cooling) and an automated procedure.

2.2 Selective Laser Melting technology

Literature reports that SLM has become more and more popular in different fields as aerospace and automotive, defence and medical. This diffusion is due to its many advantages over traditional manufacturing, as the possibility to create lighter components, have a very low material wastage, and, not requiring specific tools or moulds, it provides flexibility, high geometrical freedom and possibility to customization (Nyamekye *et al.*, 2017). The affirmation of this technology in a variety of manufacturing areas and its increasing use for final parts commercially available on the market are the reasons for which DMLM is the Additive Manufacturing technique studied in this work.

2.2.1 The SLM production process

SLM exploits a high-power laser beam to selectively scan materials in the form of powder feedstock. The laser beam is generated in the laser source, which is usually an active fiber laser (such as Yb:glass fiber laser, λ=1070nm) pumped with another diode laser. Then, the laser beam is transmitted through an optical chain to the powder bed surface. The optical chain usually comprises an optical fiber, for beam propagation; collimating lens, for beam collimation; mirrors, for beam deflection; galvo mirrors (in the scan head), for the beam displacement control; F-Theta lens, for beam focalization. The laser beam, transmitted by the optical chain, selectively scans and melts a thin layer of powder according to a 3D computer-aided design model. A new layer of powder is distributed above the previously scanned layers and the process continues until the final object is built. According to the layerwise principle of this technology, the production of complex near net shaped products is enabled. Figure 2.1 depicts the process schematisation.

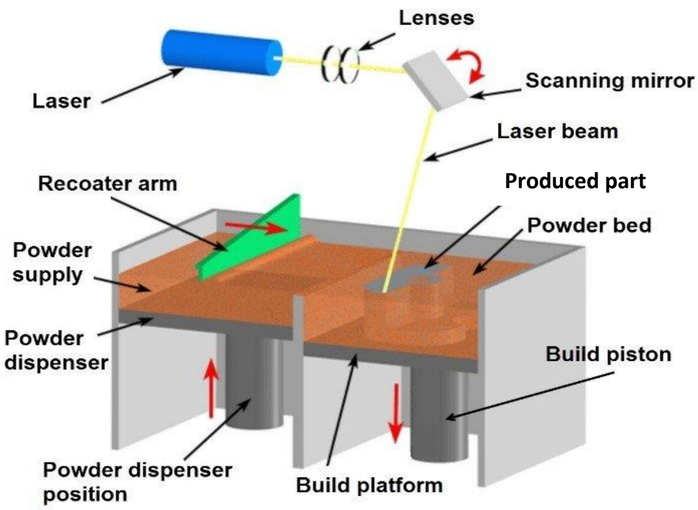


Figure 2.1: Selective Laser Melting process (Marrey et al.*, 2019).*

Being the melting process crucial for this technology, some thermal considerations have to be pointed out. During laser processing, the material is rapidly heated up and melted. The short laser-material interaction times combined with high cooling rates typical of this process, usually induce large thermal gradients within the parts, which influence the microstructure evolution and the residual stresses left after cooling. In particular, the latter may trigger unwanted defects, such as cracking, that are deleterious for the mechanical properties of the final parts. Furthermore, the previously melted and cooled layers of the 3D object continuously experience the thermal cycles owing to the heat dissipation inside the part. This effect may lead to variable mechanical properties over the part height. Therefore, the high thermal load and the sometimes long printing time (some pieces require even a week to be printed) create the need to actively control the temperature. The typical solution is the adoption of baseplate preheating. This technique involves the preheating (such as resistive heating) of the substrate on which components are built, usually up to 200°C/300°C. This operation increases the setup time of the machine before building of an amount of time that depends on the preheating temperature selected and on the height and material of the substrate on which parts are built. Finally, if the preheating is performed, the SLM machine needs to undergo to a cool down stage once the building process is completed, to re-establish the ambient temperature.

Moreover, during processing, it is of a paramount importance to work under proper environmental conditions. Indeed, high oxygen contents and spatter due to melt pool instability may reduce drastically the final part's quality. Therefore gas recirculation, such as argon or nitrogen, are used to work in an inert atmosphere (King *et al.*, 2015).

2.2.2 SLM production time steps

The production process procedures that the SLM technology follows are associated to different time step durations that characterize the technology. Furthermore, it should be underlined that the machine setup and cool down require to reserve a certain amount of time both before and after the printing process. Considering this, it is possible to describe the SLM technique as a sequence of time-steps that have to be followed (Figure 2.2):

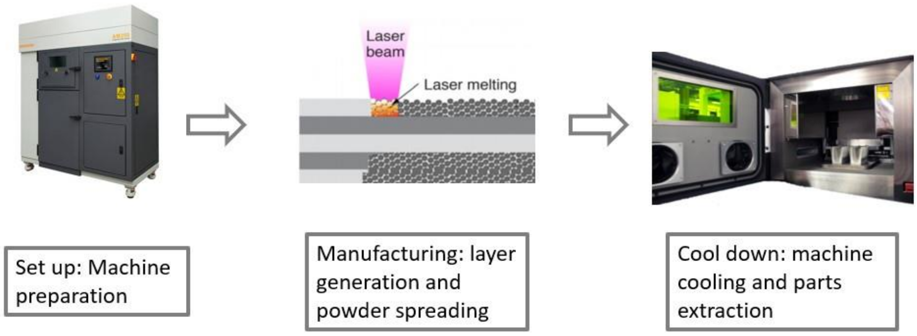


Figure 2.2: SLM production process steps (Images from Renishaw website).

i. *Setup time:* in this phase, the machine is prepared for the build job. The STL parts file is imported and scanning strategy and process parameters are assigned. The substrate is positioned, a new gas filter is placed and the powder tank is refilled with new feedstock, if needed. Finally, the baseplate is preheated and the protective gas injected to create the appropriate inert atmosphere.

ii. *Layer scanning:* this time is spent on selectively scanning the cross-section area of the manufactured part.

iii. *Recoating:* this time step is related to the spreading of the layer of powder for the subsequent layer generation, once the previous step is completed.

iv. *Cool down and finishing operations:* in this time phase, if the preheating is performed, the machine is slowly cooled down to reduce the thermal gradient with the environment. Subsequently, the plate, together with the printed parts, is then extracted. Finally, the machine is cleaned from the powder excess and prepared for the next operations. This would also require changing the filter used for the gas injection and periodically substitute the powder spreading recoater and build plate because of wear out.

After having completed the iv. step, the machine is idle again and ready to start a new job.

Actually, there are still few researches in literature with the aim of creating a generic model in order to proper estimate the time intervals requested for each step in SLM.

Examples are given by Zhang *et al.*, 2015 that used the Grey Theory integrated with a Bayesian form in order to deal with the limited availability of input parameter data set and to express the relationship between the part build time and its manufacturing and modelling factors; Di Angelo and Di Stefano (2011) exploited the Artificial Neural Network flexibility to manage the complexity in defining all the build time factors and their interdependence for the creation of a general model suitable for different AM processes; Zhang and Bernard (2014), Baumers *et al. (*2011), Brika *et al. (*2017) opted for analytical model, collecting data for parameters estimation and focusing only on SLM process; Di Angelo, Di Stefano and Guardiani (2019) extrapolated data and information directly from the GCode instruction used for governing AM computer numerical control (CNC) machines combined with Python code in order to find a quick, efficient and accurate method to determine the production time.

Despite this last cited publication that provides an excellent estimation of the build time (built time estimation error almost null), but that anyway has as a prerequisite the availability of the CAD model of the specific product to be manufactured, the build time estimation error remains quite high, reaching up to 20% (Di Angelo and Di Stefano, 2011). This is due to the high number of parameters that have to be considered at the same time, which have a great variability depending on the AM technology and specific product to be manufactured. Examples are the printing speed, the part size, the layer thickness and the build orientation (Abdulhameed *et al.*, 2019). Therefore, research is still needed to define a more accurate build time estimation model that can provide relevant information about SLM technology or even about a generic AM process: these data could be really useful and noteworthy, for example, in case of AM manufacturing scheduling and planning.

2.3 The batching problem

An interesting topic that emerges from literature is the AM production efficiency. In fact, also for this newer technology and as conventional manufacturing, studies are focused on the production optimization to reduce the total production time, the material wastage and the energy consumption.

Regarding the AM technology studied in this work, SLM, one first important consideration concerns the building time. In fact, every time a job is run, set up time and cool down time should be waited, no matter how many pieces would be printed. This means that if the production is run for just one part, this results in waiting every time that these two phases are performed, allocating their time interval totally to the production time of only one part. Another relevant insight is linked to the time required to spread the powder for every layer: if the part manufactured is just one and small with respect to the building plate, and considering that the SLM machine would spread the powder anyway all over the building plate, most of the time would be spent to distribute the metal powder in areas that would be not printed. Knowing that the AM parts are made of a quite high number of layers, this leads to a considerable loss of time (Rickenbacher, Spierings and Wegener, 2013; Piili *et al.*, 2015).

In addition to the mentioned considerations about a possible waste of building time, some studies underline the disadvantages in terms of energy consumption in case of unitary piece manufacturing. In fact, the machine set up and cool down, responsible of the control and regulation of the build chamber temperature, as well as the creation of a proper atmosphere by means of specific gas injection, have specific energy requirements for a quite long period of time that would be associated to the production of just one piece. An example is provided by Piili *et al.*, (2015) research: they reported that filling the building chamber with the maximum number of pieces possible could lead to a reduction of the energy consumption cost per piece ranging between the 80% and 90% because of the distribution of the fixed setup and cool down requirements on all the pieces in the building chamber.

Furthermore, it should be remembered that in SLM the majority of the unmelted powder in the building chamber can be recycled, but a certain percentage is wasted. This is due to the fact that the really high temperature produced by the laser source can partially melt also the powder around the surface to be scanned and the high temperature atmosphere can damage the remaining raw material. Additionally, during the process it is possible that spatters are generated, contaminating the unmelted power. Finally, part of the powder not melted is lost during the building plate extraction, powder recycling and cleaning operations. For this reason, it is generally considered a scrap factor in order to estimate the percentage of powder not melted that cannot be reused. It can be pointed out that if many pieces are printed, the not melted powder is less, and therefore also the wasted material. Knowing that the metal powder used in SLM is quite expensive, this reasoning would lead to a sensible reduction of the cost of the scrapped raw powder.

The described considerations generate as a consequence the idea of starting a production by batches: this would mean to collect a certain number of demands and manufacture the different pieces all together, placing them in the same building chamber. This would lead to allocate and distribute the set up and cool down time on a higher number of pieces, as well as optimizing the time step related to the powder spreading: if more parts are present on the same building plate, the fixed time requested by the recoating blade to distribute the material in every layer would be now divided and allocated on a higher number of pieces. In addition, the energy consumption in the set up and cool down phases would be shared by all the parts present in the machine, reducing the total energy request for every unit. Finally, the percentage of scrapped material would be reduced.

The literature reports the advantages of Additive Manufacturing batch production in terms of both building time and energy consumption saving, leading finally to a reduction of production costs. Nyamekye *et al.* (2017) studied the energy requirements and efficiency in case of Powder Bed Fusion (PBF) printing, considering the whole phases of the process: preheating, printing and cool down. They obtained that the energy consumption per piece decreases if several components can be built simultaneously. Piili *et al* (2015) compare the SLM production in case of one part per build and in case of the completely filled build platform. They report that the time saving in case of batch production is around the 80/90%, compared to the unit production one. Rickenbacher, Spierings and Wegener (2013) underline that printing more than one part in the same job can lead to a reduction up to 40% of the time required for powder spreading together with the one spent in scanning strategy and auxiliary processes, as parameters loading and adjusting. In fact, because of the simultaneous building of parts, the coating time can be divided on the number of parts printed, and more efficient scanning strategies can be used. Baumers *et al. (*2011) focused their research on the evaluation of possible energy saving considering batch manufacturing. Comparing the energy requirements for printing a single part per time or a number of parts in order to saturate the machine capacity, authors found savings of 21.7% and 28.06% for each part built in case of batch production, depending on the SLM machine used. This is motivated considering that extensive energy investments are related to set up and cool down phases that are shared in case of multiple parts manufacturing, leading to a more efficient machine utilization.

Another interesting consideration that batch production can allow is related to the cost calculation. One of the first attempt in filling up the build plate in order to utilize its maximum capacity was provided by Hopkinson and Dickens, (2003). They studied the cost related to STL, FDM and SLS which share a similar process to SLM, placing a large number of parts in each build job to minimize the final part cost. Because the costs were allocated already considering the maximum capacity utilization, the results suggest the absence of a relationship between unit cost and build volume utilization (Figure 2.3).

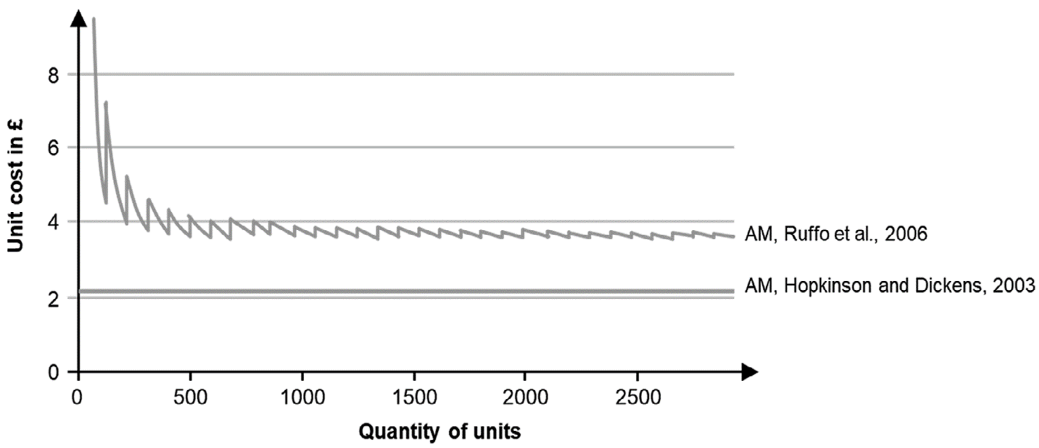


Figure 2.3: Constant and quantity dependent unit cost models in the literature for Laser Sintering technology (Baumers and Holweg, 2019).

A further development in the model was provided by Ruffo, Tuck and Hague (2006) that presented a detailed cost structure for SLS considering the impact of sub-maxima capacity utilization on unit cost, allowing in this way the possibility to establish a relationship between the number of parts in a job and the production cost. From Figure 2.3 it is possible to note that the unit cost decreases adding more parts in the building chamber. The peaks are showed whenever it is requested to add a new vertical layer to insert a new part in the job. Overall, this model development proved the effectiveness of batch production, with a cost reduction when the AM machine and the different overlapped strata printed are saturated.

Similar considerations are suggested by Baumers and Holweg (2019), that studied different packing configurations in order to fully exploit the building chamber volume in case of SLS technology. They observed a unit cost decrease of 79% in case the horizontal plate is fully occupied (Figure 2.4b), and a 52.2% considering possible vertical stacking of the piece produced (Figure 2.4c).

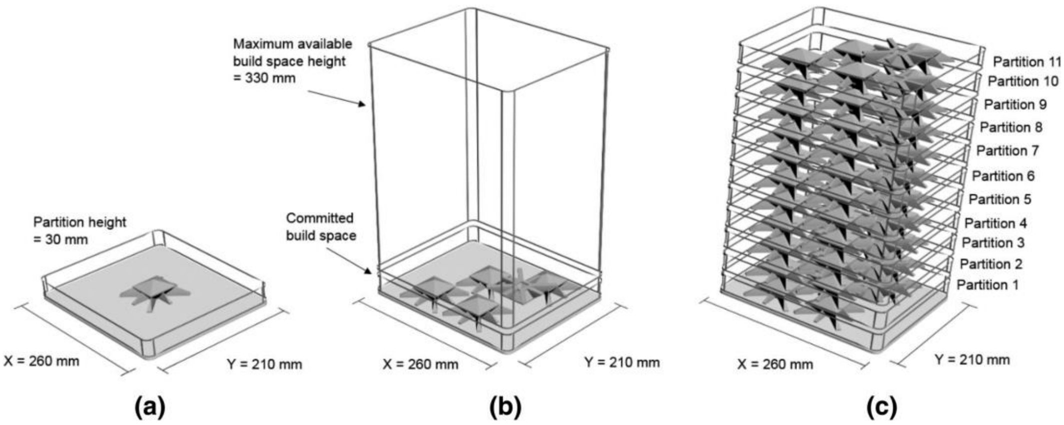


Figure 2.4: Build volume allocation and capacity utilization in the additive manufacturing process (Baumers and Holweg, 2019).

The mentioned decrease in unit production cost considering batch printing stimulated an increasing interest in literature in a proper organization and optimization of the AM batching production. Examples are given by the works of Fera *et al.* (2018), Li and Zhang (2018), Chergui, Hadj-Hamou and Vignat (2018) and Kucukkoc (2019) that have as a common general aim the creation of algorithms able to determine the optimal batch size and solve scheduling problems. In this way, the utilization of the machine can be maximized, the production makespan is reduced at minimum, having as consequences the total production costs reduction and an increased efficiency in manufacturing operations.

Literature therefore suggests that, in order to maximize the SLM machine utilization, the strategy of printing many pieces inside the building chamber is successful. In fact, in this way the set up and cool down times and the time for spreading the powder can be allocated to the number of pieces produced, optimising the total production time.

Furthermore, energy savings are reported, as well as a reduction of the unitary production cost. Therefore, it can be stated that batch production is the promoted choice to have an efficient AM production. From a technological point of view, it is reminded that anyway, placing many pieces in one job introduces the problem of maintain high part quality: in fact, it is possible that spatters generated during the printing phase can contaminate the unmelted powder, creating possible defects in the components to be built. In this respect, further test on the final batch properties should be performed to confirm the parts reliability.

2.4 AM for inventory management

Managing inventory in complex supply chains scenarios is a multifaceted challenge. Forecasting product demand, estimate lead times, predict components weekly request are some of the problems that inventory experts have to deal with in order to set adequate stock levels. These issues are further enhanced when working with parts with erratic and unstable demand and long lead times. The often chosen solution to overcome this problem is to produce following the make-to-order policy (Chen *et al.*, 2019). Anyway, this can cause a loss of protection in case of unexpected issues, having as consequence to carry extra-inventory as backup items. Although this last solution can minimize the risk in case of disruption or request of emergency parts, it significantly increases the capital tied up and the parts can become obsolete. An example of a category of components that are subjected to such problems are spare parts: these items are used to maintain equipment or original product in operating conditions and are requested both in case of preventive maintenance and corrective maintenance. This means that parts demand can occur when the component reaches its expected life or when it breaks down. This leads to intermittent and unstable demands, causing difficulty in proper managing the inventory. The consequence is a quite high inventory cost or an excessive amount of spare parts stock. Airbus maintains a 36,000-square meter warehouse for spare parts in Hamburg, Germany (White and Lynskey, 2013). The United States (U.S.) military reports spending in 2009 $194 billion on logistics operations and its spare parts supply chain with $104 billion spent on supply, $70 billion budgeted for maintenance, and $20 billion budgeted for transportation (Khajavi, Partanen and Holmström, 2014). Spare parts, which increase inventory costs, may be used infrequently and can become obsolete due to innovation (Liu *et al.*, 2014). The U.S. Navy estimates the cost of obsolescence to be $750 million annually (Khajavi, Partanen and Holmström, 2014). During a recent conflict, the Osprey V-22 program had 12 aircraft deployed and spare parts for 36 aircraft on-hand; however, only 13% of those parts were needed (Gertler, 2010; Togwe, Eveleigh and Tanju, 2019).

This growing cost of inventory stock is a significant operational challenge that firms can manage by taking innovative steps. Hence, the promotion of an improving strategy that allows flexibility and an agile approach, capable of delivering components with reduced lead time and suitable for lumpy and unstable demand is pursued. For this scope, Additive Manufacturing is seen as a possible effective solution. In fact, its capability in producing complex and high customized objects on demand would facilitate firms in delivering what the customer needs when he needs it. Therefore, AM technology's potential to revolutionize manufacturing as underlined in scientific research (Sirichakwal and Conner, 2016) and its unique capabilities could offer unprecedented opportunities for firms to improve their inventory efficiencies.

Regarding the possible application of AM in inventory management, literature offers different perspectives and research. The most significant ones for the scope of this work are reported in the following paragraphs.

AM impact on lead time, holding inventory and stockout risk

Sirichakwal and Conner (2016) studied the implications of the introduction of AM for resupply spare parts inventory in an aeronautical scenario. This industrial field is often defined critical for its lumped and uncertain demand of components, and also for the request of emergency parts that, if missing, lead in the airline industry to the grounding of airplanes. Authors studied how AM could deal with these issues. They used the (S-1,S) inventory policy for inventory replenishment using AM in production. Furthermore, they considered a constant lead time for the spare parts to arrive to the stock. Authors observed that the introduction of AM allows a reduction in holding cost. This is due to the possibility to stock less expensive and bulky raw materials with respect to the conventional manufactured final parts. The consequence is having the opportunity to store more at the same cost, leading to a reduction in the stock out risk. In addition, they affirm that AM can decrease the lead time, providing a reduction of the holding inventory. This may have as drawback an increase of stockout risk, and therefore a proper balance between the combined contributions of cost and lead time reduction in suggested. They affirm anyway that under any circumstance, the total inventory cost is reduced thanks to the positive effect brought by AM introduction.

Parallel use of AM and CM for reducing inventory cost

Knofius, van der Heijden and Zijm, (2019) used stochastic dynamic programming and numerical experiment to assess if a transition to AM becomes profitable for low volume spare parts business. In particular, they used as a case study a radar system component with a low demand that may require even more than half a year as replenishment time with CM techniques. They obtained that moving to AM leads to a cost reduction of 35% considering a service horizon of 8 years. This is due to the possibility of discarding CM tools and, thanks to a just in time production and therefore a relatively short replenishment time, to a reduction in stock level and stockout risk. In particular, they pointed out that saving holding cost is the primary benefit of the transition to AM, caused by the reduction in lead time and in the backorder probability. An additional interesting consideration is linked to the possibility of dual sourcing, so AM and CM used in parallel, as a viable solution in case of high initial AM cost per piece. In this scenario, they showed that CM would be used for regular parts, and AM for emergency ones: this cause the reduction of the stock while the high AM purchased cost is maintained within limits, until the new technology becomes completely profitable.

Switchover from TM to AM for spare parts resupply

Heinen and Hoberg, (2019) aimed to extend the literature perspective on the impact of AM introduction to manage spare parts inventory in the aerospace sector. In particular, they considered a very large portfolio of spare parts to understand the volume of parts that can be produced with AM instead of CM to reduce the total inventory cost. For modelling the inventory, they used (r,Q) policy for Traditional Manufacturing (TM), assessing that conventional manufacturing often requires set up cost, while they applied the (S-1,S) policy for AM, saying that this technology would request a lower inventory and neglecting the set up cost. They underlined that this choice can be a limitation for the model, even because the (S-1,S) model implies that the parts are produced one per time, not maximizing the machine capacity. Furthermore, they considered both the production technologies having the same average constant lead time to resupply the stock and cost approximations were assumed. Despite the underlined hypothesis and limitations, authors found that manufacturing shift from TM to AM is suggested for a percentage of stock keeping units, having as a consequence the reduction of total inventory cost and the possibility to more flexible replenishment options.

Set-up time consideration for moving from CM to AM production

Another interesting work is provided by Cestana *et al.*, (2019). Authors investigated how the duration of set-up time can favour the shift from CM to AM in case of managing a single-location, single-spare part inventory. They applied a continuous (S-1,S) inventory policy to manage both the CM and AM stocks, considering as target product the slow moving aerospace spare parts. In order to model the resupply time, Markov Chains were applied, even if with some hypotheses as infinite machine capacity, negligible AM set-up time and an overall AM production time estimated longer with respect to the CM one. Authors demonstrated that AM outperforms CM when CM set-up times are long, which often happens in case of a production that must adapt to low and lumpy demand rate as the one for spare parts. On the contrary, if the AM production time is much longer than the CM one, CM still is the optimal technology to minimize the total inventory cost, having an overall resupply time which is lower than the AM one, despite the set-up requested. Overall, they concluded that the optimal stock levels can benefit from the AM application, being lower than the CM ones, and AM can lead to considerable savings considering the total company portfolio of spare parts available.

Moving to a decentralized AM production and impact on inventory management: introduction of AM added values in cost computation

Togwe, Eveleigh and Tanju, (2019) evaluated the impact of the introduction of AM to manufacture aeronautical spare parts: their aim was to understand the consequence on the inventory management of an increasing percentage of components in the product mix produced by this technique. In order to proper evaluate the advantages that AM can bring, they introduced a very relevant concept in terms of technology evaluation and inventory costs. They suggested that, generally, literature is focused on comparing conventional and additive manufacturing using activity base costing, and therefore considering mainly machine, material and personnel cost. They underlined that categories as form factor, proximity, obsolescence and lightweight are rather considered, making AM aptly penalize in a side-by-side comparison with CM. They therefore believed that it is crucial to recognize the added value brought by this new technology, rather than focusing only on the added cost. For this reason, they calculated the annual cost rate relying on this consideration, providing to CM a higher rate in fungibility, form factor and obsolescence. To analyse the problem, a simulation model was created. In particular, the authors considered a batch reorder policy, meaning that the inventory level is continuously checked and orders are triggered when the inventory level reaches a specific reorder point. Lead times were modelled using triangular distributions from data collected by experts. Authors obtained that increasing AM percentage in spare parts allows a sensible reduction in the overall spare parts lead time (about 33% when 35% of spare parts are produced with AM). In addition, considering the value added for the computation of the annual holding cost rate calculated considering the benefits that AM would bring that very often are not quantitatively recognized, and the advantages provided by the technology, AM results in a less capital in inventory tied up and a reduced total annual inventory cost.

In depth analysis of AM production process and machine capacity

Zhang *et al.* (2019) used discrete event simulation to appreciate the effectiveness of the very acclaimed make to order production that additive manufacturing would allow, comparing it with the traditional warehouse strategies in case or spare parts demands. They performed an in-depth analysis on the PBF processes as SLM and SLS, examining all their production stages to obtain a more accurate production time estimation. The production steps listed are the setup, printing and cool down of the machine, described by means of statistical distribution. In addition, they considered the possibility of attribute priority rules to the spare parts in arrival, giving precedence to the emergency parts requested. It is worth pointing out that this work considered and studied the possible queue formation of parts waiting to be processed by the AM machine, not hypothesizing an infinite machine capacity nor a generic and comprehensive production lead time. They obtained interesting results: because of the limited AM machine capacity, if the demand rate is high, queue formation is really likely to occur, leading to an increase of penalty cost because of long waiting times. In particular, emergency spare parts are often prioritized, not giving space to the regular spare parts production. In addition, by changing the demanded spare part size, they observed that AM operations perform the best when the majority of the parts is of small size. This is because the small spare parts have a higher likelihood to enter the AM system in the volume filling process, rather than the larger one. This can be appreciated in Figure 2.5 that underlines that the small parts are always the most frequent produced, even varying the demanded part's volume.

They finally obtained that, because of the penalty cost due to the limited capacity of the AM machine, the on-demand production can be more disadvantageous than the classical warehouse strategy, especially for the larger parts. This is an interesting result because it is in contrast with the general literature outcome that reports AM as a just in time technique. This may be due to the lack in model completeness, especially regarding the limited AM machine capacity and the too generic lead time estimation, underlining in this way the importance of their consideration.

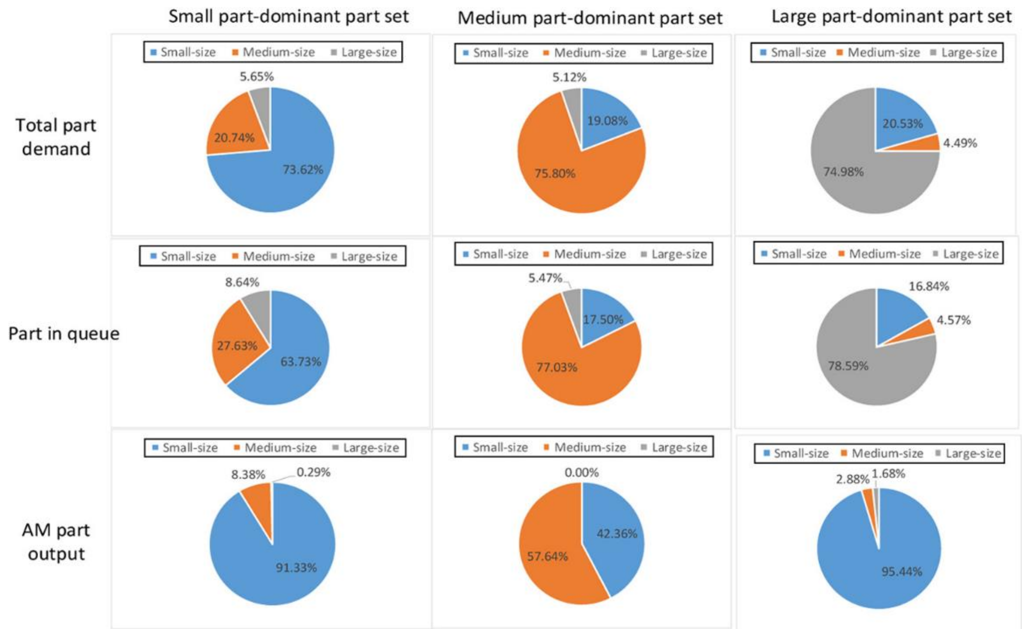


Figure 2.5: Part size characteristics in the AM operations under different spare parts size attribute (Zhang *et al.*, 2019)

AM for Maintenance, Repair and Operations: analysis for target parts

Chen *et al.* (2019) discussed the challenges in Maintenance, Repair and Operations (MRO) parts inventory management linked to erratic demand, long and irregular lead times and risk of obsolescence and how Industry 4.0 technologies can be leveraged to address them. They categorized MRO parts in four main groups, depending on their demands, lead time, usage and volume request. They demonstrated that each group may have different specific inventory policy, helped by the aid of Industry 4.0. In particular, they qualitatively suggest that Additive

Manufacturing could be useful for those parts with long lead time low volume and moderate cost. In fact, considering that in these years the AM cost is dramatically reduced, the technology could be used for make to order production, reducing lead times and avoiding to an excessive increase of stock levels and obsolescence.

AM on raw material inventory impact, considering its actual industrial degree of adoption

Kunovjanek and Reiner (2020) studied the impact on the raw materials inventory in case of AM adoption with respect to CM and pointed out possible impact on supply chain. They firstly underlined that AM introduction can lead to a reduction of raw materials requirements with respect to subtractive manufacturing, decreasing the waste and making production leaner. In addition, change in design introducing lighter lattice structures and the design freedom that this technology provides allow to reduce further the material request. This has as a consequence a possible reduction of the material inventory. It is interesting that authors considered in their research also the degree of adoption on AM. In fact, they said that the technology diffusion is still behind expectation, but the interest in it keeps increasing. This limited diffusion is also considered in the system dynamic approach used to obtain results, in order to have a more realistic perspective. Despite the low degree of adoption in AM use in production smoothens the results, they obtained that the materials inventory can be still reduced by 4%. In addition, they remarked that higher reduction can be obtained considering not only the manufacturing process, but also the interaction with all the stages of Supply Chain: if AM can allow make to order production, inventory can be further decreased having repercussion on work in progress and raw materials. They also predicted that the most total reduction will take place between 2019 and 2039, considering the technology time trends. In addition, the inventory lowering would increase the perceived usefulness of the technology, leading to a cascade of positive effects on its adoption and accelerating its impact on raw materials supply chain.

In conclusion, literature provides a quite broad perspective on the use of AM to improve the inventory management of a single supply chain location, pointing out the ability of this technology of reducing the holding costs and the lead time thanks first of all to a faster and more agile production, but also to its flexibility, that permits the disposal of old tools and the production of a great variety of product. Anyway, what appears from the mentioned researches is that, because of the complexity of the inventory system, lots of hypotheses are requested: lead time supposed constant, demand difficult to predict, infinite machine capacity and AM set up times not considered are some of them. Sometimes these assumptions can be a limitation for the model results, as suggested by Zhang *et al.*, (2019), that, proper considering AM machine capacity and developing a detailed description of its production process, provided a higher in-depth view of the system, better underlining which parts and scenarios could be suitable for AM production. It is therefore suggested that a detailed and advanced description of the AM production can provide further insight on inventory management, revealing consequences often not evident in more generic models.

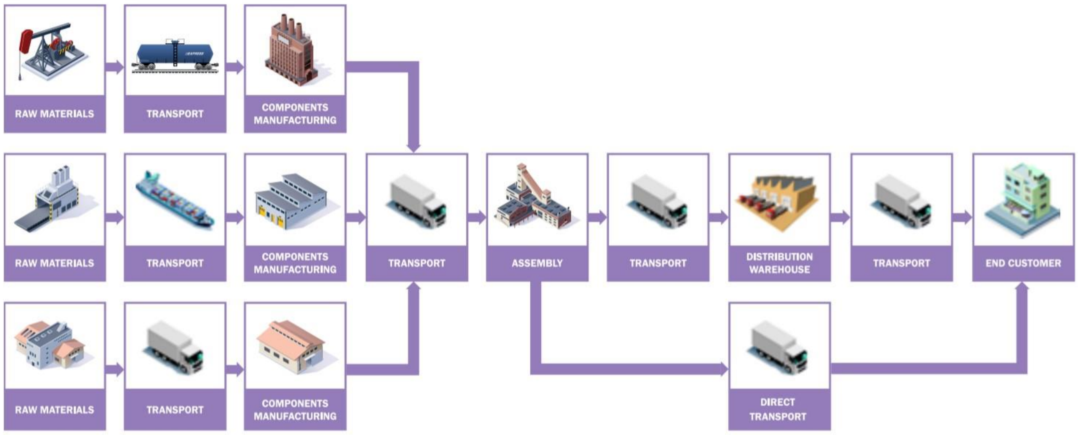
2.5 AM impact on the Supply Chain

An interesting point is to exploit the potential of AM not only on a single warehouse, but also on the whole Supply Chain (SC). In general, supply chain is designated as the network of organizations that are involved through upstream and downstream linkages through the different processes and activities which produce value in the form of products and services delivered to the ultimate consumer (Mentzer *et al.*, 2001). In other words, SC is an integrated process where the raw materials are manufactured into final products, then delivered to customers considering their requirements (Benita M. Beamon, 1999). Nowadays the biggest challenge in SC management is the efficient and effective delivery of products to fulfil customers demand and maintain high service level. This is not so straightforward, considering the volatility of the demand, the risk associated to possible disruptions or delays in supplier production, the stock out risk, the issues or slow-down in transportation and the high costs associated to it because of market distance.

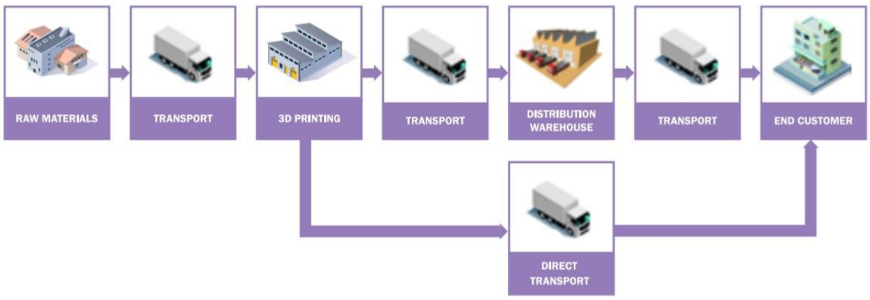
To overcome these obstacles and make SC leaner, experts considered in these recent years the opportunity of introducing AM and appreciate if the technology would be fungible for the scope. In fact, AM is known for its high flexibility and capability of producing on demand, giving the possibility of obtaining high fulfilment rate (Li *et al.*, 2017), requiring less material usage and resource wastage (Liu *et al.*, 2014). The focus of the research is generally related to the comparison of AM techniques with the CM ones, considering if this recent technology can bring as consequence SC total costs reduction. Some of the benefits underlined in literature are reported in the following paragraphs.

Leaner Supply Chain

In order to appreciate the impact of the AM introduction on the total SC cost, the first element to start the discussion is the simplified production process that the technology establishes. In fact, AM allows to redesign the part arising the possibility of making one single part instead of assembly different components as the conventional manufacturing requires. This can have a major impact in terms of reduction in labour input, tool and machining centre request, being AM, in addition, an automated technique. Furthermore, the opportunity of not requiring the assembly anymore has a direct consequence on Supply Chain Network: the number of tiers and echelons is reduced because less different raw materials are requested, and there is no need of leaning on a vast network of suppliers to find and get all the components that the final part require, being AM able to produce the final component as a whole (Figure 2.6).



a)



b)

Figure 2.6: a) Traditional Supply Chain network versus b) Additive Manufacturing Supply Chain network (Janssen et al., 2014).

Following the same reasoning, the higher level of independence from suppliers and the SC simplification have also an impact on the inventory stocks. In fact, it is possible to reduce the number of components and parts hold, as well as diminishing the work in progress. Furthermore, the safety inventory stored to be protected from possible disruption at supplier level that could lead to a higher risk of stock-out, can be reduced thanks to the autonomous and just in time production that AM introduces. Additionally, one important concept that is very often emphasized in literature is the possibility to produce on demand: it is in fact supposed that AM would allow to reduce to zero the on-hand stock, embracing the chance of producing the components following a make to order policy (Liu *et al.*, 2014; Li *et al.*, 2017; Roca *et al.*, 2019). Overall, these consequences have an economical SC impact, with the reduction on the total inventory cost. Finally, it is underlined that these results also have impact on administrative costs, that, thanks to the simplification underlined, can be reduced (Li *et al.*, 2017).

Lead time and carbon footprint reduction

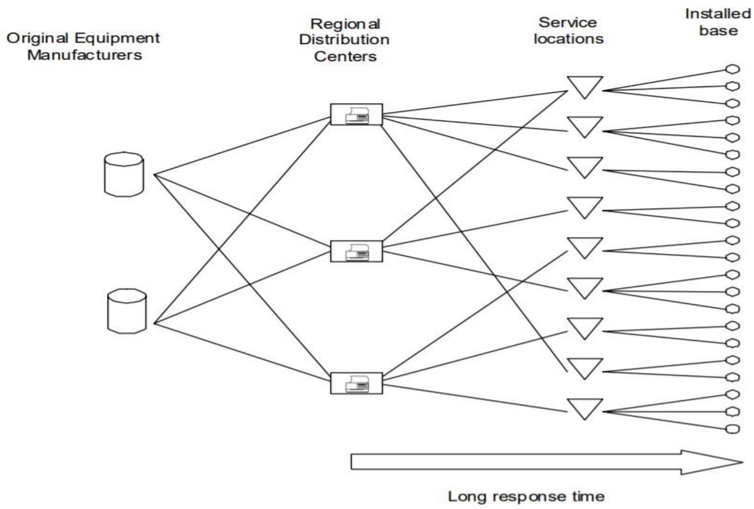
Another outcome of the more streamlined SC that AM introduction would provide is the reduction of the lead time. Researches point out that producing using AM offers the possibility to deliver parts just in time, cutting all the waiting stages related to materials and components procurements from suppliers (Liu *et al.*, 2014). In addition, the AM supply chain redesign leads also to a reduction of the transportation levels and times. Li *et al.*, (2017) compared in their work the traditional SC and a redesigned AM one, pointing out that the main cost that affect the first one is the transportation one, accounting for the 57% of the total SC cost, that could be reduced up to the 32% in case of AM introduction. The direct consequence of this transportation reduction is a decrease in SC carbon footprint. Considering the whole SC emission, both of production processes, manufacturing and transportation, authors found that AM introduction can reduce up to 76% the carbon emission. This is due for sure to the decrease in transportation, but also considering that AM has the potential of lowering the environmental impact: less raw materials are requested and waste and scraps can be reduced to a minimum, leading to a leaner manufacturing (Mashhadi, Esmaeilian and Behdad, 2015). Another interesting perspective is provided by Ashour Pour *et al.*, (2017). They published an innovative point of view comparing AM and TM inventory management and SC performance considering lot production and orders lots. This approach is quite different from the others presented in literature, more focused on the just in time production and not providing insights on AM production process and limited machine capacity. Researchers demonstrated that batch AM production can be advantageous with respect to the TM one in terms of inventory and transportation cost, underlining that an optimal production lot size can be found in order to minimize the total expenditure. Rogers, Baricz and Pawar, (2016) qualitatively described the impact that AM introduction would have from a service point of view. They underlined that AM allows build to order policy to promptly satisfy customer demands, launching the production only when customer requires specific products. Nevertheless, they considered that, to maximize the build chamber and machine utilization, one solution is to collect different orders and print them all together in one job. This scenario would be feasible considering what authors call “facility service”: this type of service provider would satisfy the customers’ requests to transform 3D modes into 3D printed objects accumulating different demands for example by means of online platform, and manufacturing the different components requested in one job.

AM flexibility and agile production systems

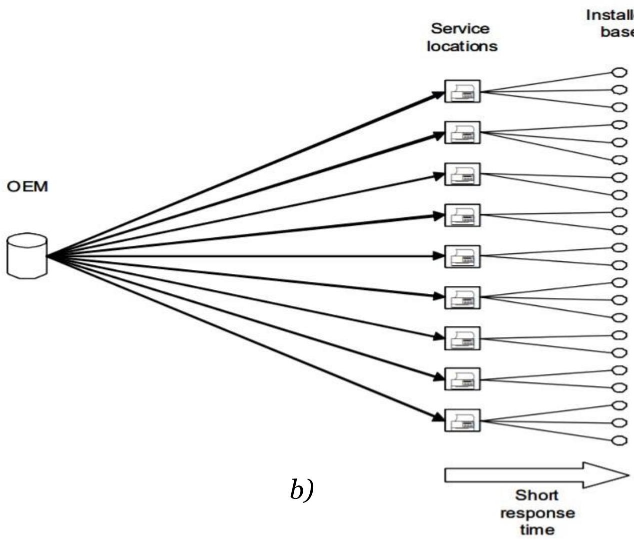
One important benefit that AM introduces is the flexibility, defined as the ability to react promptly to changing requirements (Ivanov, Das and Choi, 2018). This advantage that AM shows leads to the possibility to fast change the product design and allows the product customization on large scale. Terms as “mass-customization” are introduced, and a higher integration of customer into the manufacturing, providing the option of choosing individual features, are possible (Durach, Kurpjuweit and Wagner, 2017). Furthermore, the lead time decrease together with AM flexibility would promote the growth of an agile production system. An agile production system is recognized as a system with capability to meet the rapidly changing demand of the market both in terms of the volume and the variety of products (Mashhadi, Esmaeilian and Behdad, 2015). This system is fairly compatible with make to order strategy, that can follow rapidly the market demands changing and the quick responses often requested. Ryan *et al.*, (2017) reported in their work that AM would be indeed suitable for engineering to order and make to order strategy, identified as the 81,6% most probable AM applications scenarios, because of the facility of producing always new design with minimal setup costs.

AM and centralized or decentralized SC perspective

Another point of discussion that is often found in literature is the evaluation of AM application in two main different scenarios: the centralized one, where regional distribution centres owns the AM machine and products are then distributed in all the SC by dedicated transportation systems, and the decentralized one, where multiple AM production centres are collocated directly at service locations, nearer to customer demand (Figure 2.7). In order to discern which of the two can be the most cost effective scenario and to optimize the locations selection where to collocate the AM production centres, researchers focused their attention in developing proper algorithms, taking care of the balance between the total demand and its geographical allocations, as well as of transportation and production costs (De Brito *et al.*, 2019). The results obtained demonstrate that the decentralized solution is the best choice to decrease the transportation costs, meet the demand as quick as possible and reduce the inventory and holding costs at minimum (Li *et al.*, 2017). An additional factor that could promote the decentralized solution is highlighted by Jia *et al.*, (2016). They considered the utilization of the Additive technique applied to chocolate production. The advantage that AM can bring is the possibility to customize the product following the customer's desires directly at the retailer shop, leading to revenues increase. Furthermore, thanks to the digitalization of the technology, customized online orders where the final consumer can design his own desired final product are also possible, leading to a business improvement. Further insights are provided by Ghadge *et al.*, (2018), who studied the impact of a shift in production from CM to AM with the aim to reduce the total inventory cost in the aerospace industry. In fact, they underlined that AM could simplify the SC producing the part on demand and in a decentralized way, installing an AM machine at every service location. This would lead to a reduction in the resupply time with respect to CM, identified with a more complex SC, requiring at least raw material suppliers, first tiers suppliers that sub-assembly parts, and original equipment manufacturer that finalize the product. They obtained a significant reduction of the inventory level in the AM case, being the AM inventory about the 25% of the mean CM one. This is mainly due to the decrease of both the production time and the total resupply time all along the SC brought by the introduction of AM technique. It should be underlined that, anyway, limitations in this model can be related to the fact that infinite production capacity is hypothesized and no specific details about the AM or CM technologies analysed are provided.



a)



b)

Figure 2.7: a) Centralized Supply Chain configuration, b) Decentralized Supply Chain configuration (Holmström et al.*, 2010).*

An innovative point of view is provided by Chekurov *et al.*, (2018) who interviewed managers and manufacturers experts in the spare part sector to appreciate their willingness in adopting Digital Spare Parts (DSP). This term is used to represent the introduction of AM in production. It would allow the digitalization of spare part portfolio, reducing the need of high inventory stocks to fulfil demands and the risk of obsolescence. Experts agree that AM can reduce the repair and delivery time, decrease the inventory costs, and limit the material wastage and emissions. They suggest that a possible solution is the "replacement of central storage with model database, and the introduction of distributed manufacturer" to be able to satisfy requirements faster than with a traditional SC. Despite this, some issues that do not emerge with CM, arise. In fact, if AM allows to send digital files instead of physical components or materials, dramatically reducing the lead times, the need of intellectual property protection become an issue. Digital files security and advanced software development are required to make this transformation possible.

Despite the possible positive outcomes brought by the adoption of the decentralized solution, some researchers underline that the scenario choice is not straightforward. In fact, some factors can be crucial in the decision. One important discriminating parameter is the purchase cost: to justify the machine investment for a decentralized solution, related to the acquisition and allocation of different AM machines to cover the parts' request in different and far away locations, the demand should be high and stable. If this is not the case, the risk of bad machine capacity utilization may occur, leading the centralized scenario to provide better SC performance (Liu *et al.*, 2014; Roca *et al.*, 2019). In addition, the administrative costs can increase due to the production dislocation, and may affect the results, even if in small percentage (Li *et al.*, 2017; Roca *et al.*, 2019). Furthermore, Roca *et al.*, (2019) showed in their study on the production of non-critical spare parts for aerospace industry in USA, that the post-production costs that AM requires are a limit for the production decentralization. In fact, due to the high expenses incurred in post-processing equipment, the decentralized scenario is excluded, and only a reduction of these necessary costs would make the decentralized choice preferred.

In conclusion, literature shows how additive manufacturing can be an excellent opportunity for reducing the costs and the risks associated to the traditional SC, improving the service level and decreasing the lead time. Attention should be paid to the particular application environment, opting for centralized or distributed SC scenarios, keeping in mind that many different factors have to be considered to provide an optimal solution evaluation. Finally, because AM is a technology that progresses with a rapid pace, its scenario applicability should be regularly checked, updating it with new materials availability, machine efficiency and costs evolutions.

2.6 AM application in different production contexts

In the recent years, a growing interest about AM is perceived. Nevertheless, the research on the impact of this technology on SC is still limited and immature. It is clear anyway, that much effort is put in order to represent and hypothesise what type of scenarios could fit the best AM applications in production to eventually have positive impact on inventory management and the whole supply chain.

The most interesting literature papers on this field are summarised in Table 2.1 and classified depending on the main topic they deal with: AM impact on single-location inventory management, AM and production efficiency and AM impact on the whole SC. In particular, those papers that conduct quantitative analysis with respect to the more qualitative ones have been selected.

It is interesting to observe some particular drivers with which the different analyses are performed. One important main difference is the production strategy selected. In case of the AM application for resupply inventory or to substitute CM in the SC, the preferred production choice is the make to order one, meaning that every time a demand occurs, an AM job is scheduled. This strategy often corresponds to the (S-1,S) inventory policy. On the contrary, for a more detailed production perspective, the lot production is always selected. In this case more pieces are placed all together in the AM building chamber to maximize the machine utilization.

Other important differences between the papers dealing with production efficiency and the ones centred on AM for SC are related to the technology modelling. In fact, in case of AM for supply chain or inventory management, very often approximations on the lead time or resupply time are done, considering it described by a generic statistical distribution or even by a constant term. Furthermore, the setup and cool down times as well as post processing operation are often ignored, and the machine capacity is considered infinite. These types of assumptions are instead removed in case of studies for AM production, where a more detailed model about the technology is provided.

It is therefore evident that a discrepancy regarding the AM production modelling is present, showing general approximations in case of AM application in the SC. In all likelihood, these assumptions are justified considering that SC modelling is complex and requires to consider many parameters and factors.

Nevertheless, it could be interesting to appreciate the impact of a more detailed description of the AM technology on the SC, and in particular on the single-location inventory management, and possibly evaluate if this modelling choice would have consequences on the production strategy. In fact, it is worth noting that Zhang *et al.*, (2019), providing a detailed description of AM technology, pointed out that the make to order production can be inefficient with respect to maintain a certain inventory stock. In the same way, Roca *et al.*, (2019) who again modelled AM considering machine capacity and post processing requested, asserted that considering on demand production can limit the AM decentralized solution, being the AM post processing requirements really expensive, giving in this way additional insight about the technology application.

Paper	**Main topic**	**Quantitative Analysis**	**Qualitative Analysis**	**MTO Production**	**Batch Production**	**Constant Lead Time**	**Set up**	**Limited AM Capacity**
Sirichakwal and Conner, (2016)	**AM impact on inventory costs for single - location applications**	x		x		x		
Heinen and Hoberg, (2019)		x		x		x		
Zhang et al., *(2019)*		x			x		x	x
Cestana *et al.*, (2019)		x		x				
Togwe, Eveleigh and Tanju, (2019)		x			x			
Knofius, van der Heijden and Zijm, (2019)		x		x		x	x	
Rickenbacher, Spierings and Wegener, (2013)	**AM production efficiency**	x			x		x	x
Ruffo, Tuck and Hague, (2006)		x			x		x	x
Piili et al., *(2015)*		x			x			x
Baumers et al., *(2011)*		x			x		x	x
Baumers and Holweg, (2019)		x			x		x	x
Liu et al., *(2014)*	**AM impact on supply chain performance**	x		x		x		
Li et al., *(2017)*		x		x		x		
Roca *et al.*, (2019)		x		x			x	x
Ghadge et al., *(2018)*		x		x				
Ashour Pour et al., *(2017)*		x			x	x		x
Ryan et al., *(2017)*			x	x	x			
Rogers, Baricz and Pawar, (2016)			x	x				

Table 2.1: State of the Art of Additive Manufacturing impact on production and Supply Chain.

2.7 Supply Chain disruption risk

For the sake of completeness, even if it would be not the main topic of research, it is considered interesting to complete the present work's State of the Art considering the influence that AM can have on supply chain disruption risk, underlining the wide application scenarios that the technology can positively impact.

Supply chain system faces a great variety of disruptive events: external risks, such as the earthquake and tsunami in Japan in 2011 or the Hurricane Katrina in 2006; demand risks due to an unusual demand variability, customer bankruptcy or fragmentation; supply disruption risks due to price instability caused by financial or political crisis; time risks due to huge delays in the SC or information risks caused by communication breakdown or legal dispute (Ivanov, Dolgui and Sokolov, 2019). In addition, nowadays the global SC is impacted in its totality by a specific case of disruption risk: the epidemic one, due to the Covid-19 pandemic. This catastrophic outbreak is affecting the worldwide SC, both from operational and economical side, leading to substantial consequences in trading and commerce (Ivanov, 2020). It is therefore of primarily attention to understand how a disruption risk propagates along the SC in order to ideate mitigation action to contain the effects and re-stabilize the status quo.

Disruptive events that propagate along the SC for a long-term period, cause the so-called Ripple Effect. This phenomenon occurs when disruptions, rather than remaining localized to one part of the SC, cascade downstream, impacting at its different levels. It is related to low frequency but high impact risks, that leads to a long-term performance decrease, maybe semi-annual or annual, and it require a long recovery period, with high investments to be done. This can be considered as the opposite of the Bullwhip Effect, that concerns recurrent and operational risks as demand fluctuations that affect daily or weekly performance and can be compensated in a short-term period (Dolgui, Ivanov and Sokolov, 2018).

Supply chain weakest points identification

Because of the huge impact that the ripple effect brings to the SC, research is done to model this phenomenon, having the aim of finding the weakest SC points and connection that should be strengthened. In this way, decision-makers can analyse in a proactive stage the vulnerability of the SC and the possible disruption paths, evaluating in advance where reinforcing actions are required (Figure 2.8). In particular, Hosseini, Ivanov and Dolgui, (2020) modelled the SC with Dynamic Bayesian Network to simulate the disruption propagation behaviour, combined with Discrete Time Markov Chain to model the state of every single supplier in time (disrupted or not) and appreciate its vulnerability and ability to recover. Pavlov *et al.*, (2019) used the Genome Method in order to represent the SC network and observe the various disruption paths and optimize possible flow reconfiguration and

Blos, da Silva and Wee, (2018) exploited the potential of the Petri Net for giving a graphical and mathematical representation and easy interpretation of the network, together with Agent Based model to represent the interactions between the suppliers. The aforementioned models can be considered an effective tool to appreciate the disruption propagation and its magnitude in the possible SC paths.

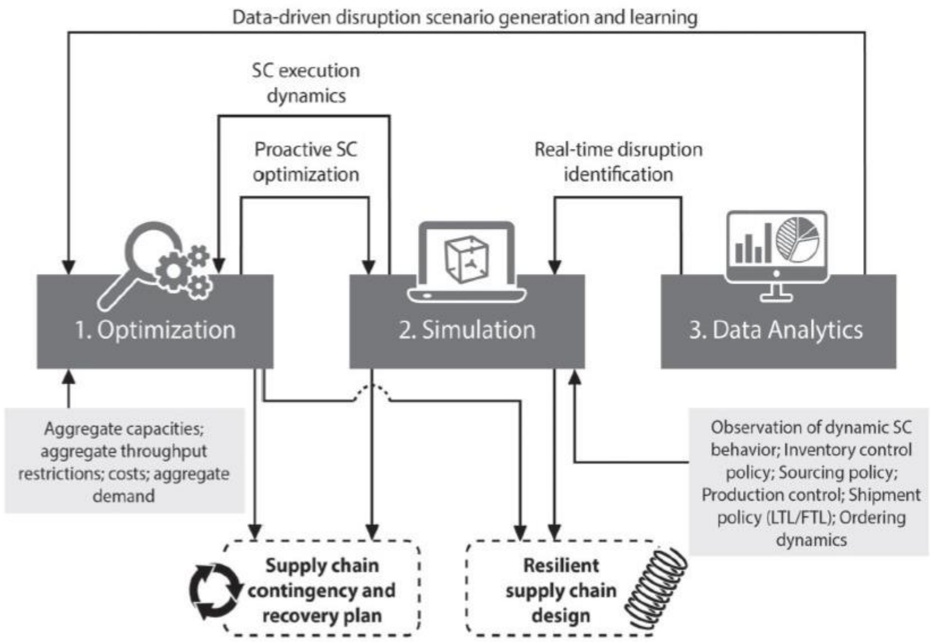


Figure 2.8: Concept of a decision-support system for Supply Chain rick analysis (Ivanov, Dolgui and Sokolov, 2019)

Mitigation actions

Having identified the critical points, it is necessary to present some possible mitigation actions to prevent and quickly react to disruptive events. In particular, the general purpose is making the SC more resilient. The *resilience* is "a property which involves the ability of the SC to change itself quickly, structurally and functionally depending on the current execution state, while simultaneously reaching SC management goals through a change in SC structures and behaviour" (Dolgui, Ivanov and Sokolov, 2018) and it is linked to "durability, recoverability and the maintenance of SC processes" (Ivanov, Dolgui and Sokolov, 2019). The SC resilience can be obtained thanks to *redundancy*. This characteristic is given by the balance between *robustness* and *flexibility*. In fact, robustness allows a direct application of redundancy, using, for example, risk mitigation inventory or backup suppliers, allowing the SC to be able to meet planned performance expectations despite disruption, whereas the flexibility is more related to an indirect usage or redundancy, considering the ability of changing the system behaviour by reallocating inventories, capacities and sourcing in a quick and effective way (Figure 2.9). Robustness and flexibility are therefore considered both necessary in order to absorb and react to possible disruption events, and can be regarded as an 'uncertainty cushion' in a SC (Ivanov, Das and Choi, 2018).

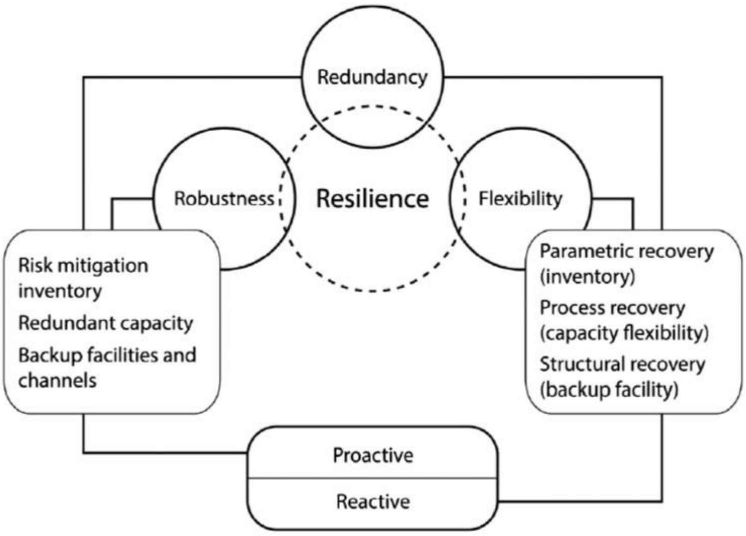


Figure 2.9: Ripple effect control elements (Dolgui, Ivanov and Sokolov, 2018)

Literature shows a quite wide variety of possible mitigation actions that can make the SC more resilient both at the proactive stage, so before the risky event happens, or even after, at the reactive stage, to respond quickly to the problem. Increase the inventory level, consider a possible *risk mitigation inventory*, so extra inventory stored to meet demand in case of uncertainty, or additional reserve capacity are one of the most typical solutions in term of robustness. Lücker, (2019) demonstrates in his works how a proper selection and balance of the order quantity, reorder point and reserve capacity, depending on disruption probability and duration, can lead to a reduction of the disruption impact in term of penalty costs in the SC; also Carvalho *et al.*, (2012) show how a proper increase of the safety stock can reduce the delay that the supply and production have because of disruption at suppliers level. Hosseini *et al.*, (2019) suggest a more strategical solution: in their model, they consider relevant to geographically segregate the supplier. This is because, in case a disruption event, for example a natural catastrophe, affects a certain area of the globe, only some nodes of the SC are impacted, allowing to keep production flowing without big consequences. In addition, they point out how a proper supplier selection is fundamental: supplier must be reliable and resilient to avoid great penalty costs because of their vulnerability and inability to resist to the risky events or to recover fast after disruption, that can lead to the ripple effect manifestation. Cockx, Armbruster and Bendul, (2019) observe how the possibility of sharing resources such machines, know-how and personnel could reduce the supply risk, avoiding possible bankruptcy among supplier, missed payment and delays, that, due to the ripple effect, can affect the whole SC, rather than being limited to only one node. The authors demonstrate how this method can lead to better results than the use of the risk mitigation inventory, but implying that competitors should be prone to cooperate and share resources, which can be a limit, considering the general companies view.

Other researchers consider the use of flexibility to implement mitigation actions to react to disruption events. In particular, Hosseini *et al.*, (2019) evaluate the best suppliers to select examining their ability to expand capacity to absorb disruption effect, utilizing fewer additional resource investments, or considering the use of back-up supplier as a contingency strategy that allows the reconfiguration of the channels for the movement of materials. In a similar way, Carvalho *et al.*, (2012) demonstrate how back-up transporters that can substitute the disrupted one, even at higher costs, can effectively mitigate the ripple effect, avoiding delays in the SC and leading to a final reduction of the total SC costs. Also Ivanov *et al.*, (2016) consider the company flexibility as a solution to avoid disruption propagation, studying the impact on a time critical SC, as the dairy ones. They showed how the *adaptability*, that is the ability to change SC plans, scheduling, or inventory policies in order to achieve a certain desired output, can be realized by means of flexible choices: using back-up distribution centres and alternative transportation means reduce the impact on SC performance in an efficient way.

Additive Manufacturing as flexible and fast mitigation action

The aforementioned mitigation actions show the ability of reducing the ripple effect by means of conventional solution. Ivanov, Dolgui and Sokolov, (2019) give in their work an innovative view about the possible mitigation strategy to follow, considering the increasing influence that industry 4.0 and digital technology are having on SC disruption risk management. They suggest applying the Additive Manufacturing flexibility and adaptability to react to disruptive events. In fact, this technology gives the possibility to produce components and modules in one place and in a faster way, without requiring special tools or a great number of suppliers (Figure 2.10).

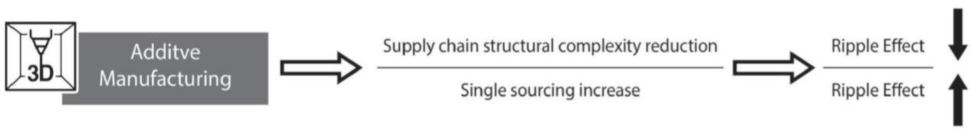


Figure 2.10: Impact of Additive Manufacturing on the ripple effect (Ivanov, Dolgui and Sokolov, 2019).

The SC chain simplification that AM causes has as consequence the reduction of the exposure to external risks and allows a faster and effective response reactions put in place (Liu *et al.*, 2014; Li *et al.*, 2017; De Brito *et al.*, 2019; Roca *et al.*, 2019). In addition, at a proactive stage, AM can avoid the implementation of robust solution as the increase of inventory level or the purchase of additional capacity thanks to the possibility to have a digital inventory always available and produce what is needed exactly when it is needed. This flexibility can therefore answer quickly to demand oscillation, avoiding the ripple effect propagation in the SC.

Additive Manufacturing is shown to have a great potential to be exploited in case of disruptive events and could be an answer to the SC issues caused nowadays because of Covid-19 pandemic outbreak. As Ivanov (2020) initial prediction study about Covid-19 impact and propagation shows, the pandemic would affect the global SC with a long-term performance consequences and great losses, and solutions to deal with them are needed. In addition, Ivanov (2019), demonstrated in his work that disruption effects are still present in SC even after the disruption itself is ended, and if the conventional mitigation actions are deactivated just after the capacity recovery, this would lead to destabilization of inventory system and backlog. This is due to the presence of disruption tails, as delayed orders and service level decrease, caused by inefficiency and unmet demand during disruption period. Therefore, only an application of revival policy in the post-disruption period can lead to a complete recovery and improve SC resilience. This last work demonstrates the big effort required in order to re-stabilize the SC, suggesting AM as prompt response, being a technology with an enormous potential to deal with this requirement in a fast, flexible, efficient and effective way.

3 Resupply time: a Markov Chain based model

The aim of this work is to provide an exhaustive analysis of the AM application in inventory management, comparing different inventory policies. One of the goals is to define an accurate model of the AM production, describing in detail all the stages that this technology requires to further develop the lead time modelling, often assumed constant or generally distributed in scientific literature. This particular focus on an appropriate production modelling is fundamental for the inventory policies. In fact, in order to completely define the inventory model, information about the lead time or *resupply time* must be provided. In particular, in inventory management, the resupply time can be defined as the interval of time that elapses between an order issue and its arrival to the inventory stock (Muckstadt and Sapra, 2010). In general, it includes the production time, set up time, possible waiting times, transportation times and material handling times, depending on the scenario and assumptions posed. This work considers a scenario with a single location manufacturer that controls the AM production, in particular SLM, to replenish the inventory stock. Hence, the lead time would involve (Figure 3.1):

- Possible waiting times since SLM machine has limited capacity and queue can be created
- SLM machine setup time
- Scanning of all the part's layers
- SLM machine cool down time
- Thermal treatments, with possible queue
- Parts removal from the build plate and support removal, with possible queue
- Finishing operations, with possible queue

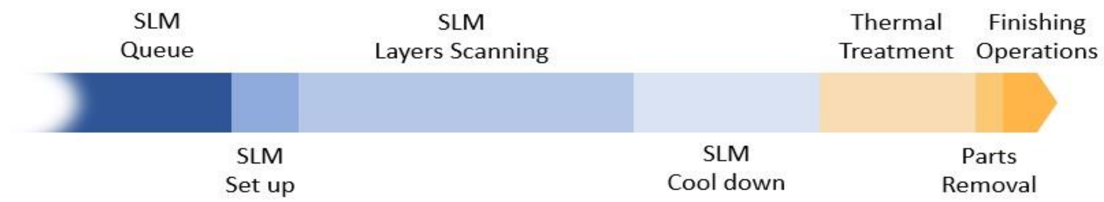


Figure 3.1: Resupply time steps.

It is clear that a sufficiently detailed model of the SLM production can have important consequences on the inventory stock optimization, being strictly correlated to the resupply time evaluation. Therefore, it is worth creating a proper model that effectively represents the AM resupply time considering all the SLM process steps. A common representation of demand arrivals, server with limited capacity and waiting times can be found in Markov Chain. These analytical models are selected and specifically designed for SLM production, as illustrated in this work section.

3.1 System description

The first point of analysis is a detailed description of the SLM production process applied to resupply the inventory stock. In particular, the technology is used to manufacture orders triggered by a specific inventory policy. The focus of this section would be on the production process only, to further analyse it.

In particular, orders arrive following a specific statistical distribution with interarrival rate λ. Subsequently, a queue can be created because the SLM machine is not idle or ready, and finally the production phase takes place, identified by the SLM machine. Note that for the moment the post processing phases as thermal treatments and finishing operations are excluded.

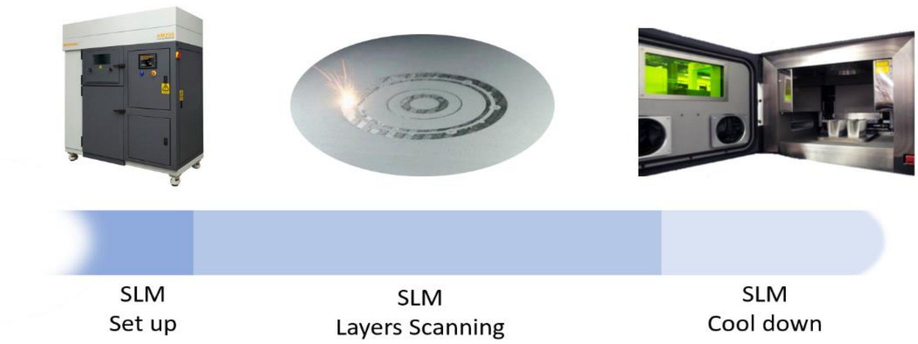


Figure 3.2: SLM production steps (pictures from Renishaw website).

The main modelling focus is on the SLM production representation. As explained in the introduction about this innovative technology, the production process is made by a sequence of steps that are briefly reminded and schematized in Figure 3.2. These are:

i. Set up phase, when the machine is heated up, the gas is injected, the STL file and all the manufacturing parameters are loaded.

ii. Manufacturing phase, when the part cross sections are scanned layer by layer and the powder is spread.

iii. Cool down phase, when the machine is cooled down, parts are removed from the building chamber and the cleaning process takes place.

The objective is to describe all these steps by designing a dedicated Markov Chain for SLM technology. In particular, it is possible to note that the manufacturing phase (ii) can be described by the sum of many other subphases, associated to every layer build. In fact, one can define a single manufacturing sub-step as the process to realize one single layer, made up by the recoating time and the scanning time. Consequently, the manufacturing phase is the sum of m different subphases, where m is the total number of layers that the part requires to be processed and it is proportional to the part height.

In addition to this m manufacturing phases, other two steps are performed: the set-up and the cool down, leading to a total of $n = m+2$ different time steps to describe the whole SLM production.

It can be noted that the time associated to each step is quite different. In fact, the set up and cool down phases require much longer time with respect to the scanning and powder spreading related to a single layer. Therefore, the focus would be now directed to represent all these time steps in a way as adherent as possible to reality, using MC as modelling tool.

3.2 Analytical model for AM resupply time evaluation

Markov Chain (MC) modelling is the selected analytical approach to represent the SLM production process to replenish the inventory stock and analyse the resupply time estimation. In particular, MC in continuous time and with discrete state space are dealt in this work. This choice is motivated considering that they allow the description of systems evolving continuously in time and assuming particular values, i.e. certain system state, with a certain probability and in a discrete fashion, so that the set of values is countable. Examples of the MC models applications are given by the number of clients that visited a shop in function of time, the number of people in queue at the supermarket, the number of orders waiting to be manufactured in a production system. Considering AM application to replenish the inventory stock, the MC model would describe the demands' arrivals, which are the orders triggered depending on the inventory policy and that can happen in any point in time, queue and waiting times because of SLM limited machine capacity, and finally SLM production, considering all the discrete time stages requested by the technology (Figure 3.3).

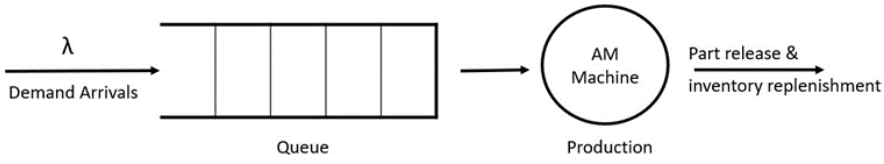


Figure 3.3: Demand and SLM production process.

3.2.1 The demand process

The AM production scenario is characterized by a low annual demand rate so that the technology can be competitive and cost-effective. Furthermore, considering a possible field of applications as the aerospace one or the production of very specific tools with peculiar characteristic, it has been hypothesized that demands are discrete, independent and arrive one per time.

To model this demand behaviour, the Poisson distribution is chosen. The Poisson distribution, in fact, is widely used in quality control, reliability, and queuing theory. It can be applied to represent the customer arrivals at a service in a certain period of time: this distribution type is very often used for describing the number of inventory demands (Law, 2007). Furthermore, it can be applied to describe Stock Keeping Units (SKU) demands used to meet request in case of components failure. In fact, in general, the assumption of Poisson processes can be justified either when lifetimes of components are exponential or when lifetimes are generally distributed and the number of machines that is served by the warehouse is sufficiently large (van Houtum and Kranenburg, 2015)

In the developed models, the AM manufacturer is therefore subjected to an independent identically distributed (i.i.d.) demand, described by a Poisson distribution with parameter λ:

$$p(x) = \frac{e^{-\lambda t}\,(\lambda t)^{x}}{x!} \qquad x = 0, 1, 2 \ldots$$

where *p(x)* defines the Poisson probability mass function.

3.2.2 The production process: an introduction to Phase Type distribution

It is reminded that one of the fundamental hypotheses on which queuing theory is based is the exponential distribution's memoryless properties, on which MC are constructed. The compliance with the Markovian process hypothesis, together with the objective of developing a sufficiently detailed model for SLM production, find the answer in the application of the phase type distribution.

Phase type (PH) distributions are an extremely versatile class of statistical distribution. It is possible to approximate any distribution on the non–negative real numbers by a PH distribution, and the resulting queueing models can be analysed almost as if one has dealt with the exponential distribution. In particular, they are defined as the distribution of the lifetime, i.e. the time spent in some transient phases before entering in an absorbing state (Buchholz, Kriege and Felko, 2014).

PH distributions family owns some well-known distributions, as the Erlang one. In particular, the Erlang distribution E_n^{λ} with n degrees of freedom (or stages) and parameter λ is the distribution of the sum of n exponential random variables with parameter λ:

$$f(x)_{Er} = \frac{\lambda^{n}}{(n-1)!} t^{n-1} e^{-\lambda x}$$

A graphical representation of the Erlang distribution as sum of exponential distributions stages is given in Figure 3.4:

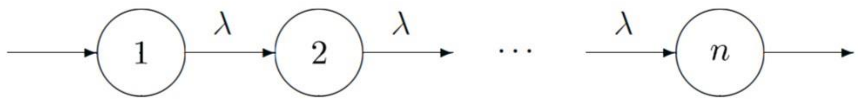


Figure 3.4: Erlang distribution.

The Erlang distribution (Er) can be represented as the holding time in the transient state set {1, ... , n} of a MC with absorbing state n + 1, and the only possible transitions occur from a state k to the next state $k+1$ (for $k = 1, \dots, n$), with rate λ each.

The first two moments of this distribution are given by the following formulas:

$$E[X]_{Er} = \frac{n}{\lambda} \qquad Var[X]_{Er} = \frac{n}{\lambda^2}$$

Another important PH distribution, which is a generalization of the Erlang one, is the Hypoexponential distribution (HE). This type of distribution is obtained if one admits the Erlang exponential stages to have different λ_i parameters. The graphical representation of the Hypoexponential distribution is given in Figure 3.5:

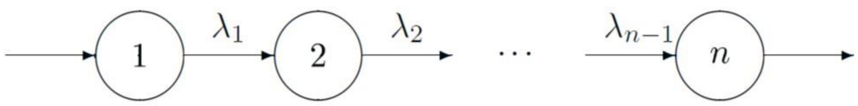


Figure 3.5: Hypoexponential distribution.

The density function of this distribution is:

$$f(x)_{HE} = \sum_{i=1}^{n} \left(\prod_{j=1, j \neq i}^{n} \frac{\lambda(j)}{\lambda(j) - \lambda(i)} \right) * \lambda(i) e^{-\lambda(i)x} \; for \; x \geq 0, \lambda(i) \neq \lambda(j) \; for \; i \neq j$$

The first two moments are:

$$E[X]_{HE} = \sum_{i=1}^{n} \frac{1}{\lambda(i)} \qquad Var[X]_{HE} = \sum_{i=1}^{n} \frac{1}{\lambda(i)^2}$$

It is possible to note that the aforementioned PH distributions are all based on the exponential distribution.

The next step is therefore substituting the exponential distribution, which is the most common distribution to represent the transition rate in MC, with these more general distributions, bringing the advantage of a more versatile description of the Markovian Process. In fact, it is possible to show that the phase type distributions, being in the same family of the exponential one, have again memoryless properties (Breuer and Baum, 2005).

3.2.3 The Markov Chain detailed model

The possibility to substitute the classical exponential distribution with a PH one in the MC leads to a greater versatility of the model. In particular, the PH distributions would be used to model the sum of different phases that describe the SLM production time. The distribution that fits the most the AM requirements is the Hypoexponential one. In fact, this distribution is the sum of *n* different exponentials that can have different rate *λ(i)*.

For this work purpose, the following scenario is presented and the assumptions made are underlined:

- SLM requires a set up phase. This step would be modelled as one of the PH phases, the first one. This phase will be associated with a particular rate, $\mu_{su.}$
- SLM requires also a cool down phase. Similarly to the set up step, this phase would be modelled as one of the PH phases, the last one. A rate μ_{cd} would be associated to this phase.
- The main manufacturing stages performed are related to the powder spreading and the scanning time. These stages are represented as phases of the PH distribution, as many as the number of layers to be realized. These steps have associated a particular production rate, called $\mu_{prod.}$

The modelling of the AM manufacturing process by a Hypoexponential distribution would lead to the following graphical representation (Figure 3.6):

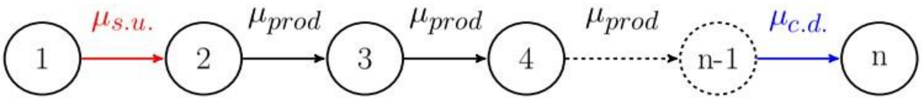


Figure 3.6: Representation of the SLM manufacturing process by means of Hypoexponential distribution.

It is underlined that the hypothesis of modelling AM manufacturing times by means of sum of exponential distributions is not so far from reality. In fact, every stage presents a certain variability in the execution depending the production steps, on the extension of the cross-section to be scanned or on the scanning strategy chosen, and in the setup and cool down stages by the volume of the building chamber or by the temperature to be reached. Therefore, the modelling of each of these stages by an exponential distribution with a suitable rate μ^* can allow the description of the variability related to the production times, permitting at the same time the employment of the MC, a powerful instrument for the system modelling.

Figure 3.6 is a part of the total MC model, the production stage. The whole MC model is now described. Some hypotheses should be taken into account:

- Arrivals are described by a homogeneous Poisson process with rate λ. Every arrival is associated to one single product.

- The manufacturing process is described by the Hypoexponential distribution, with variables rate (μ_{su}, $\mu_{prod,}$ μ_{cd})
- There is only one server and the service discipline is First Come, First Served (FCFS)
- The waiting room capacity is infinite such that there are no units lost.

Note that the arrivals are independent, i.e. demands can arrive in every stage of the production, so even when the SLM machine is already busy.

Because of the accurate description of the SLM process steps, the developed MC model would be identified as the "detailed model". Every system state is described by two variables: the first one, i, represents the number of parts in the system. It is reminded that the SLM machine has limited capacity. Hence, having $i > 1$ means that there is already a part in the manufacturing stage, making the server busy, and therefore the other orders are placed in queue waiting the machine to be idle again. The second index that identifies each stage, j, is associated to the system phase. For example, $j = 1$ is linked to the set-up stage, and $j = n$ represents the cool down step.

The state space of the system Ω_{HE} would be therefore described by a set of stages identified by couple of variables *(i,j)*:

$$\Omega_{\mathrm{HE}} = \{(\mathrm{i},\mathrm{j}), \forall \mathrm{i}: \mathrm{i} \in N\ , \forall j: j \in [0, n]\}\ = \{(0,0), (1,1), (1,2), \dots, (1, \mathrm{n}), (2,1), \dots\}$$

In the developed model, two types of states transitions are possible. The first ones are represented by order arrivals with rate λ. This type of transition increments the number of parts in resupply, increasing in this way the index i by one. The second type of transition is instead referred to the production stream. In particular, it is possible to move from state $j = 1$ to state $j = 2$ following the set up phase with rate μ_{su}; subsequently, there would be m different transitions from each of the stages with $j = 2$ until $j = n$ that represent the manufacturing of the m part layers with rate $\mu_{prod.}$ Finally, the cool down stage would be performed with rate μ_{cd}, considering a transition from $j = n$ back to $j = 1$ and reducing by one the number of parts in the system i, having completed the SLM production process.

In particular, considering that the developed model is a continuous time MC, the process behaviour can be analysed in a very small time interval of length Δt. Calling $1/v_{(ij)}$ the mean of exponentially distributed sojourn time in state *(i,j)*, it is possible to define the mentioned infinitesimal transition rates $q_{(ij),(i^*j^*)}$ from one state *(i,j)* to another state *(i*,j*)* as

$$q_{(ij),(i^*j^*)} = \nu_{(ij)} * p_{(ij),(i^*j^*)} \qquad \text{and} \qquad \nu_{(ij)} = \textstyle\sum_{(ij)\neq(i^*j^*)} q_{(ij),(i^*j^*)}$$

where $p_{(ij),(i^*j^*)}$ is the one-step transition probability of leaving the state *(i,j)* for jumping in the state *(i*,j*)*. The $q_{(ij),(i^*j^*)}$ themselves are not probabilities but transition rates. However, for Δt very small, $q_{(ij),(i^*j^*)}\Delta t$ can be interpreted as the probability of moving from state *(i,j)* to state *(i*,j*)* within the next Δt time units when the current state is state *(i,j)* (Tijms, 2003). Finally, it is possible to completely define the developed continuous time MC *{X(t)}* property:

$$P\{X(t+\Delta t)=(i^*,j^*)|\,X(t)=(i,j)\}=\begin{cases}q_{(ij),(i^*j^*)}\Delta t\ +o(\Delta t), & (i,j)\neq(i^*,j^*)\\ 1-\nu_{ij}\Delta t+o(\Delta t), & (i,j)=(i^*,j^*)\end{cases}$$

where

- $q_{(ij),(i+1\ j)}=\lambda \quad \forall i,j\ \in\ \Omega_{HE}$
- $q_{(i\ 1),(i\ 2)}\ =\ \mu_{su}\ \ \forall i\ \in\ \Omega_{HE}$
- $q_{(i\ j),(i\ j+1)}\ =\ \mu_{prod}\ \ \forall i\ \in\ \Omega_{HE}\ ,\ j=2,\dots,n-1$
- $q_{(i\ n),(i-1\ 1)}\ =\ \mu_{cd}\ ,i=1,2\ \dots$

The described detailed MC model is described by the state transition diagram in Figure 3.7.

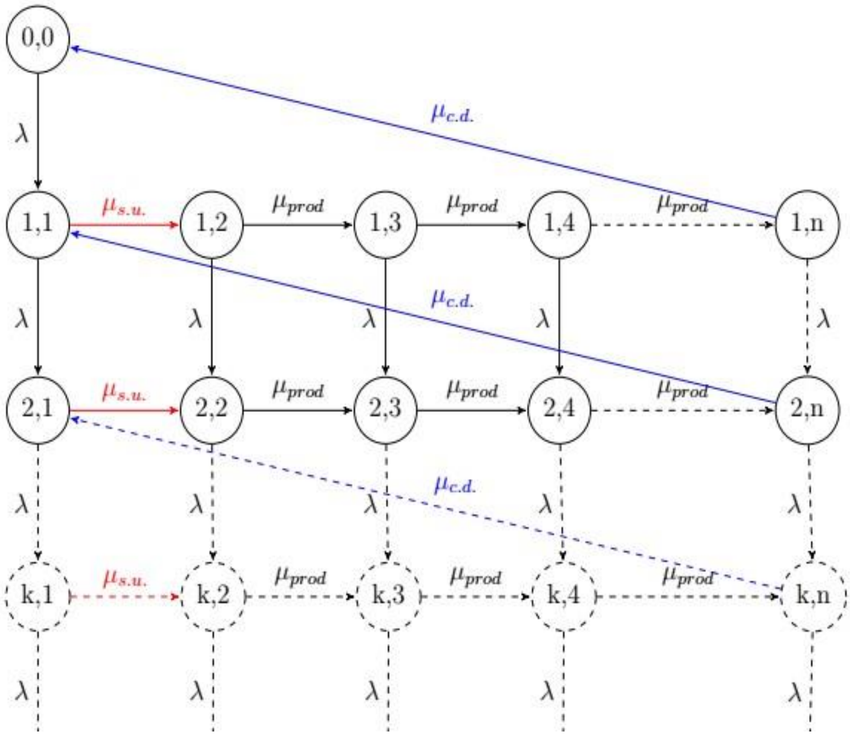


Figure 3.7: Markov Chain representation of the AM manufacturing process using a Hypoexponential distribution.

The final aim of the MC developed model is an estimation of the resupply time requested to replenish the inventory stock. To obtain this parameter, it is first necessary to calculate the MC state probabilities. For this work, the interest would be focused on steady state probabilities, so to the stationary distribution of the Markovian process at an infinite point in time. These probabilities are associated to the number of clients, i.e. parts in resupply, in the system. It is reminded that these probabilities can be calculated if the developed MC is irreducible and positive recurrent. In particular, the MC model is irreducible because there is not a closed subset of states, and therefore it is always possible to define a step transition from one state to another (Figure 3.7) In addition, it can be said positive recurrent because the stability condition holds:

$$\rho = \frac{E[S]}{E[H]} < 1 \tag{3.1}$$

where *E[S]* is the mean service time, while *E[H]* is the mean inter-arrival time. Considering the Hypoexponential case with Poisson arrival, it is given:

$$\rho_{HE} = \frac{E[S]_{HE}}{\lambda} < 1 \tag{3.2}$$

with

$$E[S]_{HE} = \frac{1}{\mu_{avg_{HE}}} = \frac{1}{\mu_{su}} + \frac{1}{\mu_{cd}} + \frac{m}{\mu_{prod}} \tag{3.3}$$

To calculate the steady state probabilities of being in every MC state, it is necessary to write and solve a system of equations. Every equation is a balance equation: the sum of the rates belonging to the arcs entering each state has to be equal to the one exiting the state.

The balance equations describing the system are derived as follow:

$$\begin{cases} -\lambda\pi_{0,0} + \mu_{cd}\pi_{1,n} = 0 \\ -(\lambda + \mu_{su})\pi_{1,1} + \mu_{cd}\pi_{2,n} + \lambda\pi_{00} = 0 \\ -(\lambda + \mu_{prod})\pi_{1,2} + \mu_{su}\pi_{1,1} = 0 \\ -(\lambda + \mu_{prod})\pi_{1,3} + \mu_{prod}\pi_{1,2} = 0 \\ -(\lambda + \mu_{prod})\pi_{1,4} + \mu_{prod}\pi_{1,3} = 0 \\ \dots \\ -(\lambda + \mu_{cd})\pi_{1,r} + \mu_{prod}\pi_{1,n-1} = 0 \\ -(\lambda + \mu_{su})\pi_{2,1} + \lambda\pi_{1,1} + \mu_{cd}\pi_{3,n} = 0 \\ \dots \end{cases} \tag{3.4}$$

In particular, the first system equation is referred to state (0,0), affirming that the sum of the transitions entering this state times their probability ($\mu_{cd} * \pi_{1,n}$) must be equal to the sum of the transitions exiting the state ($\lambda * \pi_{0,0}$). The other equations can be derived following the same logic.

The transition rate matrix *Q*, that collects all the possible transition rates from each one of the states described by the state space Ω_{HE}, is finally defined:

$$Q = \begin{bmatrix} -\lambda & \lambda & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots \\ 0 & -(\lambda+\mu_{su}) & \mu_{su} & 0 & 0 & \dots & 0 & \lambda & 0 & \dots \\ 0 & 0 & -(\lambda+\mu_{prod}) & \mu_{prod} & 0 & \dots & 0 & 0 & \lambda & \dots \\ 0 & 0 & 0 & -(\lambda+\mu_{prod}) & \mu_{prod} & \dots & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & -(\lambda+\mu_{prod}) & \dots & 0 & 0 & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \mu_{cd} & 0 & 0 & 0 & 0 & \dots & -(\lambda+\mu_{cd}) & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & -(\lambda+\mu_{su}) & \mu_{su} & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \tag{3.5}$$

Because equations 3.4 are linearly dependent, it is necessary to add the normalization equation to solve the system. This equation assures that the sum of all the probabilities in the probability vector is equal to 1.

Finally, it is possible to write:

$$\begin{cases} \pi Q = 0 \\ \displaystyle\sum_{i,j=0}^{\infty} \pi_{i,j} = 1 \end{cases} \tag{3.6}$$

Where π is the probabilities vector and Q the transition rate matrix. This system of equations is solved numerically using Matlab software. The complete codes can be found in the Appendix A.1. In this way, the vector of probabilities π is found.

Finally, calling *E[N]* the expected number of clients in the system, i.e. the number of parts in resupply, the application of Little's Law allows to calculate the expected resupply time $\bar{\tau}_{HE}$:

$$E[\tau]_{HE} = \bar{\tau}_{HE} = \frac{E[N]}{\lambda} = \sum_{i=0}^{\infty}\sum_{j=0}^{n} \frac{i\pi_{ij}}{\lambda} \tag{3.7}$$

For the sake of completeness, also other system performance measures are defined. The expected residual service time *E[R]* is the residual time that a client still has to spend in the server before completing the process when a new client arrives. In particular, it can be demonstrated that

$$E[R] = \frac{E[S^s]}{2E[S]} = \frac{\sigma_S^2 + E[S]^2}{2E[S]} \tag{3.8}$$

where $E[S]$ and σ_S^2 are the service time mean and variance. The complete calculations development can be found in Adan and Resing (2015). Therefore, the expected residual service time for the developed MC model $E[R]_{HE}$ can be defined as

$$E[R]_{HE} = \frac{\left[\frac{1}{(\mu_{su})^2} + \frac{1}{(\mu_{cd})^2} + \frac{m}{(\mu_{prod})^2}\right] + \left[\frac{1}{\mu_{su}} + \frac{1}{\mu_{cd}} + \frac{m}{\mu_{prod}}\right]^2}{2\left[\frac{1}{\mu_{su}} + \frac{1}{\mu_{cd}} + \frac{m}{\mu_{prod}}\right]} \tag{3.9}$$

Having identified the expected residual time, it is possible to derive the expected time spent waiting in queue *E[W]*. In particular, when a new order arrives, it has to wait the residual service time of the order already in process, if any, plus the time needed to serve all the orders already in queue. By the PASTA (Poisson Arrival See time Average) theorem (Tijms, 2003 for the complete derivation), it is know that the server is busy on arrival with probability ρ (Equation (3.1)). Denoting with L^q the number of orders waiting in queue, then

$$E[W] = E[L^q]E[S] + \rho E[R]$$

Applying the Little's Law on the parts in queue

$$E[L^q] = \lambda E[W]$$

it is possible to obtain the formula commonly referred as Pollaczek-Khinchin mean value formula

$$E[W] = \frac{\rho E[R]}{1-\rho} \qquad (3.10)$$

that describes the expected average time spent in queue. By substituting the derived relationship for ρ and *E[R]* (Equations (3.2) and (3.9)) in (3.10) the expected waiting time in queue for the developed MC model is found.

Finally, it possible to define the mean duration of the server busy period. In particular, a busy period begins when an arriving customer finds the system empty and ends when a departing customer leaves the system empty behind (Tijms, 2003). The formulation of the expected value of the performance is not straightforward, but, following the derivation provided by Tijms (2003) it can be demonstrated that it is equal to

$$E[BP] = \frac{E[S]}{1-\rho} \qquad (3.11)$$

By substituting *E[S]* with Equation (3.3) and ρ with Equation (3.2), the developed model expected busy period $E[BP]_{HE}$ is computed.

3.3 Alternative analytical models

The model described in Section 3.2.3 uses an Hypoexponential distribution with the aim to fit as best as possible the SLM production process. It can be interesting to compare this model, called the *"detailed model",* with other more conventional and standard queuing theory models to appreciate the differences and possible advantages.

3.3.1 The Exponential model

One of the most known models in queuing theory is the M/M/1 Markov Chain. In this case, the arrivals follow a Poisson process, hence the inter-arrival times are exponentially distributed, distribution also applied for the service time. In addition, there is only one server with limited capacity of one unit per time, and the service discipline is FCFS. This type of chain is also known as *"birth death process"*. In fact, a part arrives (birth) with a certain rate λ and is than processed by the server with a production rate μ. After that, it leaves the system (death). The state space of the system Ω_{Ex} is described by only one variable i, associated to the number of customers present in the system state:

$$\Omega_{\mathrm{Ex}} = \{\mathrm{i}, \forall \mathrm{i}: \mathrm{i} \in N\} = \{0,1,2,3 \ldots\}$$

This type of MC would be identified as the *"exponential"* (Ex) one because of its service time distribution.

This type of queuing system is well known and studied (Tijms, 2003). In particular, there exists a closed formula to calculate the average number of clients in the system, and finally the average resupply time, without the need to numerically solve the system of equations. This is a great advantage, because numerical system implementation is avoided, and computational times are saved. Therefore, it would be interesting to compare the results provided by the detailed model with the one obtained from an M/M/1 queue having the same arrival rate λ and a production rate equal to the average production rate of the Hypoexponential process.

The production rate of the M/M/1 queue would be set as follow:

$$\mu_{\mathrm{Ex}} = \frac{1}{E[S]_{HE}} = \mu_{avg_{HE}} = \frac{1}{\frac{1}{\mu_{su}} + \frac{1}{\mu_{cd}} + \frac{m}{\mu_{prod}}}$$

where $E[S]_{HE}$ is the production time expected value in case of Hypoexponential distribution (Equation (3.3)).

The resulting queuing system can be represented as follow (Figure 3.8):

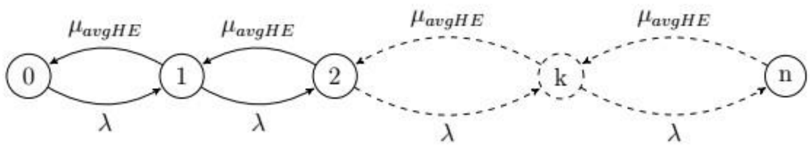


Figure 3.8: Markov Chain representation of the AM manufacturing process using M/M/1 queue.

To find the average resupply time it is sufficient to apply the following closed formula (Adan and Resing, 2015):

$$E[\tau]_{Ex} = \bar{\tau}_{Ex} = \frac{1}{\mu_{Ex} - \lambda} \qquad (3.12)$$

3.3.2 The Erlang model

Another interesting comparison that can be made is considering what would happen approximating the Hypoexponential distribution with an Erlang one having as production time expected value, and as number of stages n the same of the detailed one. This would mean developing a MC having all equal production rates, not distinguishing the set-up and cool down phases.

Imposing the mean production time to be the same as the Hypoexponential one, it is found:

$$E[S]_{HE} = E[S]_{Er} = \frac{1}{\mu_{avg_{HE}}} = \frac{n}{\mu_{prod_{Er}}}$$

therefore

$$\mu_{prod_{Er}} = n * \mu_{avg_{HE}} \qquad (3.13)$$

where μ_{avgHE} is defined in Equation (3.3). From Equation (3.13), it is seen that each one of the n^{th} stage in the M/Er/1 Markov Chain would have a rate n times higher than the average one found in the M/M/1 queue and all the Erlang states would have the same production rate.
As for the detailed model, state space of this model Ω_{Er} is described by couple of variables, i and j. The first one represents the number of parts in resupply, while the second one the machine state.

$$\Omega_{\mathrm{Er}} = \{(\mathrm{i},\mathrm{j}), \forall \mathrm{i}: \mathrm{i} \in N, \forall j: j \in [0,n]\} = \{(0,0),(1,1),(1,2),\dots,(1,\mathrm{n}),(2,1),\dots\}$$

Finally, it is possible to represent the developed model with the following transition diagram (Figure 3.9):

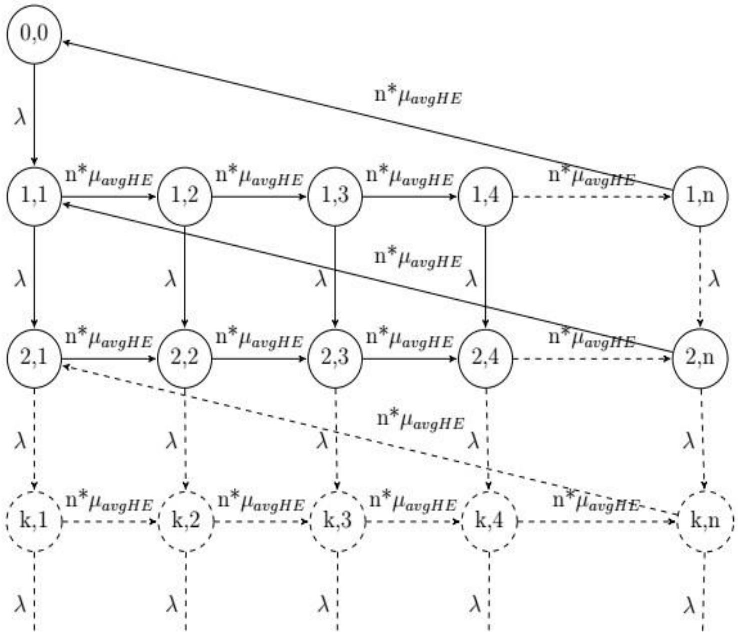


Figure 3.9: Markov Chain representation of the AM manufacturing process using M/Er/1 queue.

Note that this model would be defined as the *"Erlang"* one because of its service time distribution.

To obtain the probabilities vector from the M/Er/1 model, it is necessary again to solve the system of equations associated to the MC. In particular, this is constructed in the same way as the detailed model one (Equation (3.6)).

$$\begin{cases} \pi_{Er} Q_{Er} = 0 \\ \sum_{i,j}^{\infty} (\pi_{i,j})_{Er} = 1 \end{cases} \qquad (3.14)$$

where π_{Er} is the probabilities vector, Q_{Er} is the transition rate matrix (built similarly as in the Hypoexponential case) and the last equation is the normalization equation, needed to uniquely determine the system solution. As for the detailed model, the system is solved numerically using Matlab software (code in 0A.1, but with adjusted transition rates for the Erlang model).

Finally, having calculated the probability vector π, it is possible to estimate the average system resupply time using the Little's Law (Equation (3.7)).

3.4 Analytical models comparison

The aim of this section is to appreciate how the resupply time can vary depending on which one of the three described models is used. It is reminded that the three models are all constructed so that:

- The arrival rate λ is always the same.
- The average production time is always the same, and it is set to be the expected value of the detailed model.

3.4.1 Design of experiments

It is considered interesting to appreciate the influence of the following parameters on the final results:

- The number of strata manufactured, directly proportional to the part height.
- The arrival rate λ.
- The production rate μ_{prod}, related to the SLM machine building rate but also to the number of pieces printed. In fact, it is assumed that the scanning time is directly proportional to the number of pieces in the job, especially if the pieces are all equal and therefore have the same cross-sectional area.

In particular, it has been decided to analyse the resupply time dependency on these particular parameters because they are the ones that in a real SLM manufacturing production system can be varied the most, depending on the scenario defined, the parts geometry or the number of pieces to be printed.

The test case input parameters used are:

- Number of layers: 20, 50, 100 and from 100 up to 1400, with a discretization step of 100 layers. Considering that, in general, SLM has a layer thickness of 50μm, this would mean to print a part with height starting from 1mm up to 70mm. This choice is taken to have a sufficiently wide perspective on model performance.
- Interarrival rate: 1/100 h^{-1} [Ld], 1/50 h^{-1} [Hd]. These two demand scenarios are chosen considering the low demand that characterize the AM production.
- Production rate μ_{prod} = 1/0.003 h^{-1} (equal to 1/10.8 s^{-1}) [Lp] and μ_{prod} = 1/0.03 h^{-1} (equal to 1/108 s^{-1}) [Hp]. These two cross-sectional manufacturing rates are selected considering the data range obtained from experimental tests developed at Politecnico di Milano laboratories. The powder spreading time, which is equal to 0.001h (around 3.5s) is already considered in these input parameters.

The set up and cool down times are considered constant and equal to 55 min and 4.5 h respectively, having as reference the data collected at Politecnico di Milano Additive Manufacturing laboratory.

Four different test cases are developed (Table 3.1) from the combination on the High [H] or Low [L] demand and production rates selected:

Test case	λ [1/h]	μ_{prod} [1/h]
1 [Ld-Hp]	1/100	1/0.003
2 [Hd-Hp]	1/50	1/0.003
3 [Ld-Lp]	1/100	1/0.03
4 [Hd-Lp]	1/50	1/0.03

Table 3.1: MC models test case parameters.

3.4.2 Results and observations

The different average resupply time $\bar{\tau}$ results are shown in Figure 3.10. It is possible to appreciate that, despite all the three models have the same average production and interarrival times in each of the test cases, the average resupply time results are different. In particular, the Exponential model leads to the highest average resupply time, while the Erlang model gives the lowest results.

In all the test cases, the resupply time difference between the Erlang and the detailed model is almost constant with respect to the number of strata printed. On the contrary, the Exponential model provides an increasing resupply time difference with respect to the detailed model as of the number of layers printed increases.

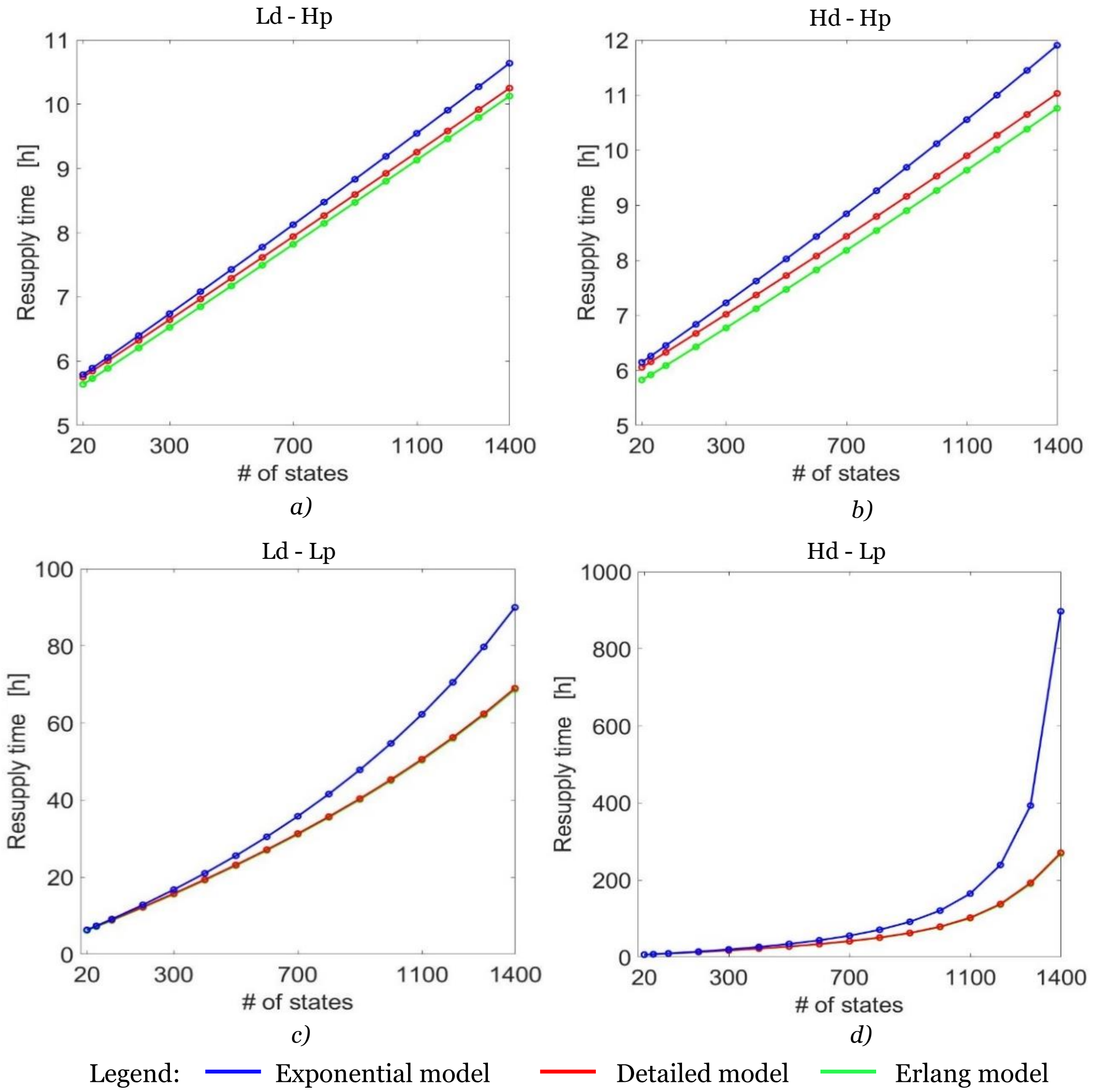


Figure 3.10: Resupply time results for a) Test case 1, b) Test case 2, c) Test case 3, d) Test case 4.

Furthermore, it has been calculated that the maximum percentage difference on the resupply time is low, around the 2-4%, and similar for all the test case in case of the Erlang model. In particular, in Figure 3.10 c) and d) the detailed and Erlang model results almost overlap, being the resupply time really similar but strongly different with respect to the Exponential model. Different are instead the outcomes in case of the Exponential model. In fact, this model seems to be more sensitive to the interarrival and production rate variation, leading to an increasing resupply time difference when the production rate is slower. This error is further pronounced, arriving up to the 230,92% in test case 4, when the slow production rate is combined with a high interarrival rate. In particular, this last scenario models a case in which the volume to be printed is high (low μ_{prod}), and the demand is high (high λ).

The motivation beside these results can be found in the statistical distribution variances.

In fact, it is possible to calculate the variances of the different phase type distributions used for modelling the production time as follow:

Detailed model production time variance

$$Var[S]_{detailed} = \sum_i \frac{1}{(\mu_i)^2} = \frac{1}{(\mu_{su})^2} + \frac{1}{(\mu_{cd})^2} + \frac{n-2}{\left(\mu_{prod}\right)^2}$$

Erlang model production time variance

$$Var[S]_{Er} = \frac{n}{(\mu_{Er})^2} = \frac{n}{\left(n * \mu_{avg_{HE}}\right)^2} = \frac{1}{n * \mu_{avg_{HE}}^2} = \frac{1}{n}\left[\frac{1}{\mu_{su}} + \frac{1}{\mu_{cd}} + \frac{n-2}{\mu_{prod}}\right]^2$$

$$\sim \frac{1}{n * \mu_{su}^2} + \frac{1}{n * \mu_{cd}^2} + \frac{n-2}{\mu_{prod}^2}$$

Exponential model production time variance

$$Var[S]_{Ex} = \frac{1}{(\mu_{Ex})^2} = \frac{1}{\left(\mu_{avg_{HE}}\right)^2} = \left[\frac{1}{\mu_{su}} + \frac{1}{\mu_{cd}} + \frac{n-2}{\mu_{prod}}\right]^2$$

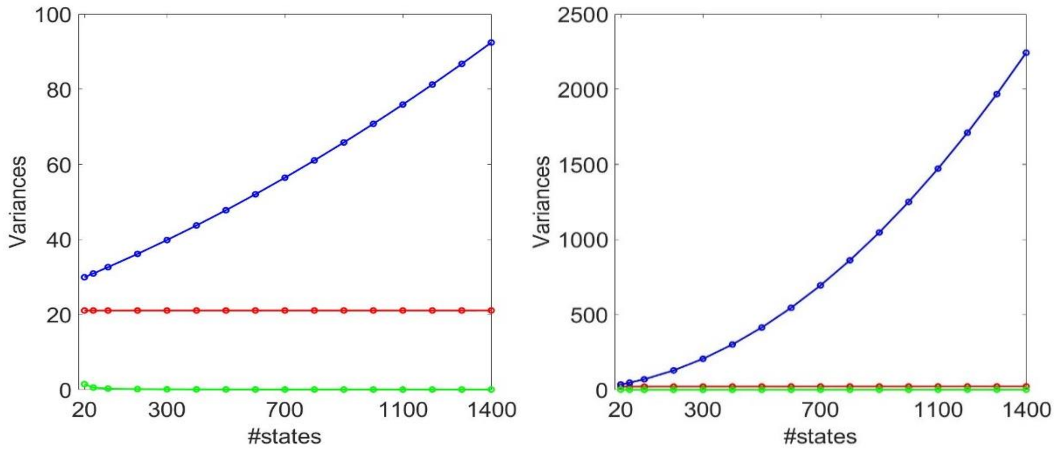
$$\sim \frac{1}{\mu_{su}^2} + \frac{1}{\mu_{cd}^2} + \frac{(n-2)^2}{\mu_{prod}^2}$$

being μ_{avgHE} the average production time obtained from the detailed model and n the total number of stages in the MC model.

The first order approximation of the Erlang and Exponential production time variance shows the Erlang one to be always lower than the detailed model one, while the Exponential to be always higher than the detailed one. This reflects the behaviour found in the resupply time results. In fact, the higher production time variability has direct repercussion on the queue generation, causing longer waiting time and therefore longer average resupply times. Furthermore, it can be noted that $Var[S]_{Ex}$ increases almost with the square of n, explaining the increase in resupply time with the increase of the number of layers. On the contrary, the $Var[S]_{Er}$ shows a more complex relationship with n: it increases with n for the term linked with μ_{prod}, while it decreases together with the terms proportional to μ_{su} and μ_{cd}. This combination of dependencies leads to an almost constant variance with the number of strata, not influencing the final resupply times behaviour. These considerations can be appreciated by looking at Figure 3.11 that compares the production time variances for both the high and low production rates.

The Exponential model

The Ld - Lp test case has shown that if the μ_{prod} is decreased, the difference between the Exponential and detailed model resupply time is further increased. This behaviour is again revealed by looking at the variance's formulation: the term proportional to μ_{prod} in $Var[S]_{Expo}$ would in fact increase its weight on the variance calculation, having as result a rapid increase of the Exponential variance (Figure 3.11 b) and finally of the average resupply time.



a) *b)*

Legend: Exponential model Detailed model Erlang model

Figure 3.11: Tested production time statistical distributions variances, analytically calculated. a) $\mu_{prod} = 1/0.003\ h^{-1}$, b) $\mu_{prod} = 1/0.03\ h^{-1}$.

The second parameter studied is the interarrival rate λ. it can be appreciated that, comparing the Ld-Hp and Hd-Hp test cases or the Ld-Lp and the Hd-Lp test cases, if the interarrival time is shortened, the resupply time $\bar{\tau}$ increases and so does the distance between the detailed and Exponential model resupply times. This behaviour is related to the system dynamics: if the interarrival is reduced, the SLM machine's utilization increases, rising the parts waiting time in queue. This justifies the $\bar{\tau}$ increase in the selected test cases comparison. Nevertheless, the higher variability shown by the Exponential model has a further impact on the system performance. In fact, the higher production time variance propagates along the SLM production models, leading to longer queue formation, especially when the system saturates. This consideration proofs the increase in the difference between the detailed and Exponential $\bar{\tau}$ when λ is increased.

Overall, this system behaviour is taken to the extreme considering Hd-Lp test case. In fact, the λ increase is combined with the lower μ_{prod} and therefore higher Exponential variances (Figure 3.11). The synergy of these two factors has as a consequence an overestimation of the resupply time of the 230% with the Exponential model with respect to the detailed model.

<u>The Erlang model</u>

The detailed and the Erlang models lead to similar behaviour. It can be observed, in fact, that the two models' production time variances approximated at the first order differs only for a n factor at the μ_{su} and μ_{cd} terms denominator. This means that these two models would lead to similar trends in resupply times, with the difference depending on the μ_{su} and μ_{cd} chosen. Anyway, it is underlined that the set up and cool down times have a lower variability with respect to the interarrival time or the production rate because they depend on the SLM machine characteristic: once the AM server is chosen, these two parameters are fixed, characterizing definitely the system behaviour. Finally, it is underlined that the Erlang model approximation would always lead to an underestimation, even if small, of the average resupply time.

Final observations

From the obtained results, it is possible to appreciate that using the Exponential approximation of the production time could lead to important resupply time overestimations, especially if the demand is high and the production rate low. Therefore, despite the Exponential model has the great advantage of using the fast and easy closed formula (3.12) for the resupply time calculation, it would be better to have a more realistic approximation of SLM build time even if it requires solving a numerical model, to reduce the resupply time estimation error. Finally, even if the Erlang and detailed models lead to similar results, the model definition and implementation for both the solutions is really similar, and the detailed one does not introduce evident complications with respect to the Erlang one, that has all the strata with equal production rate. To avoid underestimation of the average resupply time, it is therefore suggested to apply the detailed model.

Overall, it is possible to conclude that a sufficiently detailed modelling of the SLM production times, avoiding assuming constant or generally distributed resupply time, can significatively impact the resupply time calculations, especially for high part volume and increasing demand.

3.5 SLM machine analysis using the detailed MC model

Having underlined sensible differences applying the developed models and the importance of considering the different time steps that SLM requires in a detailed way, it is now interesting to appreciate the impact of the input parameters in the detailed model on the final resupply time result.

3.5.1 Design of experiments

The cool down time is a time interval that can vary depending on the operator's experience in parts extraction and machine cleaning, on the number of operators that performs the steps and on the machine model used. Regarding this last consideration, an example is provided by the Renishaw AM250 SLM machine utilized in the Politecnico di Milano laboratories. For this time step, it requires at least 3h or even more depending on the build plate thickness, because this procedure is performed waiting until the machine autonomously cool down at the open air. On the contrary, more performing machines as the Trumpf TruPrint3000, have implemented a fans system to speed up the cool down phase, reducing sensibly its duration around to 15 or 20 min.

For this reason, three different cool down times are tested: 1.5h in case of SLM machine as Trumpf TruPrint3000 and expert operator, 4.5h which is an average value sum of the Renishow AM250 cool down time and cleaning operations, and a final worst scenario with a longer cool down time requirements, for a total of 6h.

Another interesting parameter is the layer printing time. This value strongly depends on the geometry to be printed, on the number of parts placed in the building chamber and on the SLM machine building rate. Also for this case, three different parameters are tested selected considering data provided by the Politecnico di Milano laboratories (Table 3.2). It is reminded that the data reported in Table 3.2 already consider in their definition the time $t_{speading}$ = 0.001 h (around 3.5s) requested to spread the metal powder.

Finally, a variation on the setup time is considered. Similarly to the cool down time, this phase duration depends on the SLM machine performance and on the ability of the operator to load the input data. Despite this, generally this parameter has a lower variation with respect to the cool down one. The parameter tested are reported in considering the information provided by experts at the Politecnico di Milano AM laboratories.

The parameters testes are collected in Table 3.2.

Cool down time [h]	1.5	4.5	6
Printing time [h]	0.003	0.015	0.03
Set up time [h]	0.75	0.92	1.5

Table 3.2: Sensitivity analysis Set up, Printing and Cool down parameters.

A reference test case has been defined, with the parameters in Table 3.3:

	Set up	**Cool down**	**Printing time**	**Interarrival time**
Reference test case parameters	0.92h	4.5h	0.015h	100h

Table 3.3: Reference test case parameters.

To evaluate the impact of the set up, cool down and printing times, these parameters are varied one per time in the experiments, maintaining the others constant and equal to the reference test case ones. For all the test case, the number of layers printed is varied starting from 20 and reaching 2500. This last value corresponds to the production of a part with maximum height of around 10 cm depending on the layer thickness selected.

3.5.2 SLM machine analysis results

Figure 3.12 shows the test case results obtained using the detailed MC model.

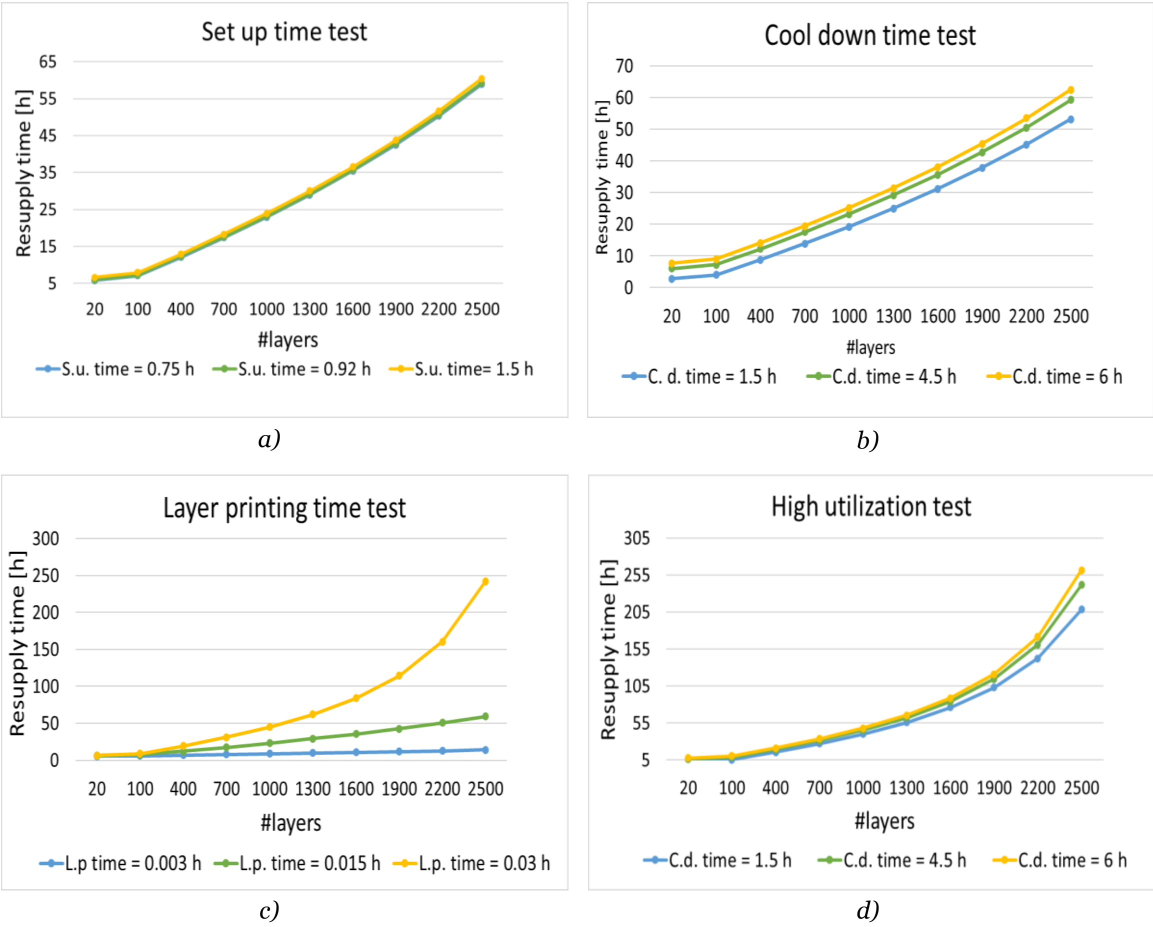


Figure 3.12: Test case for sensitivity analysis a) Set up time test case b) Cool down time test case c) Layer printing time test case d) High SLM machine utilization test case.

It can be generally pointed out that the main differences in terms of resupply time are observed with the higher number of layers printed, 2500. In fact, being the production volume and therefore the production time for this case higher, also the SLM machine utilization is increased, reaching around the 42% and, possibly, queues are created. Hence, the input parameters' variability is further perceived. For this reason, the following analysis reports detailed data focusing more on this particular case, where the resupply times differences emerge strongly. In addition, 2500 layers could correspond to a part height of around 10 cm, which is a typical AM manufactured part dimension.

Set up time test

Figure 3.12a shows the resupply time $\bar{\tau}$ results when the set up time is varied. It can be noted that $\bar{\tau}$ almost overlaps, being the variation of the input set up time minimal. It has been computed that the final resupply time difference on the base case is the double with respect to the absolute delta tested (for example, in case of reducing the set up time by 0.17h, it is obtained a resulting delta on $\bar{\tau}$ of -0.35h). Despite this, the total influence on the resupply time is minimal. In fact, reducing the set up time by 0.17h leads to a resulting $\bar{\tau}$ percentage difference with the base case of just the -0.59%, while increasing the set up time lead this difference to be 2.63%. These results are also related to the small set up time possible variability.

<u>Cool down time test</u>

Figure 3.12b shows the results obtained when the cool down time is varied. Also for this case, it has been calculated that the final resupply times difference with respect to the base case is around the double of the absolute delta tested. For example, reducing the cool down time by 3h leads to a final delta on $\bar{\tau}$ of 6.03h with respect to the base case. In addition, the impact of the variation of the cool down time on the total resupply time is significative, especially when the cool down selected is 1.5h, leading to a percentage difference on the total resupply time equal to -10.18%. This result can be a significative index of the positive impact of the choice of a more performing SLM machine with a reduced cool down time.

<u>Layer printing time test</u>

Figure 3.12c shows the results when the layer printing time is varied. Comparing this result with the previously described ones, it is easy to appreciate a higher impact of this parameter on the total resupply time. In fact, the increase (or decrease) of this parameter, strongly influences the SLM machine utilization. In case of layer printing time = 0.003h, the utilization is around the 12.9% with 2500 layers printed, while when the layer printing time is 0.03h, the machine utilization increases to 80.4%. This results in a higher SLM machine busy time (or idle time), with a respective longer (or lower) waiting time. This finding significantly impacts the final resupply time, with an estimated total difference on the base case of –76.37% when the printing time is reduced to 0.003h, and of 308.31% when the printing time is increased to 0.03h.

<u>High SLM machine utilization test</u>

Having proved that the layer printing time and the cool down time can significatively impact the resupply time, it has been decided to test the case of high utilized SLM machine (printing time = 0.03h) while varying the cool down time. In particular, this test can represent a scenario where multiple pieces are placed in one job to optimize the SLM building chamber utilization, or when one high volume piece is manufactured. Results are provided in Figure 3.12d.

The cool down time reduction has a positive impact on the resupply time $\bar{\tau}$. In fact, it has been obtained that reducing it of 3h leads to a total saving on $\bar{\tau}$ of 33.20h, being this result around eleven times higher with respect to the tested time difference. This decrease also sensibly impacts the final resupply time, with a percentage decrease of 13.73%. On the contrary, increasing the cool down time by 1.5h has as consequence a resulting delta on final $\bar{\tau}$ of 19.72h, with a percentage increase of 8.15% with respect to the base case.

Final observations

In conclusion, this analysis showed that the resupply time is strongly influenced by the layer printing time, and therefore by the number of pieces to be printed, by their volume, or by the SLM machine building rate performance, having significative consequence on machine utilization. It can be underlined that, when the system is particularly saturated, a positive solution to reduce the resupply time is the cool down time decrease. This can be done selecting a high performing machine as the Trumpf TruPrint3000, with an embedded cool down system, and expert operators' employment that can accelerate the removal and cleaning stage. The selection of such a machine could also positively impact the production time rate, offering a possible higher build rate with respect to the Renishaw AM250, meaning that, for the same printing time, the productivity can be increased, or, for realizing the same components, the resupply time can be reduced. On the contrary, the set up time modification is less significant for the total resupply time calculation, even because it was found that this parameter is less subjected to time variation and it has also a smaller absolute value compared with the cool down and total printing time ones.

4 Inventory models

The scenario defined for this work involves a single-location inventory managed by the AM manufacturer, who uses this technology to replenish the stock and owns both the inventory stock and the production workshop. Basically, in a real manufacturing system, clients ask to the producer a specific product: if the component is available on the stock, the warehouse owner picks the part from the stock satisfying just in time the client. This operation leads to a decrease in the inventory level. Hence, the need of developing and optimizing the right inventory policy for stock replenishment arises.

In this work, two different policies are considered: the (S-1,S) and the (r,Q) ones. In both of the cases a continuous time inventory review is assumed. This is not a strict or unrealistic assumption, knowing that nowadays the inventory level monitoring is often automated thanks to the use of designed inventory software combined with scanning systems as bar code or QR code that track the parts and make possible to update and control the inventory just in time.

In addition, it is assumed that the unmet demands are backordered: if the part required is not available, the customers are asked to wait for their requests to arrive, and eventually their orders are satisfied. In contrast, other models apply the lost sales policy: if the demand cannot be immediately satisfied by the parts already available in the stock, the request is definitely lost. Considering backorders possible leads to the introduction of penalties costs. In fact, manufacturer would incur in a higher part cost due to discounts offered to clients because of miscarriage, clients' unsatisfaction risk, discredit and eventually loss of new clients.

A general assumption for all the suggested inventory policies regards the demand process. As reported in Section 3.2.1, the AM production scenario is characterized by a low and unstable parts demand. Very often, in fact, AM is proposed in literature as the flexible solution to promptly respond to the unpredictable demand of aerospace spare parts, or to realize customize components for different client requirements. Considering the scenario defined, the Poisson process is therefore again selected to model the demand arrivals in the inventory system. All the distribution characteristic and peculiarities can be found in Section 3.2.1.

4.1 Analytical models

In this Section, the selected (S-1,S) and (r,Q) inventory policies analytical models are presented. The approach followed is the one suggested by Muckstadt and Sapra (2010).

4.1.1 The (S-1,S) policy under Poisson demand

The first inventory policy studied in this work is the so-called (S-1,S) policy. It is based on placing an order matching exactly the demand size whenever a demand of one or more units of items occurs. This procedure explains the policy nomenclature: calling S the *inventory position,* which measures the inventory hold, or *on-hand* inventory (OH), plus the outstanding orders (O) minus the backorders (B), once it is decreased by at least one unit (S-1,S), an order matching the demand size is immediately placed to bring it back to S.

$$S = OH + O - B \tag{4.1}$$

In the particular case of this work, demands are considered of unitary size. This can be motivated since AM application for inventory replenishment is advocated in scenarios where the annual demand is very low, lumpy and uncertain, as the case of aeronautical spare parts or customized components, that are often asked one per time and when needed for a specific application. The (S-1,S) inventory logic applied to a SLM production scenario is represented in Figure 4.1.

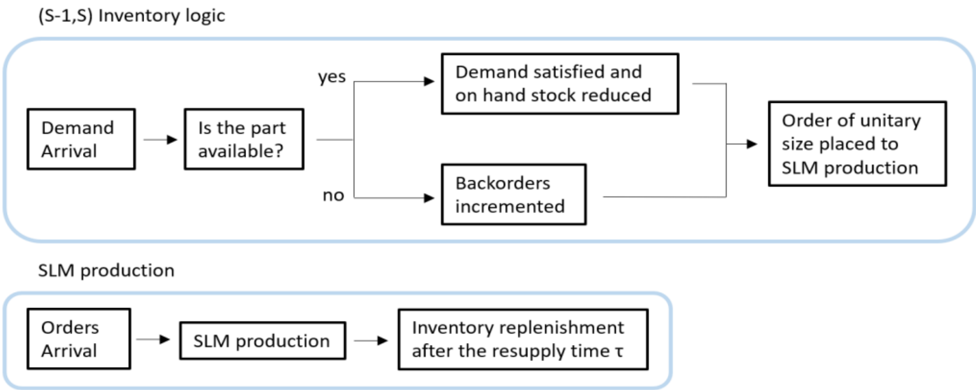


Figure 4.1: (S-1,S) inventory logic.

The *(S-1,S)* policy has often been promoted for controlling inventories of expensive, slow-moving items, i.e., in situations where the demand for the item is infrequent and the item is expensive, so that the cost of ordering is negligible when compared to the costs of holding and shortages (Schultz, 1990). Hence, this policy models the make to order (MTO) or piece-by piece production mentioned in scientific literature regarding AM application in supply chain. In fact, in this research field, the AM technology is used to produce just in time the demanded component with the aim to avoid high stock levels and to incur in risk of obsolescence.

The (S-1,S) model is characterized by an objective function developed to minimize the total annual inventory cost, having the inventory position S as the only decision variable. To delineate the objective function, the costs expressions should be found. In particular, the costs incurred for this inventory policy are the holding costs and the backorders cost. The procurement or order costs are not considered because an order is placed whenever a demand occurs and therefore, they do not depend on the decision variable *S*. Furthermore, this kind of one-for-one replenishment strategy is generally applied when the fixed order costs are generally irrelevant with respect to the final part value. If the fixed order costs start to be significative, other inventory policies where it is possible to accumulate demands and place fixed order quantity should be considered, as the (r,Q) one (van Houtum and Kranenburg, 2015).

The formulation of the cost function follows Muckstadt and Sapra, (2010) approach. It should be noted that the model assumes that the Palm's Theorem, together with its hypotheses, is verified. In particular, this remarkable result is referred to the number of units in resupply calculation and it is briefly reported.

Suppose S is the inventory position for an item whose demands are generated by a Poisson process with rate λ. Suppose further that the resupply time random variables have density functions g(τ) with mean $\bar{\tau}$, and have distribution functions G(τ). Suppose further that the resupply times are independent and identically distributed from customer order to customer order. Then the steady state probability that x units are in resupply is given by

$$P\{X = x\} = p(x|\lambda\bar{\tau}) = e^{-\lambda\bar{\tau}} * \frac{(\lambda\bar{\tau})^x}{x!} \tag{4.2}$$

Being *X* the random variable defining the number of units in resupply. Thus, the probability that there are *x* unit in resupply is Poisson distributed with mean $\lambda\bar{\tau}$, i.e. it is not needed to know the exact distribution of the resupply or procurement time, but just its mean value.

Considering that the Palm's Theorem holds, it is possible to develop the (S-1,S) inventory policy objective function. The whole procedure can be found in Appendix A.2.1. It is recalled that the demand follows a Poisson process with mean λ.

In particular:

- h is the holding cost per year associated to every single unit stock in the inventory.
- b is the backorder cost associated to every unit backordered.

The optimization problem to find the optimal inventory position S^* that minimizes the total annual inventory cost C_{tot} can be defined as follows

$$\begin{gathered} min\ C_{tot}(s) = C_{OH}(S) + C_B(S) = h[\,S - E[X] + B(S)] + bB(S) \\ s.t. \\ F(S) > \alpha \\ S \geq 0 \ \forall\, S \in N \end{gathered} \tag{4.3}$$

where $C_{OH}(S)$ and $C_B(S)$ are the expected holding and backorder annual costs computed calculating the expected number of units in resupply *E[X]* and the expected annual number of backorders *B(S)*. In addition, the fill rate *F(S)* (i.e. the expected fraction of demands that can be satisfied immediately from on-hand stock) constrained is considered In particular, it is assumed that this metric has to satisfy a target level α to be able to promplty meet demand and to satisfy customer requirements.

4.1.2 The (r,Q) policy

One of the possible limitations of the (S-1,S) policy in the designed AM scenario is that an order of unitary size is placed every time a client demands a piece from the inventory. This means that production is run for manufacturing just one part per time. In the case of AM, this results in waiting every time for machine set up and cool down. Furthermore, if the part build is small with respect to the total printing area, most of the recoating time is spent to distribute the powder on areas that would not be scanned. Additionally, it should be remembered that a percentage of the unmelted powder cannot be recycled and must be scrapped: considering a low building volume utilization, this would mean to waste a great percentage of material. Regarding power consumption, the energy requested in the set up and cool down phases would be spent to produce just one piece. Finally, the manufacturer would incur in the order cost every time an order is issued and if the order costs are high, this would lead to high annual expenditure. Considering these observations, literature suggests that the batch production is the best choice to optimize SLM machine capacity and reduce the part production time. Bearing in mind the limitations of the MTO production and the literature outcomes regarding SLM production efficiency, it has been therefore considered interesting to evaluate the batch production repercussions on inventory management. For this reason, the reorder point, lot size (r,Q) inventory policy, which is suitable for batch orders, has been analysed.

In particular, when employing this model, it is assumed that significant order costs are incurred whenever an order is placed. In the case of AM manufacturing, these costs can be related to energy requirements for set up and cool down phases. In addition AM, and in particular SLM, requires the substitution of some consumables as the gas filter, the recoating blade and the build plate that constitutes incurring fixed costs every time a job is given to production (Previtali *et al.*, 2017), and, therefore, an order is issued. Similarly, also a fixed amount of inert gas is consumed in every job. Additionally, waiting every time for both the set up and cool down phases, as well as for post processing operations, increases the total resupply time, having as consequence the requirement of higher stock level to maintain the desired service level.

To reduce the impact of the ordering costs on the total annual inventory cost, the (r,Q) policy suggests to accumulate demands and place batch orders of size *Q* instead of producing every time a demand arrives. The timing of the orders placement depends on the underlying demand process and on specific policy parameters settings. In particular, when the inventory position, which is again (Equation (4.1)) the on-hand inventory (*OH*), plus the outstanding orders (*O*) minus the backorders (*B*), reaches a particular value, called the reorder point *r*, an order of size *Q* is sent. The underlying assumption for the development of this model is that the demand process is memory-less, and this is the case of the Poisson process. Otherwise, for determining the best time to place an order, it would be necessary to know when the last demand occurred, requiring therefore a more advanced solution (Muckstadt and Sapra, 2010). Figure 4.2 represents the (r,Q) inventory logic applied to a SLM production scenario.

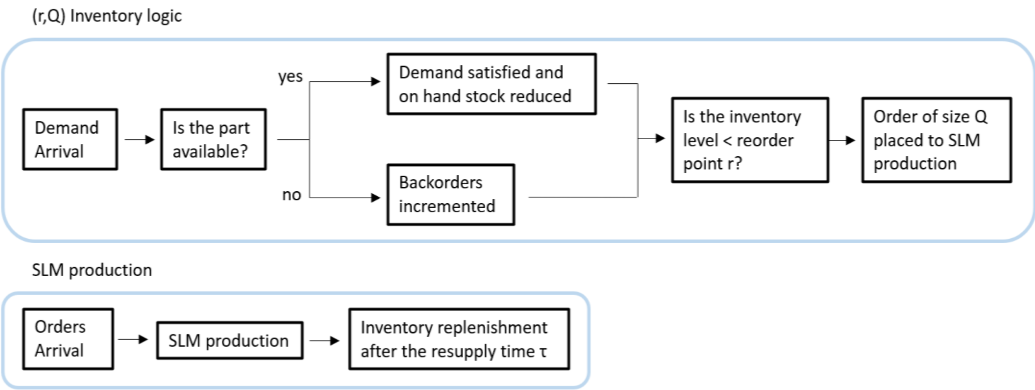


Figure 4.2: (r,Q) inventory logic.

Also for this type of inventory policy, a continuous inventory review is considered, and the demands, which follow a Poisson process, occur for a single unit at time. Because of this hypothesis, the reorder point *r* is met exactly every time an order has to be placed and it is an integer number.

When the number of units demanded in a customer order are more than one, policies other than the (r,Q) should be employed, like the (s,S) one (Axsater, 2000). Anyway, for the considered scenario, where AM is used for inventory replenishment, it is reminded that the demand hypothesized is low and lumpy: for the same motivations already provided for the (S-1,S) model, also for the (r,Q) one the unitary demand size is selected.

The goal of the (r,Q) policy is to minimize the total annual inventory cost, which depends on the (r,Q) decision variables. Literature presents different analytical models to deal with this inventory policy, providing more or less stringent hypothesises and accurate solutions. The one used as reference in this work is suggested by Muckstadt and Sapra (2010) and Hadley and Whitin (1963) and brifly reported in the following section.

An approximate model when backordering is permitted

The first model described is an approximation based on many assumptions. This model is easy to understand, the resolution methods are quite simple, and can be suitable for very low demands. Nevertheless, one must be aware of the impact of the assumption on the policy parameters resulting values.

In fact, the key assumption in this model is that there is never more than one single order of size Q outstanding in any point in time. This implies that, whenever the reorder point r is reached, there are no orders outstanding, or, in other words, that the demand over the resupply time never exceeds Q. This model therefore does not allow the case in which, in a certain resupply time, the demand is so high that exceeds the size Q more than once, leading to the need of placing more than one order of size Q. In addition, it is also assumed that the average number of backorders at a random point in time is negligible.

These assumptions allow to formulate the (r,Q) inventory policy optimization problem for selecting the optimal (r,Q) parameters to minimize the total annual inventory cost. The model is first computed considering a constant resupply time τ. The complete model formulation can be found in 0.0A.2.2.

The cost accounted in the model are:

- a fixed order cost, called K, associated to every order placing.
- the backorder cost b, related to every unit backordered.
- the holding cost h, the cost of carrying a unit of stock for a year.

It is therefore possible to define the (r,Q) optimization problem as:

$$min\ C_{tot}(r,Q) = \frac{K\lambda}{Q} + \ h\left[\frac{Q}{2} + r - \mu\right] + b\frac{\lambda}{Q}\int_{r}^{\infty}(x-r)f(x)dx$$

$$\begin{gathered} s.t. \\ r \geq 0 \ \ \forall\, r \in N \\ Q > 0 \ \ \forall\, Q \in N \end{gathered} \qquad (4.4)$$

where μ is the expected resupply time demand and *f(x)* is the demand density function over the resupply time.

<u>Stochastic resupply time</u>

The above results were obtained considering a constant resupply time. It is possible that the replenishment time is stochastic, hence some modifications of the already described model are necessary. Considering the *resupply time independent from the demand,* the marginal distribution of the resupply time demand is:

$$h(x) = \int_{0}^{\infty} f(x,t)g(t)dt$$

being *g(t)* the density function of the stochastic resupply time.

Following the calculations reported in the Appendix A.2.2, it can be demonstrated that the total annual average cost C_{tot}, considering stochastic resupply time, can be calculated as:

$$C_{tot}(Q,r) = \frac{K\lambda}{Q} + h\left[\frac{Q}{2} + r - \mu^*\right] + b\,\frac{\lambda}{Q}\int_r^{\infty}(x-r)h(x)dx$$

with μ^* the expected resupply time demand i.e. $\mu^* = \int_0^{\infty} xh(x)dx$.

The new total cost equation can be finally substituted in the optimization problem (4.4).

The detailed model

The assumptions made in the aforementioned model can be sometimes unrealistic. In fact, it can happen that there is more than one outstating order and the expected number of backorders is not always negligible. Considering the objective of describing the SLM process as much near to reality as possible, a second model is introduced. In particular, it gives an exact representation of the stationary probability of the net inventory, without approximation.

Also in this case, the demand is assumed to follow a stationary Poisson process with annual rate λ, and customers arrive one by one. The model is firstly developed considering a constant resupply time τ. For synthesis purpose, all the formulas derivations can be found in the Appendix A.2.3.

For this model, the costs incurred are:

- a fixed order cost, called K, associated to every order placing.
- the backorder costs b, related to every unit backordered.
- the holding cost h, the cost of carrying a unit of stock for a year.
- the backorder cost $\hat{b}$ related to every unit backordered during a year.

Following the equation derivation found in the Appendix A.2.3, it is possible to write the exact (r,Q) model optimization problem as follow:

$$min\ C_{tot}(r,Q) = \frac{\lambda K}{Q} + h\left[\frac{Q+1}{2} + r - \mu + B(r,Q)\right] + b * E(r,Q) + \hat{b} * B(r,Q)$$

$$s.t. \qquad (4.5)$$

$$r \geq 0 \ \ \forall\, r \in N$$
$$Q > 0 \ \ \forall\, Q \in N$$

In particular, two types of backorders expression are present: *E(r,Q)* which is the expected number of backorders incident per year, and *B(r,Q)* which represents the number of unit backordered in a generic point of time. Finally, μ is the expected demand over the resupply time.

The exact model and stochastic resupply time

The exact model formulated in the previous section considers the resupply time τ as a constant. Despite this, very often the resupply time is a stochastic variable. In particular, in case of the (r,Q) model application to the SLM production scenario, this time would be linked to the stochasticity of the manufacturing process. The introduction of the resupply time stochasticity in the exact (r,Q) formulation is not straightforward. In fact, in the case of the reported approximated (r,Q) model, it was sufficient to substitute the demand probability distribution with the marginal distribution of the resupply time demand $h(x)$. This was possible because of the hypothesis of just one outstanding order. This is not the case of the exact model, that removes this assumption making impossible to rigorously assert that just one single order is outstanding.

In case of more than one order is allowed to be outstanding, difficulties are encountered in properly represent the resupply time random variable (Hadley and Whitin, 1963). In fact, to provide simpler calculation, one would like to treat the resupply times as independent. This means that outstanding orders can cross, i.e. orders placed in a subsequent moment can be delivered before the first order placed. In practice, it is almost always true that this cannot happen, considering also that, generally, orders are placed in queue and wait in line the other orders to be fulfilled. Hadley and Whitin, (1963) suggest in their book some solutions to consider the resupply time stochasticity. Nevertheless, they are forced to define particular scenarios in which resupply times can be treated as independent or to consider the resupply time equal to its mean, removing in way the variability. Considering that the aim of this work is to provide a detailed analysis of the SLM production process, accurately describing its time steps, these types of assumptions are too strict. Simulation modelling can be a valid solution to represent a real system reducing the above hypotheses as minimum and allowing a fair good representation of reality.

4.1.3 The (r,Q) model in case of lot size dependant resupply time

The (r,Q) mathematical models described in the previous paragraphs consider, with more or less strict hypothesises, the resupply time constant or generically distributed. Despite this, very often in manufacturing systems the resupply time strongly depends on the lot size Q. In fact, it is easy to imagine that, while machines set up time is usually independent from Q, the processing time could well show a dependency on the number of parts to be produced, leading the resupply time to vary depending on the lot size. This particular behaviour is proper of the SLM manufacturing, where the printing time can be considered proportional to the number of parts to be realized. Hence, one major drawback of the models presented is that this outcome is not taken into account.

In particular, Karmarkar, (1987) showed in his work the importance of considering the processing time as dependant on the lot size *Q* starting to analyse just the impact on resupply time. He used Markov Chains to prove the impact of the production time, considered directly proportional to *Q*, on the system time. He found that the system time is strictly correlated to the lot size and exist a point of minimum dependent on *Q* to minimize it. If *Q* is too large, machine saturations and system congestion are caused, leading to performance decrease. Therefore, considering the resupply time constant can bring to not fair estimation of system capabilities.

Having appreciated the relevance of the production time dependency on *Q*, one of the first attempts in studying the impact of the lot size *Q* on the inventory management can be found in the work of Kim and Benton, (1995): they suggest that in manufacturing environments *Q* can have significative consequence on the resupply time and therefore on the inventory costs. They analysed continuous review (r,Q) inventory policy introducing a linear dependency between *Q* and the resupply time with waiting times accounting as a proportion of the resupply time. They developed an iterative algorithm based on an adjusted economic order quantity and demonstrated that significant saving can occur if firms consider the interrelationship between the lot size and safety stock decision, finding more appropriate optimal inventory parameters.

Çakanyildirim, Bookbinder and Gerchak, (2000); and Glock, (2012) underlined in their research the necessity of finding an appropriate (r,Q) model to consider the dependency between lot size and resupply time. They therefore posed as goal the development of a continuous review (r,Q) inventory model considering stochastic demand and variable, lot size dependant resupply time. Both of the models allow backorders. Çakanyildirim, Bookbinder and Gerchak, (2000) discuss two cases: the resupply time linear and concave in lot size, finding for both the models a closed form solution. Glock, (2012) considers instead the resupply time as function of setup and transportation time, plus the manufacturing time, expressed as the production rate times *Q* and provides an iterative algorithm to find the optimal inventory parameter. Despite the results obtained from the two previous models, it is worth pointing out that both the works are based on the hypothesis that only one outstanding order per time is possible every cycle, i.e. no overshooting of orders is considered. These models are therefore an evolution of the approximated (r,Q) model (Section 4.1.2) and still, in order to find a possible solution, assumptions on the number of orders outstanding are required.

An interesting research is done by Noblesse *et al.*, (2014). They assumed a continuous review (s,S) inventory policy, underlining the importance of the impact of the lot size in determining the resupply time. In particular, they constructed a queueing system model that determines resupply time, and it is connected endogenously to a second analytical model that describe the inventory orders placing. In particular, the order quantities generated by the inventory model determine the lot production size and therefore the (production) lead times. They tested the model comparing its result with the ones obtained by the often-used Economic Order Quantity one (EOQ). They showed that setting the inventory parameters based on the EOQ and ignoring the lot size impact on the resupply time, may results in significantly higher expected total cost. Table 4.1 reports the two models results comparison evaluating different decision variables in each of the cases. In particular, they underlined that savings can be obtained considering the dependency of the resupply time on *Q* because lead time variability strongly influences the total cost, and a proper selection of it can reduce the total annual expenditure.

Cases	**(s,S) Endogenous**	**C(s,S) Endogenous**	**(s,S) EOQ-based**	**C(s,S) EOQ-based**
Case 1	(36, 62)	42.9430	(108, 115)	116.8844
Case 2	(46, 74)	54.1630	(143, 150)	151.3624
Case 3	(35, 63)	45.4753	(41, 56)	51.4359
Case 4	(46, 76)	55.5563	(54, 69)	64.2507

Table 4.1: (s,S) parameters and corresponding expected costs per hour C(s,S) when lead times are treated endogenous to the inventory policy, versus the traditional EOQ-based approach (Noblesse et al., 2014).

Despite the model computation complexity and the necessary hypothesises to allow the use of Markovian processes, the model results clearly underline the importance of considering the dependency of resupply time on the order size, especially in case of limited capacity. The production and therefore the resupply times are in fact influenced by the number of units to be machined and to the machine utilization: if the system starts to be saturated and too high queue levels are created, drawbacks can be caused on the system and so inventory performance.

In conclusion, it can be said that considering the resupply dependency on *Q* can result in a turning point for a proper total inventory cost estimation and eventually leading to a more accurate selection of optimal inventory parameters that bring to a reduction of total inventory costs. It is suggested therefore to develop an inventory model that could remove the stringent assumptions regarding the number of outstanding orders and resupply time formulations, with the objective to provide an exhaustive analysis of the AM impact on inventory management, considering that the resupply time, also for this technology, is strongly dependant on the number of pieces to be produced in each job.

4.2 Discrete events simulation models

In this section, an alternative modelling tool given by discrete event simulation is introduced. In particular, the motivations beside this choice are first reported, and the developed models are subsequently illustrated.

4.2.1 Motivations

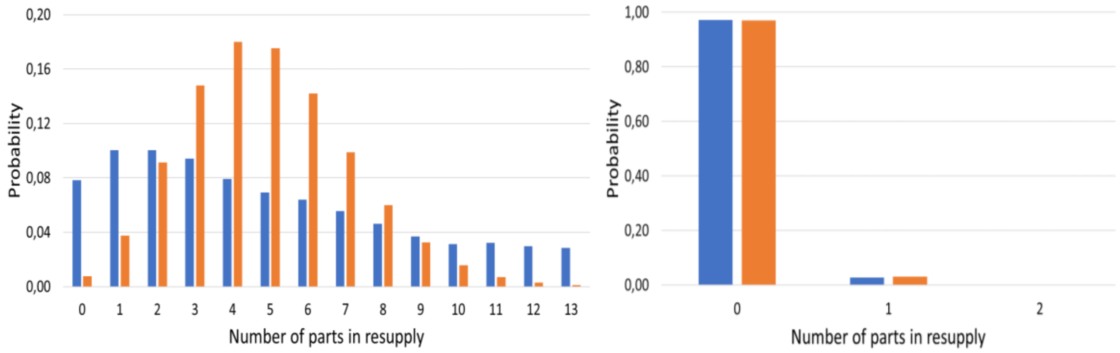
From the analytical models reported in Section 4.1, it can be noted that the main difference between the (S-1,S) and the (r,Q) policies is strictly related to the SLM production process. In fact, in the (S-1,S) policy an order is placed every time a demand arrives, leading to a *make to order* (MTO) or piece-by-piece production, very often advocated in the AM Supply Chain literature. On the contrary, in the (r,Q) policy, orders are issued only when the inventory position reaches the reorder point *r*, and the orders size is equal to the lot size Q, having as consequence a *batch* production, chosen in scientific literature when the scope is gaining AM production efficiency. Being the production strategy the focal driver to characterize the inventory policies, an accurate modelling of the SLM manufacturing is necessary.

It is underlined that the SLM production is quite complex. In fact, the AM machine is characterized by three different phases: set up, printing and cool down. Furthermore, the manufactured components need to undergo to specific thermal treatments, parts must be removed from the build plate with wire EDM and also support structures have to be cut off, and finally finishing operations as grinding are often required. Overall, from modelling point of view, this would mean to consider an articulated process flow, characterized by different stages with specific time durations and peculiarities.

Considering that the aim of this work is to remove the literature assumptions done to model SLM manufacturing, providing instead a detailed view of the AM production, the analytical models' hypothesis described in the previous section may be too strict for this purpose.
First, the (S-1,S) model is constructed on the Palm's theorem (Equation (4.2)), that assumes that the resupply times are all independent and identically distributed. This is not always true: very often, in a real manufacturing system, if the machine has a high utilization rate, parts and orders are often requested to wait in buffers until the machine is idle for production. This means that, if the system is particularly saturated, the order resupply times can be dependent one from another. A proof of this behaviour is given Figure 4.3.

Figure 4.3 compares the probabilities of having different number of parts in resupply calculated with Palm's Theorem and with a simulation model. Both the models are constructed having the same demand and production times. In particular, it has been hypothesized to produce a part with 1250 layers and a mean layer production time of

0.003h. The set up phase is estimated to be 55min, while the cool down 4.5h. Two different demand interarrival times are tested: 10h and 300h.



a) *b)*

Legend: Simulation probabilities ■ Palm's probabilities ■

Figure 4.3: Palm and simulation probabilities comparison for a) interarrival time =10 h, b) interarrival time = 300 h

It can be noted that, in case of low interarrival time (Figure 4.3b) Palm's Theorem and the simulation model provide comparable probability distributions, with a maximum and negligible error of 0.28% in case of 1 part in resupply. On the contrary, in case of higher interarrivals (Figure 4.3a), the two discrete probability distributions give sensible different results, with a maximum error of 10.5% in case of 5 parts in resupply. This is due to the fact that, if interarrival rate is increased, the system gets saturated and orders are forced to wait in queue until the server is idle. This leads to the loss of the independence assumption of Palm's Theorem, leading to significative errors in probabilities estimation.

Another interesting point of discussion is referred to the (r,Q) analytical model's formulations existing in literature. For this case, one of the main issues is related to the resupply time evaluation. In fact, it is often assumed to be constant, and when it is represented by a stochastic distribution, the assumption of independent resupply times is necessary. In a real manufacturing system, when machine limited capacity is considered, this assumption can be too strict. Furthermore, literature suggests that the dependency between the resupply time and the lot size can be very significative in a production scenario, impacting in a remarkable way on the inventory total cost (Section 4.1.3)

Considering all the possible limitations that the analytical inventory policy models present, considering the SLM process flow complexity, and remarking that the aim of this work is to describe a production scenario as much near to reality as possible to collect accurate results, it has been decided to make use of discrete event simulation to model the system. In fact, Kunovjanek and Reiner, (2020) underline that "mathematical analysis, in general, is not powerful enough to provide analytical solution in such complex systems. Experimental approaches are alternatives to that, but they cannot always be performed in real life, so simulation can be sometimes the only option".

4.2.2 Inventory simulation models description

Considering the previous reasoning, the SLM production scenario, together with the inventory management different approaches, are modelled using discrete event simulation in Arena by Rockwell Automation software. A brief model description, together with the assumptions considered, is reported in the following paragraph.

SLM production

SLM machine

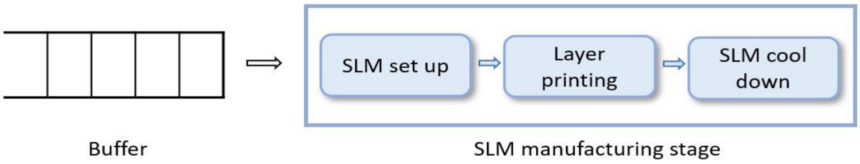


Figure 4.4: SLM server modelling.

Considering the low demand that characterizes the AM production scenario, only one SLM machine is considered. The SLM machine is modelled as a server with a production time that is a sum of the three different stages that the technique requires: set up, printing and cool down (Figure 4.4). The Markov Chain results (Section 3.4.2) have underlined that the choice of a detailed statistical distribution to model the SLM production process can lead to accurate resupply time estimation, as opposed to the generic ones. For this reason, it has been decided to stick with this outcome describing the SLM machine by again a hypoexponential distribution. In particular, this distribution is the sum of exponential distributions with different means, which, for the SLM modelling, are: two different exponential stages for set up and cool down with mean $1/\mu_{su}$ and $1/\mu_{cd}$ respectively, and m intermediate stages with equal mean $1/\mu_{prod}$ for building n layers. It is reminded that $1/\mu_{prod}$ is given by the sum of the powder spreading time t_s and the layer printing time t_p. Hence, this distribution has been modelled in Arena by means of two exponential distributions for set up and cool down and an Erlang distribution with mean m/μ_{prod} for the production phase.
Additionally, considering that all equal parts are produced, it is supposed a linear dependence between t_p and the number of pieces to be printed. This assumption's validity is confirmed by SLM experts' opinion at Politecnico di Milano AM laboratories. Therefore, in case of batch production, the layer production time would be

$$t_{prod} = t_s + Q * t_p = \frac{1}{\mu_{prod}(Q)}$$

In order to characterize the SLM machine as a single server with limited capacity, a buffer of infinite capacity is placed in front of it, so that orders can wait in queue if the machine is busy. This buffer is symbolic: it is a way to state that potentially an infinite number of orders can be collected, meaning that a digital queue of STL file is scheduled, waiting to be concretely produced.

Post-processing operations

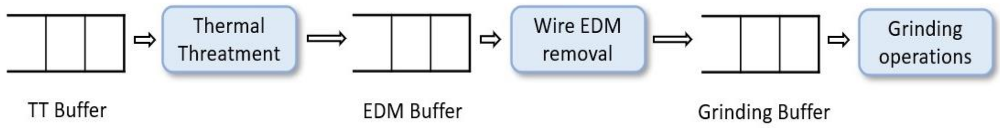


Figure 4.5: Post-processing operations modelling.

After the three-step printing operation, the parts produced undergo to thermal treatment, wire Electric Discharge Machining (EDM) for parts and supports removal and finally grinding for finishing operations. These three subsequent operations are represented by three different single servers with an exponential production time, with a mean respectively of $1/\mu_{TT}$, $1/\mu_{EDM}$ and $1/\mu_{G}$ preceded by a buffer with infinite capacity (Figure 4.5). In particular, the thermal treatment time is considered independent of the number of pieces in the oven, while the EDM and grinding times are directly proportional to the parts to be manufactured. Therefore, as for the SLM machine, in case of batch production of size Q, the EDM and grinding servers would have a busy time Q times longer than the single piece one. It can be noted that there is no such a detailed attention in depicting these three operations (tool changes, downtimes…), being such a description beyond the scope of the work. Anyway, the stochasticity introduced by the selected exponential distribution can easily comprehend all these additional times.

The approximation of considering infinite buffer capacity is not so unrealistic. In fact, the system described is constituted by distinct servers not positioned in a line, where instead one busy server can condition the behaviour of the other servers causing eventually their blocking. Hence, this assumption means that it would be always possible to find space (on the shop floor, pallets, trucks…) where to allocate parts to wait in case of the required server is not idle.

Demand arrivals and inventory policy

As hypothesised in Section 3.2.1, the demands are of unitary size and their arrivals follow a Poisson process with exponentially distributed interarrival times with mean $1/\lambda$. Once a client arrives, the demand can be satisfied just in time if parts are available in the stock, incrementing the counter of clients immediately satisfied. Otherwise, clients are asked to wait and the demand is backordered. Depending on the inventory policy chosen, two different ways of placing orders are requested. In case of (S-1,S) policy, there would be a variable that checks the inventory position I: when it reaches a value equal to S-1, a signal would be emitted and an order would be placed in the production stream. Differently, in case of (r,Q) policy, an order signal would be emitted when I reaches the reorder point r, placing an order of size Q. When the manufacturing stage completes the order production, the components would be used firstly to satisfy backorders, if any, and the remaining parts (if any) would increment the on-hand inventory. It is underlined that, thanks to the use of simulation, no constraint on the number of pending orders is present. Furthermore, it is assumed infinite stock capacity, i.e. there is always enough space to stock the components. Again, this is not such a stringent assumption, considering that the parts manufactured with SLM have small volume and have a low demand, leading to limited stock level and hence limited inventory space requirements.

Collected KPIs

To compute the total inventory cost together with the system performance, the following Key Performance Indicators (KPIs) are collected:

- <u>On-hand inventory over time OH:</u> this is a time persistent statistic computed considering the number of parts available on stock at each simulation instant of time divided by the total time.
- <u>Backorders over time B:</u> this is a time persistent statistic computed considering the number of parts backordered at simulation instant of time divided by the total time.
- <u>Number of orders O:</u> the total number of orders placed every year.
- <u>Number of clients satisfy Just In Time (JIT) and total number of annual demands:</u> these two variables' ratio allow to compute the service level.
- <u>Resupply time:</u> the time that elapses between an order issue and its completion, with the arrival to the stock.
- <u>Server utilization:</u> percentage time fraction in which the server is busy over the total time.
- <u>Buffers waiting times:</u> time that orders spend in queue.

4.2.3 Optimization problems

The implementation of proper simulation models for both the (S-1,S) and (r,Q) policies allows the inventory KPIs computation, overcoming the analytical models assumptions and permitting a detailed description of the SLM production scenario. It is therefore possible to formulate the two inventory models' optimization problems based on the collected parameters.

<u>(S-1,S) optimization problem:</u>

The optimization problem objective function is constructed following the approach presented in Section 4.1.1, using Equations (4.3) as reference. The objective is the total annual inventory cost C_{tot} minimization, which is the sum of the average holding C_{OH} and backorder C_B costs. These costs are computed multiplying the expected on-hand and backorder inventory over time ($E[OH(s)]$ and $E[B(s)]$) obtained from simulation, times the holding and backorder costs per unit. These costs are considered proportional to the unitary production cost $C_{u.p.}$ times a factor, h and b for on-hand and backorders respectively, that define the impact of the unitary production cost on the final cost computation.

$$\min C_{tot}(S) = C_{OH}(S) + C_B(S) = h * C_{u.p.} * E[OH(S)] + b * C_{u.p.} * E[B(S)]$$
$$s.t.$$
$$S \geq 0 \; \forall \, s \; \in N$$

<u>(r,Q) optimization problem</u>

The optimization problem objective function is developed following the exact model approach presented in Section 4.1.2, using Equations (4.5) as reference. The objective is minimization of the total annual inventory cost C_{tot}, which is the sum of the average holding C_{OH}, backorder C_B and order C_O costs, that depend on the decision variables r and Q. The holding and backorder cost are computed in the same way as the (S-1,S) simulation inventory model. The order cost is defined as the product of the fixed unitary order cost $C_{u.o.}$ times the expected number of orders per year N_{orders}, collected from the simulation model. It is reminded that the lot size has an upper bound Q_{max} depending on the maximum number of pieces that can be printed in the same job using an SLM machine.

$$\begin{aligned} \min C_{tot}(r,Q) &= C_{OH}(r,Q) + C_B(r,Q) + C_O(r,Q) \\ &= h * C_{u.p.} * E[OH(r,Q)] + b * C_{u.p.} * E[B(r,Q)] + C_{u.o.} * N_{orders}(r,Q) \end{aligned}$$
$$s.t.$$
$$r \geq 0 \; \forall \, r \; \in N$$
$$1 \leq Q \leq Q_{max} \; \forall \, Q \; \in N$$

4.2.4 Simulation models validation

To check the simulation model correctness, a verification and validation phases are performed. The simulation models are verified thanks to the Arena command *trace* that reports all the events happening during the simulation runs, together with the use of animations that help in visually following the process flow.

The second step is the model validation. This phase is useful in order to determine if the simulation model is an accurate representation of a real system. Considering that it was not possible to find a real industrial case suitable for this work, being the application of AM in inventory management still immature, and taking into account also that no analytical models able to describe such a complex system without the imposition of many assumption exist, it was decided to validate the simulation models following a step-by-step procedure, with an increasing level of hypotheses removal and better representation of the real SLM production and inventory management case.

The different steps followed for the simulation model validations are:

i. Validation of the SLM production stream

In this step, the SLM production process is considered focusing only on the SLM server, together with its production queue and neglecting the post processing phases. In this way, it is possible to compare the simulation resupply times with the one obtained from the detailed MC model and validate SLM the production stage.

ii. Validation of the (S-1,S) inventory model

Also in this step, the simulation production process includes only the SLM machine together with its waiting queue. Additionally to this, the (S-1,S) logic for managing the inventory is added, and inventory KPIs are collected. These results are compared with the ones analytically calculated using both the developed MC detailed model and the analytical (S-1,S) model formulation (Sections 3.2.3 and 4.1.1). In fact, to overcome the limitation underlined in Palm's Theorem application linked to i.i.d. resupply times, the probability of having x units in resupply is computed from the probability vector π that the MC model gives as result. In particular, having as reference the MC model notation of Section 3.2.3, it can be noted that the system state is defined by two variables: the number of units in resupply i and the SLM production stage j. By summing up all the n different stage probabilities π_{ij} for a certain number of units in resupply i, it is possible to find the probability of having i units in resupply π_i:

$$\pi_i = \sum_{j=1}^{n} \pi_{ij}$$

Knowing the π_i probabilities, it is possible to compute the annual average on-hand and backorders levels as expressed in the (S-1,S) analytical model, and compare them with the simulation model results.

iii. Validation of the (r,Q) model

In order to validate the (r,Q) inventory model, the simulation results are compared with the ones analytically calculated with the detailed (r,Q) model described in Section 4.1.2. It is necessary to point out that this model assumes a constant resupply time to consider the possibility of having more than one outstanding order. For this reason, the simulation model is modified to obtain a constant resupply time: the SLM machine is duplicated many times so that the system can be considered having infinite capacity and no queue formation is verified. In this way, it is ensured that the resupply time is constant and equal to the value of the analytical model. The KPIs collected in the simulation model are than compared with the ones calculated analytically.

It is pointed out that the hypoexponential resupply times validation is already provided in the validation step *i*. Hence, the positive validation of both the models in steps *i.* and *iii.* is a positive feedback also for the correctness of the (r,Q) simulation model with stochastic resupply time, that cannot be numerically validated because of missing analytical models.

Models validation results

i. Validation of the SLM production stream

In order to validate the SLM production, the model is tested considering different input parameters in terms of interarrival times, setup, cool down, production durations, and number of strata to be printed (Table 4.2). Every time ten different simulation runs of length 200,000h and a warm-up period of 60,000h are performed, from which the average resupply time and its confidence interval are extracted. The confidence intervals (CI) are calculated with a confidence level of 95%. The comparison between MC and simulation outputs and the validation proof are reported in Table 4.3. The complete table reporting all the simulation runs' results is found in the Appendix A.3. From Table 4.3, it is possible to appreciate that the model is validated for all the different test cases.

	Interarrival time [h]	Set up time [h]	Layer production time [h]	Cool down time [h]	#strata
Test case 1	10	1	0.003	1	1000
Test case 2	24	0.5	0.003	0.5	500
Test case 3	100	1	0.003	4.5	1000
Test case 4	100	1	0.03	4.5	1000

Table 4.2: SLM resupply validation test case input parameters

	MC average resupply time [h]	Simulation expected resupply time [h]	Half CI	Validated?
Test case 1	7.7009	7.7011	0.0592	Yes
Test case 2	2.1023	2.1003	0.0146	Yes
Test case 3	9.0110	8.9951	0.1193	Yes
Test case 4	45.4411	45.4760	0.6206	Yes

Table 4.3: SLM resupply time validation: MC and simulation results of different test cases (CI calculated with a confidence level of 95%).

ii. Validation of the (S-1,S) inventory model

To validate the simplified (S-1,S) model described in the previous section, the same test cases utilized for the SLM resupply time validation are used, with the addition of the inventory position *S* parameter (Table 4.4). In particular, ten different simulation runs of one year (i.e. 8760h) length are performed, considering a warm-up period of 200,000h. The backorders (B) and on-hand (OH) inventory KPIs are collected, and their expected value, together with the CI (confidence level 95%) are calculated. The results are compared with the ones obtained from the analytical model based on the MC. The validation proofs are showed in Table 4.5 and Table 4.6 while the whole simulation runs results are collected in Appendix A.3.

	Interarrival time [h]	Set up time [h]	Layer production time [h]	Cool down time [h]	#strata	S
Test case 1	10	1	0.003	1	1000	3
Test case 2	24	0.5	0.003	0.5	500	3
Test case 3	100	1	0.003	4.5	1000	2
Test case 4	100	1	0.03	4.5	1000	3

Table 4.4: (S-1,S) model validation: test cases input parameters.

	B. analytical results	Simulation expected B.	Half CI	Validated?
Test case 1	0.0271	0.0245	0.0100	Yes
Test case 2	0.0003	0.0003	0.0001	Yes
Test case 3	0.0003	0.0003	0.0003	Yes
Test case 4	0.0032	0.0031	0.0033	Yes

Table 4.5: (S-1,S) model validation: Backorders (B) analytical and simulation results for different test cases (CI calculated with a confidence level of 95%).

	OH. analytical results	Simulation expected OH.	Half CI	Validated?
Test case 1	2.2570	2.2619	0.0285	Yes
Test case 2	1.8896	1.8906	0.0052	Yes
Test case 3	1.9101	1.9041	0.0120	Yes
Test case 4	2.5487	2.5641	0.0597	Yes

Table 4.6: (S-1,S) model validation: On-Hand (OH) analytical and simulation results for different test cases (CI calculated with a confidence level of 95%).

iv. <u>Validation of the (r,Q) model</u>

Also the simplified (r,Q) model with constant resupply time is validated. To prove the correctness of the model independently from the input parameters chosen, different test cases are selected, varying the reorder point r, the lot size Q, the interarrival rate and the resupply time (Table 4.7). Ten different replications for each test case are run, with a length of one year (i.e. 8760h) and a warm-up period of 250,000h. The final results obtained from the simulation and the analytical model are shown in Table 4.8, Table 4.9 and Table 4.10, proving the model validation. The KPIs confidence intervals have been calculated considering a confidence level of 95%. The whole simulation runs result are reported in the Appendix A.3.

	Interarrival time [h]	Resupply time [h]	Q	r
Test case 1	4	30	15	4
Test case 2	3	60	20	10
Test case 3	10	30	20	5
Test case 4	6	40	10	6

Table 4.7: (r,Q) model validation: test cases input parameters.

	B. analytical results	Simulation expected B.	Half CI	Validated?
Test case 1	0.5339	0.5213	0.0386	Yes
Test case 2	2.7473	2.7280	0.1684	Yes
Test case 3	0.1072	0.1091	0.0138	Yes
Test case 4	0.0137	0.0146	0.0053	Yes

Table 4.8: (r,Q) model validation: Backorders (B) analytical and simulation results for different test cases (confidence intervals calculated with a confidence level of 95%).

	OH. analytical result	Simulation expected OH.	Half CI	Validated?
Test case 1	5.0339	5.0536	0.1003	Yes
Test case 2	3.2473	3.1959	0.1743	Yes
Test case 3	9.6072	9.5946	0.1439	Yes
Test case 4	7.5137	7.5459	0.0992	Yes

Table 4.9: (r,Q) model validation: On-Hand (OH) analytical and simulation results for different test cases (confidence intervals calculated with a confidence level of 95%).

	O. analytical result	Simulation expected O.	Half CI	Validated?
Test case 1	146.000	145.800	2.101	Yes
Test case 2	146.000	146.300	2.238	Yes
Test case 3	87.600	87.400	1.589	Yes
Test case 4	87.600	86.900	1.767	Yes

Table 4.10: (r,Q) model validation: Orders (O) analytical and simulation results for different test cases (confidence intervals calculated with a confidence level of 95%).

4.3 Cost model

In order to provide a proper estimation of inventory costs, it is necessary to define a cost model for Additive Manufacturing technology, and in particular for SLM.

The costs associated to the selected inventory policies are listed in the following paragraphs and briefly described.

4.3.1 Inventory policy costs

Order costs

The scenario considered in this work is represented by an AM manufacturer that controls both the production and the inventory system in the same location. For this reason, the order costs are all the fixed costs that incur every time an order is issued. In the SLM production case, the set up and cool down energy cost as well as the cost of the consumables that are requested every time a job is produced, fall into this category.

Inventory holding costs

The inventory holding costs are related to carry one unit in the inventory stock for a certain amount of time. In particular, it can be said that the instantaneous rate at which inventory carrying cost occurs is proportional to the investment in inventory at that point in time (Hadley and Whitin, 1963). For this reason, the physical dimension of inventory holding cost would be cost per unit time. This type of cost incurs because capital is tied up in inventory: if units stay in a warehouse possibly for a long time, there is an opportunity cost that one could have invested otherwise. Furthermore, holding costs are also related to insurance charges, breakage and pilferage at the storage site, soil rental and occupation space in a warehouse and obsolescence, if the part stored reveals not to be sellable anymore.

Generally, this type of costs are considered as a percentage of the final part unit cost (Axsater, 2000), leading to the need of a proper estimation of the SLM manufacturing costs.

Backorder costs

The stockout costs occur when demands are not satisfied because the system is out of stock. In this work, this would lead to backordering the request: the clients are asked to wait until the products required are available again and the demand can be satisfied. This cost is often difficult to be estimated: it is related to loss of customer goodwill (because of the unsatisfied demand, the customer could decide to make business with other manufacturers) and the loss of image that a company can have not being able to satisfy requirements just in time. Therefore, backorder cost strictly depends on the relevance of the consequences that the inability to satisfy demand can have for a manufacturer. Furthermore, they can be associated to a penalty cost that the manufacturer has to incur because offering discount to clients due to the service inefficiency offered.

In this work, the backorder costs are considered proportional to the average amount of units backordered in every instant of time. For this reason, the physical dimension of backorder costs would be cost per unit time, as the holding cost one. As the holding cost, this type of costs is generally considered as a percentage of the final part unit cost, therefore a proper estimation of the unitary SLM production cost is necessary.

4.3.2 SLM costs

To evaluate the inventory costs, a proper estimate of the SLM fixed costs and unitary part costs is necessary, as described in the previous section. Hence, the need to develop a cost model appropriated for SLM emerges. It is reminded that the evaluation of production cost for a company represents an operation of considerable importance, because it directly impacts on the total annual expenditure. For this reason, it has been decided to follow as guideline the cost models developed by Previtali *et al.*, (2017) and Rickenbacher, Spierings and Wegener, (2013), which provide a detailed and accurate cost estimation of SLM, considering different cost drivers.

The costs of interest, collected in Figure 4.6, are:

- Fixed SLM costs for order cost evaluation.
- SLM unitary production costs for holding and backorder cost evaluation.

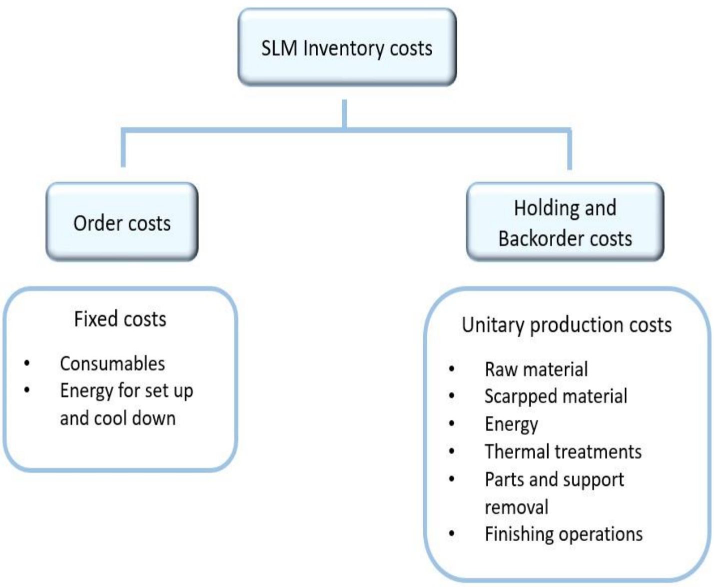


Figure 4.6: SLM inventory costs.

Fixed SLM costs

The SLM production is characterized by fixed costs that incur every time a job is realized. These costs are listed below.

<u>Consumable costs</u>

During SLM production, auxiliary tools and materials are partially or totally consumed during each build job. These components, referred as consumables, are:

- the re-coater blade, which partially wears out in every print.
- the build plate, which is partially consumed when SLM are removed.
- the inert gas injected, which is contained in a tank partially consumed in every job.
- the gas filter cartage that has to be substituted at every job.

The costs of consumables would be calculated as follow:

$$C_{Cons} = \frac{C_{re}}{L_{re}} + \frac{C_{plate}}{L_{plate}} + \frac{C_{gas}}{L_{gas}} + \frac{C_{filter}}{L_{filter}}$$

where C_{re}, C_{plate}, C_{gas} and C_{filter} are the cost of the re-coater blade, the building plate, the inert gas per cylinder and of the gas filter, while L_{re}, L_{plate}, L_{gas} and L_{filter} are the respectively lives expressed in numbers of build-jobs.

<u>Energy requirements in set up and cool down phases</u>

When the SLM machine is heated up or cooled down, a certain amount of energy is required, no matter how many pieces would be printed. Furthermore, an additional energy requested by the computer station is necessary when the machine parameters are set up.

The cost of energy consumption has been calculated as follow

$$C_{en_{fix}} = E * M_{cons} * t_{su} + E * M_{cons} * t_{cd} + E * PC_{cons} * t_{prep}$$

where E is the cost for energy supply, in €/kWh, M_{cons} is the SLM machine energy consumption in set up and cool down phases, in kWh, and t_{su} and t_{cd} are the time spent respectively in set up and cool down, measured in h. The last term is associated to the PC energy consumption that Rickenbacher, Spierings and Wegener, (2013) consider in the preparation phase, when the input parameters are loaded on the SLM machine, requiring a time t_{prep} and a PC energy consumption PC_{cons}, measured again in kWh.

SLM unitary production costs

This cost is associated to the sum of all the specific incurring costs that a manufacturer has to bear when producing a part with SLM. It can be underlined that overhead costs (as, for example, human labour) or machine capital cost are not costs specifically associated to SLM final part manufacturing cost, being the AM machine already bought and there already exists a stable work force hired by the plant owner in the considered scenario.

The relevant SLM manufacturing costs are listed below.

Raw material costs

This cost includes the powder used to selectively melt the parts' layers in the job, together with the cost of the unmelt powder that cannot be recycled and therefore is discarded. It is calculated as follow:

$$C_{mat} = V_p * \rho_p * C_M + V_{scrap} * f * \rho_{pow} * C_M$$

where C_M is the material cost in €/kg, V_p is the final part volume [m^3], ρ_p is the final part density [kg/m^3], ρ_{pow} is the powder density [kg/m^3], V_{scrap} is the volume of unmelt powder, which is calculated as the build job volume V_b minus the volume of the printed parts:

$$V_{scrap} = V_b - V_p$$

being V_b the overall printed area times the height of the tallest part in the build job. Finally, f is known as the scrap factor, which is the powder percentage that cannot be recycled and therefore is wasted.

Energy costs

This cost is related to the amount of energy requested for melting the metal powder and build the final part.

$$C_{en_{print}} = E * M_{cons} * t_{print}$$

where E is the cost for energy supply, in €/kWh, M_{cons} is the SLM machine energy consumption in the printing phase [kWh], and t_{print} is the printing time, measured in hours.

Thermal treatment costs

SLM parts require thermal treatments to reduce residual stresses and improve mechanical properties. The thermal treatments cost is generally given as [€/kg] and therefore is directly proportional to the total parts's weight printed in a job.

$$C_{TT} = TT * W_P$$

where TT is the thermal treatment cost [€/kg] and W_p is the part weight [kg], computes as the part volume V_p times the part final density ρ_p.

Parts and supports removal

Generally, the SLM parts are removed from the build plate with wire EDM and in the same way also the support structures.

$$C_{EDM} = EDM_{proc} * t_{EDM}$$

Where EDM_{proc} is the EDM process cost [€/h] and t_{EDM} is the time requested for removing the part from the build plate, considered proportional to the contact area between the part and the plate, plus the time for support removal.

Finishing operations

To respect quality requirements, SLM parts often necessitate grinding operations to improve the surface roughness.

$$C_{FO} = G * t_G$$

where G is the grinding process cost [€/h] and t_G is the grinding time, proportional to the final part external surface.

Overall production cost

It can be noted that, in case of lot production, these cost drivers are directly proportional to the number of pieces printed. Only the material cost presents a peculiarity. In fact, the cost of the scrapped powder is proportional to the volume of powder unmelted, so to the total building chamber volume, which is constant if the part height is fixed, minus the volume of the pieces printed, which varies with the lot size. Overall, the total variable printing cost in case of batch production is made by a cost factor C^* directly proportional to the lot size Q, plus to a constant term:

$$\begin{aligned} C_{prod}(Q) &= Q * C + C_M * \left(V_b - V_p * Q\right) * \rho_{pow} * f \\ &= Q * C^* + C_M * V_b * \rho_{pow} \\ &= Q * C^* + C^*_{scrap} \end{aligned}$$

with $$C^* = C - C_M * V_p * \rho_{pow} * f$$

where C includes $C_{enprint}$, C_{TT}, C_{EDM} and C_G, and the material cost directly proportional to the number of parts printed. Finally, C^* is the new unitary production cost which considers the savings brought knowing that a portion of the total building chamber volume is occupied by the printed part (V_p), leading to a minor powder volume scrapped and therefore a minor unitary production cost. C^*_{scrap} is instead the cost incurred in case no parts are present in the SLM machine and therefore the whole building chamber volume scrap fraction is wasted.
Finally, by dividing the production cost by the lot size, it is possible to formulate the unitary production cost:

$$C_{u.p.}(Q) = C^* + \frac{C^*_{scrap}}{Q}$$

5 Case study: description

The scientific literature focused on the AM production for inventory management is still limited and consequently it is even more difficult to find real industrial applications for this scenario perspective. Nevertheless, it has been considered interesting to test the developed models on the production of a real SLM manufactured industrial component. For this reason, it has been identified a possible case study suitable for the work purpose.

Fubri is an Italian Company specialized in the precision gear tool manufacturing, offering a complete range of them: hobs, shapers, shaving are some examples. In particular, gear hobs are one of the most widely used cutting tools in gear industry. To increase the production, the classical single-thread hobs evolved in multi-thread hobs. The limitations that often verifies is that they are not as accurate as single-thread hobs. Furthermore, very often these tools are customized for fitting specific gear profiles, and, therefore, they require specific manufacturing technique. An example is provided in Figure 5.1. In order to obtain such features, conventional machining often requires the employment of very advanced machines, as four axis CNC machines, followed by specific grinding operations.

Figure 5.1: Three-dimensional view of Gear Hob (Ayathamaraju and Demir, 2020).

Another important consideration is the tools request: being these components so case-specific, the annual demand is quite low. For example, Fubri presents a model that is requested just 8 times a year. Hence, considering the specialized requirements that these tools have, together with the customization that is often requested and the low annual demand, it is possible to assess that the gear hobs are perfect candidate for AM production.

One previous work (Ayathamaraju and Demir, 2020) developed at Politecnico di Milano Mechanical Engineering Department, confirmed the production feasibility of the Fubri gear hobs by means of SLM, evaluating the best process parameters in order to respect the quality of conformance requirements that the Company asked.

Having demonstrated the tools production feasibility by SLM, it would be interesting to use this result to appreciate the impact of the adoption of this new technology on the annual inventory costs.

In order to perform the analysis, it is of primary importance a proper estimation of the gear hobs' production times by SLM. Having this purpose in mind, technical data of the already performed experimental campaign are collected, and SLM experts as Politecnico di Milano interviewed for the not available information. The data are than elaborated and the models input parameters calculated. The results are presented in the following sections.

5.1 Gear hobs manufacturing scenario

5.1.1 SLM machine

The production time calculations are performed considering the use of the Renishaw AM250 machine, already employed in the experimental campaign that revealed the tools production feasibility (Ayathamaraju and Demir, 2020). The SLM system operates with a Yb:glass fiber laser source with a wavelength of λ=1070nm, maximum power of 200W and a beam diameter in the focus position of 70μm. The laser source operates with a pulse wave emission (PW). The system can be equipped with a reduced build volume system (RBV) that reduces the building volume to 78 x 78 x 50 mm^3 for allowing the test of new raw materials with low powder quantities. Other important technical specifications are reported in Table 5.1.

Maximum building volume	245mm x 245mm x 300mm
Building rate	5 – 10 cm^3/h
Layer thickness	20 – 100 μm
Inert Gas utilized	Argon

Table 5.1: Renishaw AM250 technical specifications.

5.1.2 Production lot characteristics

The reference gear hob models studied in this work are related to Fubri's production data for the year 2019-2020. In particular, Fubri has manufactured 121 different components with a demand ranging from 1 up to 15 units/year, for a total number of 296 components. In order to better focus the work analysis, the different products are grouped into five categories, depending on their main dimensions and peculiarities. The production lot, with all the specifications, is reported in Table 5.2.

Model	Length [mm]	Diameter [mm]	Quantity [year]
M1	50	45	24
M2	100	80	64
M3	150	100	106
M4	180	120	94
M5	260	160	8

Table 5.2: Gear hobs production categories (Ayathamaraju and Demir, 2020).

It is remarked that both the (S-1,S) and the (r,Q) models developed are single-item, single-location inventory policies. Therefore, the focus of the analysis would be on the production and inventory management of a single selected gear hob model per time.

The M1 and M3 models are identified as the two more appropriate test cases. In fact, it can be noted that M1 is the smallest one: this would allow to better exploit the SLM machine building chamber volume, by adjusting many different gear hobs one near the others, permitting to have a good range of variation of the lot size Q in case of the (r,Q) model. M3 is instead selected because it has higher volume with respect to the M1 and it is the most frequently demanded.

5.1.3 Raw material

The metal powder used for the gear hobs production is the MC90 Intermet, a designation given for Fe-Co-Mo grade by Böhler Uddeholm (Lancaster *et al.*, 2016). This is a carbon-free tool steel appropriate to be used for cutting tools, possessing high hardness at elevated temperatures.

5.2 Data collected

5.2.1 SLM production times

The SLM expert's interview provided a detailed description of the time steps that the production of both the M1 and M3 models requires. In particular, for both the cases, the setup phase consists in:

- Build file importing and process parameters assignment.
- Recoater mounting and alignment procedure.
- Baseplate preheating.
- Building chamber preparation (injection of inert gas (Ar) and inert atmosphere creation).

For a total average time t_{su} = 55 min.

After the printing, both the models incurred in the same time steps, identified globally with the term *cool down:*

- SLM machine and build plate cool down: 3h, on average.
- Build plate removal: 10 min, on average.
- SLM machine cleaning: 1h 20 min, on average.

For a total time t_{cd} = 4 h and 30 min.

It is reminded that the baseplate preheating and cool down times depend on the material and dimensions of this component. The data collected are related to the average SLM machine utilization at Politecnico di Milano Laboratories.

After these stages, the SLM machine is ready to start printing again. All the data are collected in Table 5.3.

	M1	M3
Set up time [h]	0.92	0.92
Cool down time [h]	4.5	4.5

Table 5.3: Gear hobs set up and cool down time using Renishaw AM250.

It can be noted that the cool down time is sensibly higher with respect to the set up one. This result is obtained considering the Renishaw AM250 machine, which is not equipped with a refrigerating system that would accelerate this phase. The cool down is simply performed by leaving the machine to reach autonomously the ambient temperature.

It is underlined that it is not strictly necessary to load powder at the beginning of the printing: the SLM machine already has metal powder in the tank from previous jobs, and the raw material can be easily re-loaded while the machine is working. Therefore, this time interval is not considered.

The printing time t_p was calculated for each model considering an average machine production rate equal to 7.5 cm^3/h and taking into account the average cross-sectional area that has to be scanned. From Table 5.2 it is possible to appreciate that M3 has a diameter more than double than the M1 one: for this reason, the layers printing time are quite different, being the M3 one more than four times higher than the of the M1 one.

The recoating time is then estimated. The Renishaw AM250, for every layer to be printed, first positions the recoater blade, then unload the powder that is subsequently spread by the tool. The total evaluated time t_s is recoating time around 3 or 4s = 0.001h.

Furthermore, the optimal layer thickness of 40μm selected in Ayathamaraju and Demir, (2020) work has been kept to obtain a good final part quality. Knowing the part height and this parameter, it was possible to estimate the number of layer that each model requires.

The data collected for the printing phases are showed in Table 5.4.

	M1	M3
Average cross-sectional area [mm²]	535.02	2642.08
Layer scanning time [h]	0.003	0.014
#layers	1250	3750
Volume [mm³]	26751	396332
Average total scanning time [h]	3.75	52.5

Table 5.4: M1 and M3 experimental parameters.

Last line of Table 5.4 describes the average total scanning time for M1 and M3, calculated as the layer scanning time times the number of layers. It is assumed that the parts produced are defect-free, and so the whole manufactured lot can be sellable. This hypothesis is based on the reliability of the results obtained from the Ayathamaraju and Demir, (2020) experimental campaign, that tailored the SLM printing parameters for obtaining high quality components, and considering that no data are available on discarded parts.

5.2.2 Post processing durations

In order to have a complete view on the SLM resupply times, the post-processing durations are collected. In this case, data are estimated from the already executed experimental campaign (Ayathamaraju and Demir, 2020).

In particular, all the pieces need to undergo to an aging thermal treatment for 3h fro residual stress relief. Subsequently, parts are removed from the build plate using wire EDM: in this case, the removal time has been considered proportional to the connection area between parts and build plate (Rickenbacher, Spierings and Wegener, 2013) and an additional time is considered for support removal. Finally, grinding is performed to obtain the final finishing requirements. The post-processing durations are summarized in Table 5.5.

	M1	M3
Thermal Treatment [h]	3	3
Wire EDM removal [h/part]	0.084	0.415
Grinding [h/part]	0.8	2.4

Table 5.5 Post-processing duration for M1 and M3 parts.

5.2.3 Batching

An additional point of discussion is related to the (r,Q) model's specific case. In fact, this inventory policy introduces the problem of creating batches. Hence, it is necessary to estimate the maximum number of pieces that can be allocated in the Renishaw AM250 building chamber. This procedure is called "nesting" because the different pieces have to be nested on the build plate in order to maximize the space utilization, considering a certain tolerance in order to avoid the components contact or difficulties in part removal. Considering the M1 and M3 base area, the total build plate surface and tolerance requirements, the Renishaw AM250 building chamber can accept:

- 25 max pieces of the M1 model
- 5 max pieces of the M3 model

Therefore, these two results are considered the upper bound of the maximum lot size Q possible. It is assumed that for all the lot dimensions the parts are defect free. No experimental results are in fact available on the possible defects generation in case of high lot size. Instead, the Ayathamaraju and Demir, (2020) work experimental campaign reveal the gear hobs SLM production feasibility with final desired quality.

5.3 Gear hobs costs estimation

The cost model developed in Section 4.3, specifically constructed for SLM, is now applied to Fubri gear hobs, used as case study to evaluate the impact of AM on inventory costs. A detailed description of the cost estimation is listed below.

5.3.1 Fixed costs

Consumable costs

Table 5.6 reports the consumable cost in case of using SLM as AM technique. Data are collected in accordance with Ayathamaraju and Demir, (2020) work. The re-coater blade is made by an extruded polymeric material and laboratory research reports that it can be used for three different prints before wear-out. The inert gas used is Argon, and it is estimated that a tank of 200l at 200bar lasts for 4 prints. The metal plate is rectified after the end of each print because of parts removal, and it has an estimated life of 12 prints before its thickness is too low for sustain the printed part. Finally, the filter, which filters the inert gas holding the volatile particles developed during the printing, is replaced every print.

Consumable	**Life [Build jobs]**	**Costs [€]**	**Cost/life [€/build job]**
Re-coater blade	3	2.5	0.83
Inert gas cylinder	4	60	15
Build plate	12	241.5	20.12
Gas filter	1	27	27

Table 5.6: Consumable costs (Ayathamaraju and Demir, 2020).

Therefore, a total cost of 62.95€ per print is estimated.

Energy requirements in set up and cool down phases

This cost is proportional to three different terms: the energy requested for heat up the SLM baseplate to reduce the thermal stresses on the parts produced, the energy for cool down phase, when the SLM is cooled down before parts removal, and finally the PC energy requirements when input parameters are loaded. These costs are calculated considering the requirements of the Renishaw AM250 machine, used for realizing the test case pieces. In particular, this machine is not equipped with an internal cool down system, but, at the end of each print, the machine is left cooling down autonomously for about three hours. Hence, the energy costs related to this step is not considered.

The necessary data for the energy requirements calculation are provided in Table 5.7. The data reported has as reference the average Italian energy supply cost, while the machine and PC workstation requirements as well as the time durations are obtained considering the Renishaw AM250 performances used in the test (Table 5.1).

E: Electricity cost [€/kWh]	0.22
M_{cons}: Machine consumption [kWh]	4
PC_{cons}: PC workstation consumption [kWh]	0.385
t_{su}: setup time [h]	0.92
t_{prep}: preparation time [h]	0.3

Table 5.7: Energy requirements in setup and cool down phases. Energy consumption data are referred to Ayathamaraju and Demir, (2020).

The total cost related to energy requirements in setup and cool down stages is therefore $C_{en_{su-cd}} = 0.83\ €/print$.

Finally, the total fixed cost for the Fubri test case, that will be used in the order cost computation, is equal to:

$$C_{u.o.} = C_{cons} + C_{en_{su-cd}} = 63.78 \frac{€}{print}$$

5.3.2 Unitary production costs

Material cost

Fubri gear hobs are made by MC90 powder, which is a mixture of Fe-Co-Mo that shows mechanical properties appropriate for gear hob tools. The material cost, density data and the scrap factor obtained from the experimental campaign are listed in Table 5.8.

C_M: material cost [€/Kg]	50
$\rho_{p:}$ Final part density [g/cm³]	8.05
ρ_{powder} = powder density [g/cm³]	4.3
f: scrap factor	0.15

Table 5.8: Material costs input data (Ayathamaraju and Demir, 2020).

Scrap cost

It is necessary to estimate the C^*_{scrap} defined in the cost model as the cost associated to the wasted material in case no parts are printed in the building chamber. This cost is directly proportional to the building chamber volume V_b, which depends on the Renishaw AM250 build plate dimensions and on the part height. Data for the M1 and M3 Fubri gear hobs are reported in Table 5.9.

	M1	**M3**
V_b [cm³]	3001.25	9003.75

Table 5.9: M1 and M3 gear hobs total building chamber volume.

This parameter is than multiplied by the material cost C_M, scarp factor *f* and powder density ρ_{powder}, already provided in Table 5.8.

Energy cost

After having interviewed experts at the Politecnico di Milano AM laboratories, the energy cost *E* and the machine consumption M_{cons} for the printing phase are estimated to be the same as the set up one (Table 5.7) for the Renishaw AM250 machine. In this case, the energy cost would be proportional to the printing time of every single piece realized in a job.

Post-processing costs

Thermal treatment costs *TT*, EDM process cost EDM_{proc} and grinding cost *G* are reported in Table 5.10. These data are assumed from previous research (Ayathamaraju and Demir, 2020). These parameters would be than multiplied by the respective time durations (t_{EDM} and t_G) or by the part weight (W_p) for computing the final post-processing cost.

TT: thermal treatment cost [€/kg]	1.5
$EDM_{proc:}$ EDM processing cost [€/h]	45
G: Grinding cost [€/h]	25

Table 5.10: Post-processing parameter cost (Ayathamaraju and Demir, 2020).

M1 and M3 unitary production cost

The M1 and M3 models time durations data and the geometrical characteristic collected in Section 5.1 are finally applied to compute the tools unitary production cost following the developed cost model. It is reminded that the unitary production cost is a function of the lot size, considering the specific characteristics of the SLM technology: $C_{u.p.}(Q) = C^* + C^*_{scrap}/Q$. The two parameters that allows the unitary production cost calculation for the M1 and M3 models are reported in Table 5.11.

	M1	**M3**
C* [€/piece]	37.32	276.63
C^*_{scrap} [€]	96.79	290.37

Table 5.11 M1 and M3 gear hobs unitary production costs parameters.

It is worth pointing out that the C^*_{scrap} is comparable with the C* parameter for the M3 model, while is sensibly higher in case of the M1 model. This difference is due to the smaller M1 volume, which is far more small with respect to the total building chamber volume. This characteristic leads to a higher material waste in case on small batch jobs. Hence, the cost of material wastage in case of small lot size has a high impact on the unitary cost computation, suggesting that the strategy of producing more pieces in the same build job would allow a less material wastage and consequently costs savings.

6 Case study: inventory policies comparison

This chapter has the objective to analyse the impact of the SLM application to resupply the inventory stock. In particular, two different policies are compared in terms of total annual inventory costs using of the developed simulation models: the (S-1,S) and the (r,Q) ones. Both the policies are continuous review ones, and a single-location, single-item scenario is assumed, with unitary size demand. It is reminded that the first policy is linked to the often advocated in literature *make to order* production, considering that an order of unitary size is placed every time a demand arrives. On the contrary, the (r,Q) one is based on the production of *batch* orders of size *Q*, with the aim of better exploiting the SLM machine capacity as the literature focused on AM production suggests. To strengthen the validity of the results, the models are applied to the production of the Fubri gear hobs introduced in the Case Study (Chapter 5). It is reminded that, aware of the outcomes obtained with the Markov Chains models (Section 3.4.2) the SLM production is modelled in a detailed way using an hypoexponential distribution, considering in this way all the time steps that the technology requires.

6.1 Design of experiments

The test case products analysed are the M1 and M3 Fubri gear hobs models, assuming their production on the Renishow AM250 SLM machine. Their specification can be found in Section 5, together with the production time parameters and the estimated unitary production and order costs.

The demand assumed for the M1 and M3 gear hobs is coincident with the one provided by Fubri annual production. The data are reported in Table 6.1, where the calculation of the interarrival times and interarrival rate are also listed.

	Annual demand [parts/year]	**Interarrival time 1/λ [h]**	**Interarrival rate λ [1/h]**
M1	24	365	0.0027
M3	106	82.64	0.012

Table 6.1: Annual demand, interarrival time and interarrival rate of gear hobs M1 and M3 models.

A particular attention is given to the inventory cost definition. In fact, it is often difficult to provide an estimation of the factors *h* and *b* that define the percentage of the unitary production cost (u.p.c.) impacting the on-hand and backorder costs, respectively (Sections 4.2.3 and 4.3.1). In fact, their value is strongly correlated to

Companies' evaluation of the opportunity cost of having capital tied up for the *h* parameter, and of the impact of a possible customer loss for the *b* parameter. Furthermore, the order cost calculation comprehends energy and consumable costs, that can vary from country to country, of from supplier to supplier.

For this reason, different cost scenarios are tested, considering high (H) and low (L) values for the *h* and *b* parameters, and defining an order factor *o* that, multiplying the order cost, would again determine a high and a low order cost scenario. The parameters are estimated by looking at literature values (Heinen and Hoberg, 2019) and interviewing experts from Management Engineering Department of Politecnico di Torino. The results are reported in Table 6.2.

	L	H
***b* [% of the u.p.c.]**	150 %	200 %
***h* [% of the u.p.c.]**	8 %	20 %
***o* [order cost factor]**	80 %	120%

Table 6.2: Evaluation of the backorders on-hand cost and order parameters, considering different scenarios.

The combinations of these values lead to define eight possible costs scenarios collected in Table 6.3.

Scenarios	*b*	*h*	*o*
1	L	L	L
2	H	L	L
3	H	H	L
4	L	H	L
5	L	L	H
6	L	H	H
7	H	L	H
8	H	H	H

Table 6.3: Different costs scenarios.

Simulation models were run testing different (r,Q) and *S* parameters, to obtain their optimal values to minimize the total inventory cost. *S* has been selected ranging from 1 up to 25 parts, while the (r,Q) parameters used are reported in Table 6.4. All the possible combination between the chosen *r* and *Q* parameters are checked.

	r	Q
M1	1, 2, 3, 5, 8, 10	5, 10, 15, from 20 to 25
M3	From 4 to 11	From 1 to 5

Table 6.4: (r,Q) input parameters tested.

The data are collected at the steady state. For this reason, for the (r,Q) model a warm-up period of 800,000h is selected, and the results are related to 30 simulation runs of one year length (1year = 8760h). For the (S-1,S) model, a warm up period of

1,000,000h and 300 simulations runs are considered. Confidence intervals are calculated with a confidence level of 95%.

6.2 Results and discussion

This section presents the results obtained for both the (S-1,S) and (r,Q) inventory policies, applied on M1 and M3 gear hobs management described in the Fubri case study. The analysis starts with the presentation of significant results in terms of unitary resupply time and subsequently showing the two inventory policies costs obtained. Regarding this, it is pointed out that in case of the (S-1,S) policy it is known that the order cost is not influencing the objective function, being constant for every order issue. Nevertheless, to provide a coherent comparison with the (r,Q) model that instead considers the order cost in the total annual inventory cost calculation, it is necessary to include this cost driver also in the (S-1,S) model total cost.

6.2.1 Unitary resupply time analysis

An interesting analysis can be performed calculating the unitary resupply time (u.r.t.). This time duration is obtained collecting the time interval that elapses between the order issue and its arrival to the stock, and finally dividing it by the number of parts produced in every order. Hence, the total resupply time includes the SLM production time plus the post processing durations together with all the waiting times that can incur because of servers limited capacity. The unitary resupply time is calculated for orders of size one, so the one placed in the (S-1,S) model, and for every *Q* parameter tested in the (r,Q) model runs (Table 6.4). The results are showed in Figure 6.1.

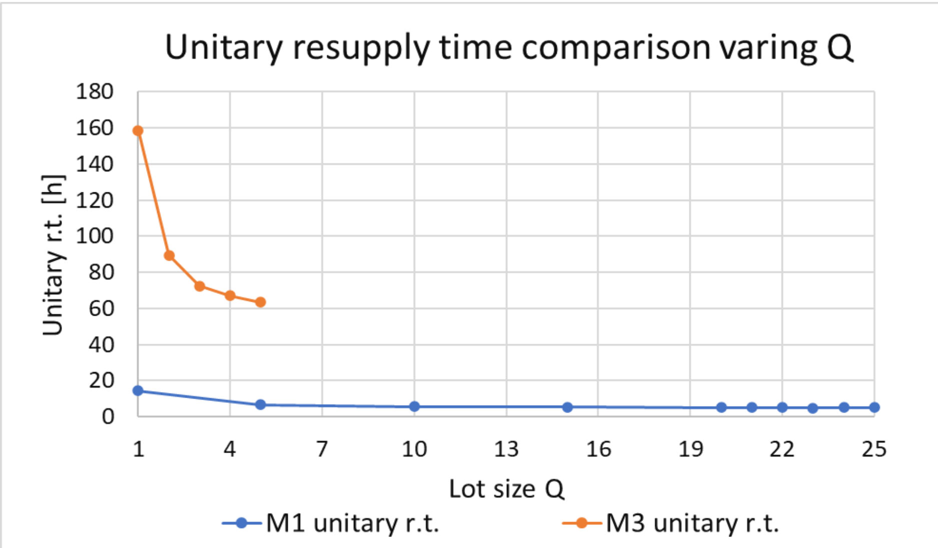


Figure 6.1: M1 and M3 unitary resupply time varying the lot size Q. CI: 3% of the mean value.

It can be noted that the unitary resupply time decreases by increasing the number of pieces to be printed. As significant example, Table 6.5 reports the unitary resupply times calculated in case of only one piece and of the full lot size (25 pieces for M1 and 5 pieces for M3) printed.

	Q = 1	Q max	Delta % u.r.t.
M1 u.r.t [h]	14.48	5.05	186.97
M3 u.r.t. [h]	158.65	63.43	150.12

Table 6.5: M1 and M3 SLM unitary resupply time for minimum and maximum lot size, together with the delta% on their difference. CI: 3% of the mean.

By looking at Table 6.5, it is possible to note that significant savings on the unitary resupply times can be gained by placing multiple pieces inside the SLM building chamber, with a total difference of 186.97% in case of the M1 model and of 150.12% in case of the M3 model between the unitary and completely filled lot. These results are coherent with the ones presented in literature: by placing many pieces inside the building chamber, it is possible to allocate the set up and cool down times, as well as the recoating time on all the parts to be produced, optimizing these fixed durations. Additionally, the overall unitary resupply time is decreased considering the decrease in the total waiting time. In fact, by analysing the SLM machine time spent in buffer in case of just one part printed and multiple parts printed, a sensible difference is noted (Table 6.6).

	Q = 1	Q max
M1 SLM waiting time [h]	0.18	0
M3 SLM waiting time [h]	91.11	30.06

Table 6.6: M1 and M3 SLM buffer waiting time for minimum and maximum lot. CI: 10% of the mean.

From Table 6.6, it is clear that increasing the lot size, the AM buffer waiting time is strongly decreased. This result is obtained considering that, increasing the lot size, orders are placed less frequently, leading to fewer arrivals to the SLM machine and better balancing the system. Nevertheless, increasing the lot size means also increasing the production time. Despite this, times as set up, cool down and powder spreading remain fixed no matter how many components are produced. This would mean waiting at least 0.92h + 4.5h for every order. In addition, the total powder spreading time would require around 1.25h for M1 and 3.75h for the M3 model. Considering that M1 has an average scanning time of 3.75h and M3 of 52.5h (Table 5.4) the fixed times would sensibly impact the total production time and the machine utilization, causing longer queues and waiting times in case of small order size.

In conclusion, it can be observed that, considering the limited SLM machine capacity and taking into account the fixed times that the SLM production requires, the choice of increasing the batch size can have positive repercussions on the unitary resupply time, causing sensible savings on this parameter.

6.2.2 Total annual inventory cost comparison

M1 model

Table 6.7 shows the results obtained from the simulation models when the M1 model is produced. It is possible to observe that the (S-1,S) model leads to total annual inventory costs sensibly higher with respect to the (r,Q) ones. In particular, the total difference can reach up the 1462.09% in scenario 7, being the (S-1,S) total cost around fifteen times higher with respect to the (r,Q) one.

Scenario [b,h,o]	r	Q	Total (r,Q) Cost [€]	S	Total (S-1,S) Cost [€]
1 - LLL	1	24	92.32	1	1221.78
2 - HLL	1	24	92.32	1	1221.84
3 - HHL	1	21	154.66	1	1237.31
4 - LHL	1	21	154.66	1	1237.26
5 - LLH	1	24	116.99	1	1827.44
6 - LHH	1	24	181.49	1	1842.91
7 - HLH	1	24	116.99	1	1827.49
8 - HHH	1	24	181.49	1	1842.96

Table 6.7: Optimal inventory policy parameters and total annual inventory cost in case of Fubri M1 gear hob production. CI inventory KPIs used for costs calculation: 3% of the mean value for the (r,Q) model and 1.5% of the mean value for the (S-1,S) model.

Optimal inventory parameters	Average Backorders [parts]	Average On-hand [parts]	Average Number of Orders
(r,Q) 1, 21	$1.45*10^{-4}$	11.75	1.10
(r,Q) 1, 24	0	13.00	0.97
(S) 1	$7.90*10^{-4}$	0.96	23.74

Table 6.8: Average backorders, on-hand and orders level evaluated in the (r,Q) and S optimum points for M1 model. CI: 3% of the mean value for the (r,Q) model, 1.5% for the (S-1,S) model.

The reason of such results is found considering the total annual cost composition. In case of the (S-1,S) model, around the 99% of the total annual inventory cost is made by order cost (Appendix A.4 reports the detailed results). This can be clearly appreciated by looking at Figure 6.3 that represent the (S-1,S) inventory cost, using the scenario 1 as example. This is due to the fact that every time a demand arrives, an order is issued to replenish the inventory stock. This consideration is confirmed observing the average number of annual orders placed in Table 6.8, that corresponds to the annual demand. This means to incur every time in the order cost. On the contrary, the order cost in case of the (r,Q) model is less than 50% of the total annual inventory cost, on average (Appendix A.4 reports the detailed results). This can be appreciated by looking at Figure 6.2, noting that at the optimum point the order and on-hand costs almost have the same values.

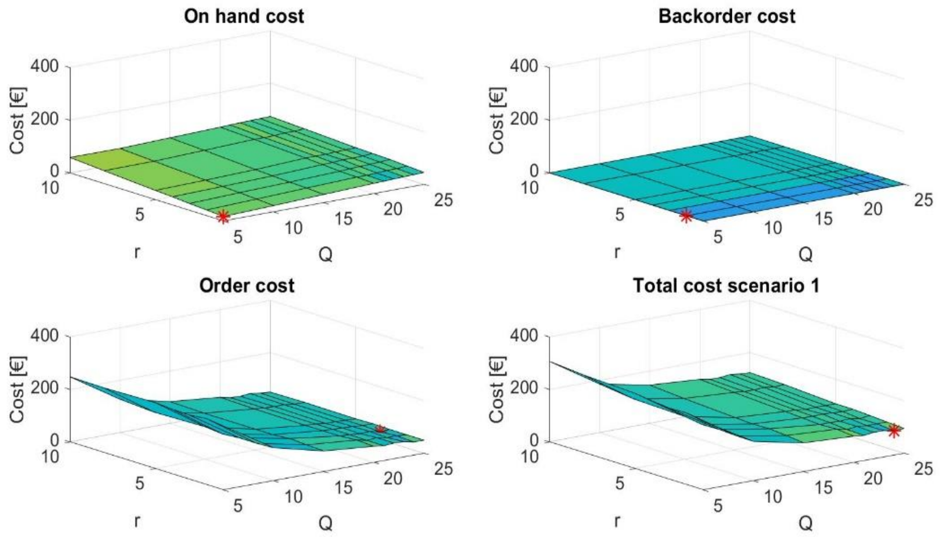


Figure 6.2: On-hand, Backorder, Order and Total (r,Q) inventory costs for the M1 model considering the cost scenario 1. The red star indicates the minimum cost point. CI inventory KPIs used for costs calculation: 3% of the mean value.

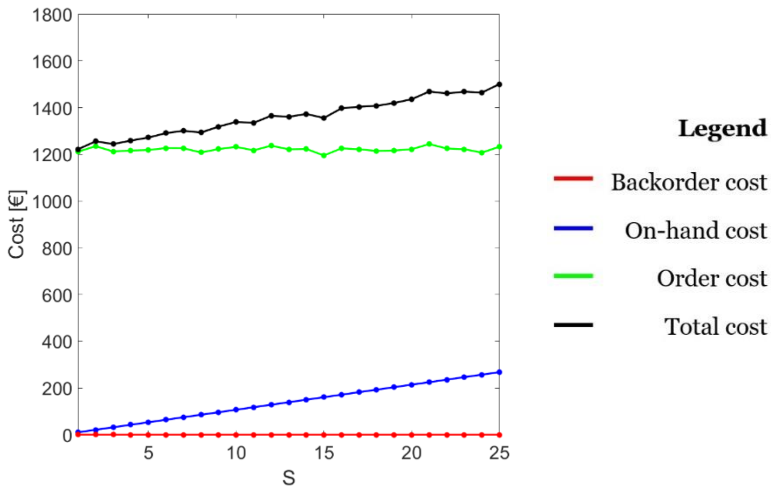


Figure 6.3: On-hand, Backorder, Order and Total (S-1,S) inventory costs for the M1 model considering the cost scenario 1. CI inventory KPIs used for costs calculation: 1.5% of the mean value.

It can be noted that the scenarios 3 and 4 are the ones with lower lot size. These scenarios are characterized by a high (H) on-hand cost and a low (L) order cost. Considering the relevant impact on the on-hand cost on the total inventory cost (63.7%), it is preferred to reduce the lot size and stock less units a year, eventually placing more than one order per year.

For both the models, it is always selected the lowest reorder point, equal to 1. This result can be easily justified considering the M1 low annual demand and its low production time (total printing time per piece of 5h), that can allow to maintain low on-hand inventory while successfully keeping the average number of backorders negligible, as can be observed by looking at Table 6.8.

M3 model

Table 6.9 reports the results obtained from the simulation models when M3 is produced. Also in this case, the (S-1,S) policy leads to a total annual inventory cost which is higher with respect to the (r,Q) one, being the maximum total percentage difference equal to 374.29% and being the (S-1,S) cost 4.7 times higher in scenario 5.

Scenario [b,h,o]	r	Q	Total (r,Q) Cost [€]	S	Total (S-1,S) Cost [€]
1 - LLL	5	5	1224.200	6	5644.82
2 - HLL	5	5	1237.69	7	5664.32
3 - HHL	5	5	1411.13	4	5892.46
4 - LHL	5	5	1397.64	4	5835.94
5 - LLH	5	5	1758.25	6	8339.27
6 - LHH	5	5	1931.69	4	8538.11
7 - HLH	5	5	1771.74	6	8361.43
8 - HHH	5	5	1945.18	4	8594.62

Table 6.9: Optimal inventory policy parameters and total annual inventory cost in case of Fubri M3 gear hob production. CI inventory KPIs used for costs calculation: 3% of the mean value for the (r,Q) model and 1.5% of the mean value for the (S-1,S) model.

Optimal (r,Q) parameters	Average Backorders [parts]	Average On-hand [parts]	Average Number of Orders
(r,Q) 5, 5	0.081	4.32	20.93
(S) 4	0.20	2.31	105.92
(S) 6	0.08	4.18	105.61
(S) 7	0.03	5.13	105.81

Table 6.10: Average backorders, on-hand and orders level evaluated in the (r,Q) and S optimum points for M3 model. CI: 3% of the mean value for the (r,Q) model and 1.5% for the (S-1,S) model.

The reason beside this result can be found again in the order cost, which is in this case the 95% of the total annual inventory cost for the (S-1,S) model (Appendix A.4 reports the detailed results). It should be pointed out that also for the (r,Q) model the order cost is one of the main drivers in determining the total annual inventory cost, being around the 84% of it (Appendix A.4 for the complete results). The importance of the order cost on the total cost computation can be appreciated looking at Figure 6.4 and Figure 6.5. In fact, the total cost has mostly the order cost shape.

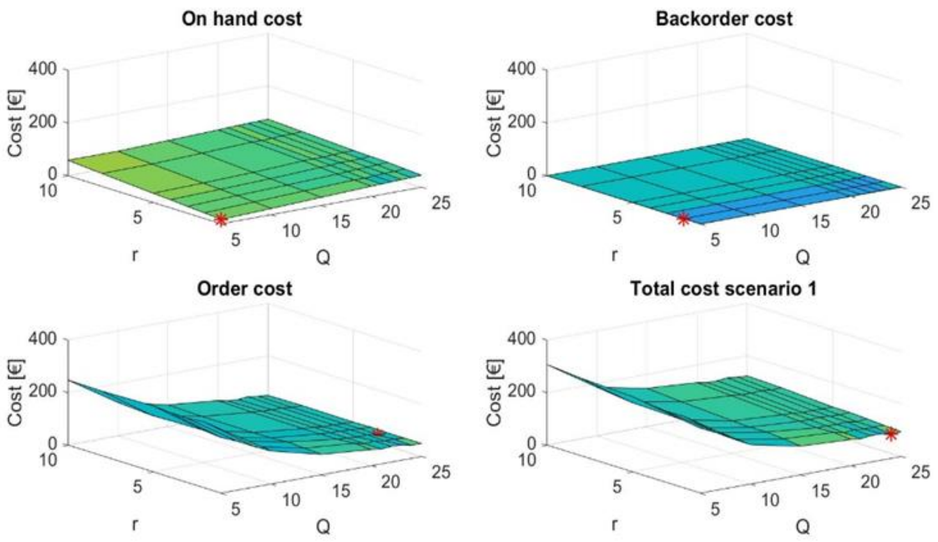


Figure 6.4: On-hand, Backorder, Order and Total (r,Q) inventory costs for the M3 model considering the cost scenario 1. The red star indicates the minimum cost point. CI inventory KPIs used for costs calculation: 3% of the mean value.

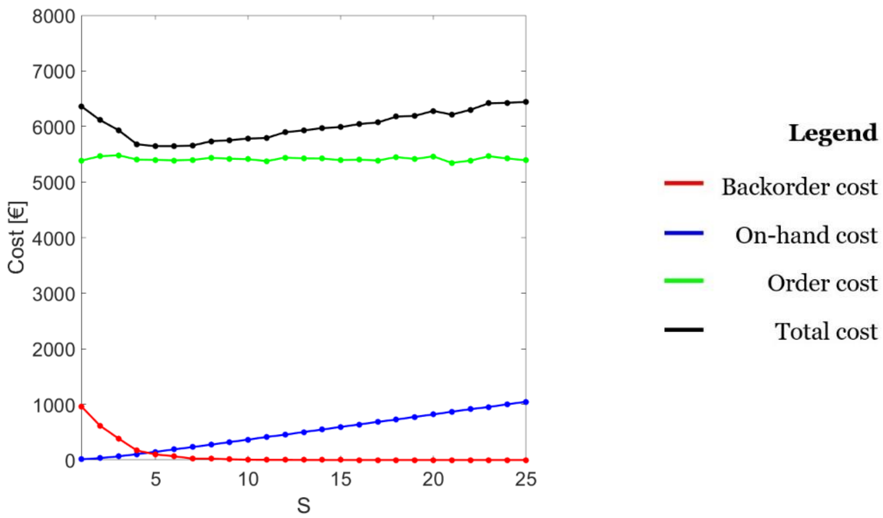


Figure 6.5: On-hand, Backorder, Order and Total (S-1,S) inventory costs for the M3 model considering the cost scenario 1. CI inventory KPIs used for costs calculation: 1.5% of the mean value.

This consideration justifies the decrease in the total inventory cost difference between the two models. In fact, being the M3 annual demand higher with respect to M1, and being the maximum order size equal to 5, it is necessary to place many more orders to satisfy the annual demand, incurring in higher order cost. This observation can be proved by looking at Table 6.10, noting that on average 20.93 orders per year should be placed in case of M3 production, in contrast with the around 1 order for M1. It is important to note that for every scenario, the maximum lot size *Q* has been selected. This indicates that optimizing the SLM building chamber has a positive impact on reducing the inventory cost.

Furthermore, higher reorder points are selected with respect to the M1 model. This result reflects the increase of the component annual demand, being of 106 pieces for M3 and only 24 for M1. In order to be able to satisfy the higher incurrence of demand, it is necessary to maintain a higher safety stock, especially when the backorder cost is high and the on-hand cost is low. In fact, it can be observed from Table 6.10 that increasing the target level S, the average number of backorders decreases and the safety stock increases, leading consequently to a decrease of the backorder cost while keeping low the on-hand costs.

Even if the general optimal reorder point result has the same cause found in the increase of annual demand, it should be pointed out that two different production strategies are designed. In particular, the reorder point *r* equal to 5 in the (r,Q) model is strictly correlated to an increase of the resupply time mainly due to a longer production time. In fact, the production of five different M3 models in the same job would require 266.25h for just the SLM printing phase. This leads to stock more in order to be able to satisfy the demand. Different is the case of the (S-1,S) model. In fact, in this scenario the production time is related to just one M3 piece per time, being equal to 56.25h on average for the printing phase. Despite this, also this policy leads to select a target level *S* comparable with the (r,Q) policy. This result is caused by the longer queue and waiting times that characterize this production policy, as already mentioned in Section 6.2.1. Being the waiting times so impactful, these parameters provide a longer average resupply time in case of a make to order policy. This leads to select a higher target inventory level to maintain a sufficiently high stock level to promptly satisfy the demands, especially in those scenarios when the backorder cost is high and the on-hand cost is low.

6.2.3 Analysis conclusions

In conclusion, both the M1 and M3 test cases have demonstrated that applying the (r,Q) inventory policy consent to an overall decrease of the total annual inventory cost. This is mainly due to a reduction of the order cost impact on the total calculation, considering that the (r,Q) model would allow to place a reduced number of orders per year to satisfy the demand. Furthermore, considering the low annual demand, it is not necessary to have high safety stock level, leading to a low reorder point selection. On the contrary, for both the gear hobs models production, it is suggested to order the maximum lot size or a little less of it. This is justified considering that increasing the lot size means reducing the number of orders per year and therefore the order cost. In addition to this consideration, is has been observed that the lot size increase allows also to decrease the unitary production time and better exploit the SLM production capacity, obtaining in this way a more efficient system.

A further element that would suggest the selection of a higher lot size is the unitary production cost. In fact, it has been pointed out that a percentage of the unmelted metal powder is scrapped, quantity that decreases when many pieces are placed inside the building chamber. For this reason, the unitary production cost has been considered as a sum of a constant term, proportional to material, energy and post-processing operations cost, plus a term with an inverse proportionality with the lot size, linked to the scrap cost (Section 4.3.2). Therefore, by increasing the lot size, the unitary production cost tends to decrease, positively impacting the final on-hand and backorder costs calculation. On the contrary, in case of an (S-1,S) production, the scrap cost is paid as a whole meaning, especially for the M1 model, to spend more for the wasted material rather than for the one utilized for the part production. One solution to lower this cost in case of a piece-by-piece production can be the selection of a smaller building chamber volume. Having generally the industrial SLM machine a building plate dimension around 250x250mm^2 or even more, there exists some systems to reduce the building chamber printing volume. One of this, often used for research purpose to test different raw materials, is the Reduced Build Volume system (RBV), suitable for the Renishow AM250 machine. It consists in a smaller build plate equipped with specific powder tanks and a system of pistons to distribute the metal powder on a reduced area. The RBV application would therefore bring the advantage of reducing the wasted powder in case of piece by piece production, but has different drawbacks. One above all is the impossibility of preheating the baseplate, that can represent a technological problem. Indeed, this procedure is beneficial to reduce thermal stresses, and, for the processing of some metal alloys, it is desirable to avoid unwanted defects in the parts. Furthermore, having the RBV a complex dynamic for power spreading, the time requested for pistons movements and powder spreading is sensibly longer with respect to the one in the full-size machine: experts at Mechanical Engineering Department of Politecnico di Milano estimate it around 20 or 30s, on the opposite of the 3 or 4s for the full size. This parameter would therefore create difficulties in the RBV industrial applications, and also it would sensibly affect the total resupply time, probably leading to a further increase on the total inventory cost difference between the two inventory models.

In conclusion, the choice of a lot production shows positive outcomes from different points of view: it is possible to reduce the unitary resupply time, better optimizing the SLM building chamber and allocating the set up and cool down times on more pieces, reducing also in this way the waiting times; the unitary production cost can be reduced because of a reduction in material wastage and finally the total annual inventory cost is decreased because of a reduction of the order and stock costs. These results are the outcomes of the accurate SLM process analysis and modelling, that allowed the removal of the assumptions often found in literature as infinite machine capacity, negligible set up and post processing time, and generic or constant distributed resupply times.

7 Case study: impact of SLM modelling on inventory costs

The simulation tool has been selected to model the AM production together with the inventory resupply logic in their entireness. In this way, assumptions often found in literature as infinite SLM machine capacity, absence of set up and cool down phases, constant or generically distributed resupply time and neglected post-processing operations can be discarded, providing instead a more comprehensive representation of the scenario.

In Chapter 3, it has been underlined how a sufficiently detailed SLM production modelling can sensibly impact the total resupply time. It is considered interesting to appreciate if such an accurate representation could also have consequence on the total inventory cost. Therefore, in this chapter the results obtained with the MC study on the production and resupply times are extended considering the AM application for inventory management in its totality, including the post processing phases and the inventory logic. In particular, the analysis is focused on the (r,Q) policy, shown to be the most cost effective, and permitting also to broaden the spectrum of analysis considering different batch size and consequently different production scenarios.

7.1 Design of experiments

In this section, the different SLM machine modelling assumptions are described, and the definition of the analysed test cases reported.

7.1.1 SLM modelling

The MC models have demonstrated that the application of the hypoexponential distribution, and therefore of a detailed description of the SLM modelling, can sensibly reduce the error in the resupply time estimation with respect to the use of a more general distribution, as the exponential one. It has been decided to apply these two different SLM stochastic representations also in the developed (r,Q) simulation model to appreciate if they possibly impact the total annual inventory costs.

The two models are constructed in the following way:

i. **Detailed model:** the SLM production process is represented by the sum of different time steps: the set up and cool down phases, considered as

Exponential distributions with mean $1/\mu_{su}$ and $1/\mu_{cd}$, plus the printing phase, modelled as an Erlang distribution with mean m/μ_{prod}, where m is the number of strata to be printed. It is reminded that μ_{prod} includes both the strata scanning time, proportional to the number of pieces to be printed, plus the fixed powder spreading time. The sum of these phase type distribution corresponds to the hypoexponential modelling used in MC, as proved in the validation phase (Section 4.2.4).

ii. **Exponential model:** in this case, the SLM production is modelled using an Exponential distribution, made up of only one stage having mean equal to the average production time obtained in the detailed model:

$$\frac{1}{\mu_{Ex}} = \frac{1}{\mu_{avg}} = \frac{1}{\frac{1}{\mu_{su}} + \frac{1}{\mu_{cd}} + \frac{m}{\mu_{prod}}}$$

The post processing phases are modelled with exponential distributions with mean $1/\mu_{TT}$, $1/\mu_{EDM}$ and $1/\mu_{G}$ for thermal treatments, wire EDM and grinding phases respectively. The demand process is Poisson, with mean 1/λ (Section 3.2.1). For an accurate description of the assumptions' justifications, the reader is reminded and 4.2, where the simulation model is described in detail.

7.1.2 Test cases

The analysis of the suggested different SLM modelling will be applied on the production of Fubri's M1 and M3 gear hobs presented in the Case Study (Chapter 5), where all the input data utilized in terms of production and post processing duration, and costs estimation can be found.

In order to have a broader analysis perspective, different demand scenarios are considered for the components production. In this respect, four different test case have been identified, as reported in Table 7.1.

	M1 [Lv]	**M3 [Hv]**
Low annual demand [Ld]	87	87
High annual demand [Hd]	1752	135

Table 7.1: Different annual demand and production test cases.

As it can be noted from Table 7.1, the different scenarios are identified by the volume of production and the annual demand. In particular, the M1 gear hob has been labelled as a low volume [Lv] part, while the M3 as a high volume [Hv] one, being its volume around 14 times higher with respect to the M1 one (Table 5.4). This characteristic influences the production time: the M3 printing phase, in fact, has a total printing time around 14 times higher with respect to M3 model. In this way, different production rates are tested.

Additionally, two different demand scenarios are reported: a low annual demand case [Ld], with a correspondent mean interarrival time of 100h, and a high demand case [Hd], developed with the aim of saturating the SLM machine. It can be noted that the M3 model has a lower high annual demand with respect to the M1 one because, having a longer production time, the SLM machine gets saturated for lower demands rate and therefore higher interarrival times of orders issuing [65h for M3 compared with 5h for M1].

Considering the difficulty in estimate the inventory cost parameters as *b* and *h* or the SLM manufacturing costs, it has been decided to maintain the approach already described in Section 6.1 to calculate the total annual inventory costs. In this respect, eight different cost scenarios are defined (Table 6.3) considering all the possible combinations of the *b, h, o* parameters' values (Table 6.2).

The developed (r,Q) simulation models were tested with different (r,Q) parameters to find their optimal values. They are reported in Table 7.2.

	r	Q
M1 [Ld]	1, 2, 3, 5,6, 10	5, 10, 15, 20, 23, 24, 25
M1 [Hd]	From 24 to 30, from 32 to 37	20, 22, 23, 24, 25
M3 [Ld]	From 1 to 12	From 1 to 5
M3 [Hd]	From 6 to 11, from 14 to 16, from 19 to 27	From 1 to 5

Table 7.2: (r,Q) parameters tested for the different scenarios.

The data are collected at the steady state. For this reason, a warm-up period of 800,000h for the detailed model and of 1,000,000h for the exponential model are selected applying the Welch method. The results are related to 30 replications runs of one year length (1year = 8760h). Confidence intervals are calculated considering a confidence level of 95%.

7.2 Results and discussion

In this section, the impact of the different SLM modelling on the inventory costs is analysed and reported. In particular, it has been considering interesting to identify the bottleneck of the AM production system, being this machine the one determining the system performances.

7.2.1 Bottleneck identification

Before going through the SLM modelling and inventory cost analysis, it is interesting to observe some AM production line performance. In this respect, a study is conduced to identify the bottleneck machine between the SLM and the post-processing operations servers. The bottleneck represents a constraint in a production system because it conditions the overall line behaviour and performance. For this reason, its operations are the main focus of a manufacturing system analysis, to understand the general system behaviour and eventually improve its KPIs.

The bottleneck machine can be detected considering different performances: it is the server having the slower production rate, it is the most utilized and generally creates longer queues and waiting times (Urban and Rogowska, 2020). These main drivers are utilized in this work to identify the system bottleneck.

First, an analysis on production time is done. Table 7.3 reports all the servers' mean production times. In particular, the SLM one has been calculated as:

$$SLM_{pt} = t_{su} + t_{spread} * n_{layers} + t_{prod} * n_{layers} * Q + t_{cd}$$

considering the data reported for M1 and M3 test cases reported in Section 5.2.

	M1	M3
SLM	$6.67h + 3.75h * Q$	$9.17h + 52.5h * Q$
Thermal treatment (TT)	$3h$	$3h$
EDM	$0.084h * Q$	$0.415h * Q$
Grinding (G)	$0.8h * Q$	$2.4h * Q$

Table 7.3: Servers' production times for M1 and M3 production.

It can be noted that the SLM machine shows an average production time sensibly higher to the other ones, especially considering the case of having a high lot size Q.

It is also considered interesting to observe the machine utilizations, the queue length and buffers waiting times. These performances analysis is reported for the M1 production with high annual demand test case [Lv-Hd], being characterized by a high part request together with the possibility to place many pieces inside the SLM building chamber. In this way the system capabilities are fully exploited, observing with more evidence the results searched. It is anyway reminded that all the other test cases lead to the same conclusion. The analysis is performed considering the detailed model which better describes the SLM production process.

Different lot sizes are tested (Q = 6, 8, 10, 15, 20, 23, 24, 25) to observe the general system behaviour. Ten independent simulation runs for each Q parameters are done, considering a warm-up period of 800,000h and a running period of 30 years (1year = 8760h). The results confidence intervals (CI) are calculated with a confidence level of 95%. The KPIs collected are:

- SLM and post-processing machines utilization
- SLM and post-processing average orders in queue
- SLM and post-processing average waiting time in queue.

Server utilizations analysis

Figure 7.1 reports the average machine utilization for all the different servers, obtained with a confidence interval of around 1% of the mean. It is clear that the SLM machine is the most utilized server for all the lot sizes, while the other servers show significantly lower utilization.

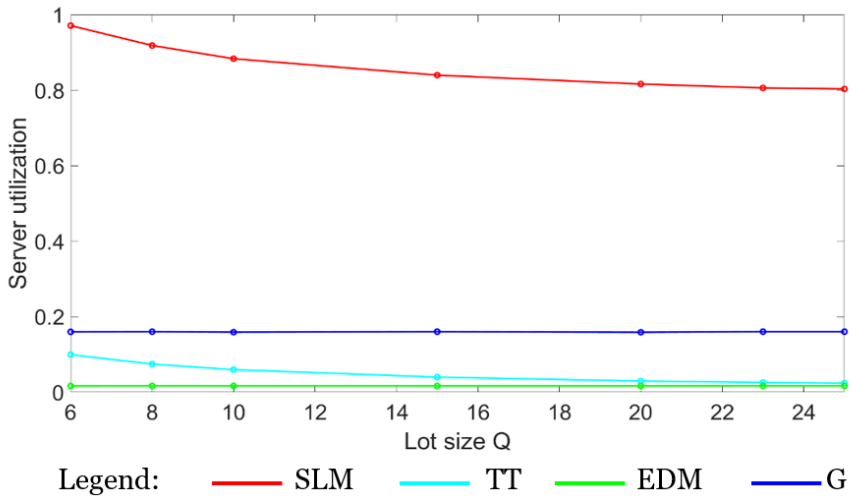
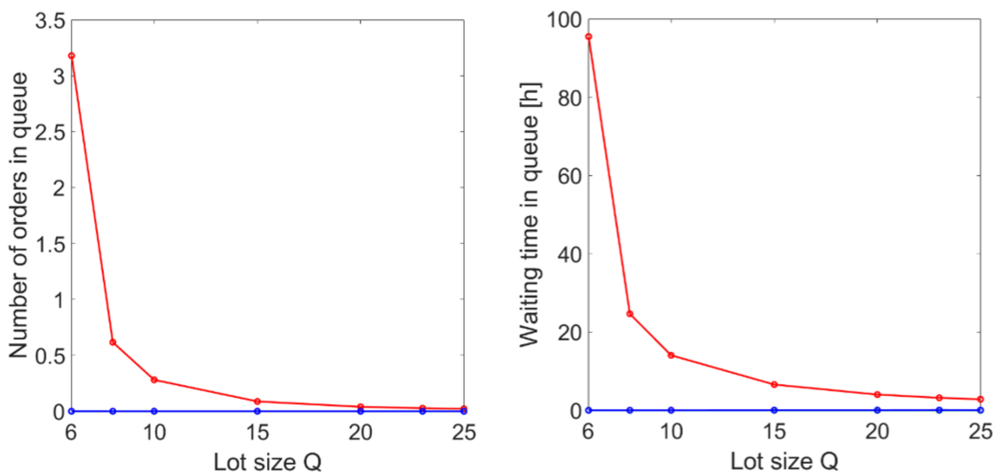


Figure 7.1: Server utilization (CI: 1% of the mean value).

Queue analysis

Figure 7.2 shows the average number of orders in queue (a) and the orders average waiting time (b) for all the AM production server. The results have a CI of around 10% over the mean. It is possible to note that these two KPIs for the thermal treatment (TT), wire EDM and grinding (G) operations overlap in figure, and are all null. On the contrary, the SLM machine presents up to 3.18 orders in queue with an



a) *b)*

Figure 7.2: a) Number of orders in queue; b) Waiting time in servers' queue. CI: 10% of the mean.

average waiting time of 95.48h when Q is equal to 6, and shows for all the lot size higher KPIs value with respect to the other servers.

The results obtained from the machine production time analysis, machine utilizations, number of orders in queue and orders waiting time in queue all agree in selecting the SLM machine as the bottleneck of the AM production system, being the slowest, most utilized, with longer queue and higher waiting times server.

7.2.2 Comparison between SLM production models

It is significant to appreciate the extent of the impact on the total annual inventory cost selecting the (r,Q) optimal parameters with an approximated model as the exponential one on the real production scenario, modelled instead by the hypoexponential distribution. This is done collecting the (r,Q) parameters that minimize the total annual cost using the exponential model, and evaluating the detailed model for the same (r,Q) parameters. The annual cost found with this method is then compared with the one obtained with the detailed model evaluated in its optimal parameters. It is reminded that this analysis is focused on the proper modelling of the SLM machine because this server is the bottleneck of the line: its performances influence the whole system behaviour, and therefore an accurate process study is necessary.

From a managerial point of view, this would mean to evaluate the possible error that an approximated model would cause in the ex-ante analysis performed for the optimal (r,Q) parameter selection, applied then on the real industrial SLM production. A possible total inventory cost error is calculated comparing this scenario with the one that uses a detailed model to optimize the inventory policy.

The maximum total cost differences obtained considering all the cost scenarios tested are reported in Table 7.4. The complete optimal parameters and total cost calculation can be found in the Appendix A.5.1.

		Total annual inventory cost difference	**Machine utilization**
M1	**Ld**	0.00%	4.01%
[Lv]	**Hd**	6.09%	80.18%
M3	**Ld**	13.43%	55.32%
[Hv]	**Hd**	65.04%	83.40%

Table 7.4: Maximum annual total inventory cost difference between the two SLM models and SLM machine utilization (CI: 3% of the mean value) in the different test cases.

It can be noted that the two models give the same results in case of low part volume and low annual demand, while the costs discrepancy increases by increasing these two parameters, having its maximum value in case of high volume and high annual demand.

The reason beside these results can be motivated considering the statistical distribution properties together with the SLM machine utilization, reported in Table 7.4 for the optimal exponential (r,Q) parameters obtained. It is firstly recalled that the MC model has already pointed out that the exponential and detailed models provide sensible different resupply times value when the production volume is high and the systems is saturated, so it is subject to high annual demand. This observation can easily reflect the results obtained in terms of total inventory cost difference. In fact, the inventory cost is strongly influenced by the average resupply time, being this duration the time requested to resupply the stock and therefore to be able to promptly satisfy customers demands. If the production volume is low, as in case of the M1 model, the error brought by the exponential model is limited. Additionally, considering this scenario in combination with a low annual demand [Lv-Ld], the direct consequence is a low SLM machine utilization, as reported in Table 7.4. From a practical point of view, these conditions are translated into a negligible SLM machine queue and waiting times, leading the average resupply time to be almost equal to the sum of the average production times. These observations can be proved by looking at the two models resupply times obtained for the optimal (r,Q) parameters of the detailed model, reported in Table 7.5. It can be noted, in fact, that the test case Ld-Lv shows an average resupply time difference equal to 1.0%, almost negligible from a statistical point of view. For the mentioned reasons, the two models lead to select the same (r,Q) optimal parameters, having a null difference on the total annual inventory cost.

		Detailed model		**Exponential model**	
		Optimal (r,Q)	Resupply time	Optimal (r,Q)	Resupply time
M1 [Lv]	**Ld**	(1,25)	126.39h	(1,25)	128.19h
	Hd	(29,25)	128.33h	(35,24)	317.13h
M3 [Hv]	**Ld**	(4,5)	301.85h	(9,5)	416.91h
	Hd	(7,5)	394.23	(25,5)	1009.20h

Table 7.5: Optimal (r,Q) detailed and exponential models parameters for the scenarios analyzed; Detailed and exponential models resupply time comparison calculated considering the optimal exponential (r,Q) parameters in the scenarios of analysis. CI: 3% of the mean for the detailed model, 6% of the mean for the exponential model.

On the contrary, when the demand or the production volume are increased (Lv-Hd and Hv-Ld cases), the total annual inventory cost starts to diverge, with a maximum difference of 6.09% and 13.43% (Table 7.4). This result can be motivated again by looking at the statistical distribution used. In fact, the exponential distribution leads to higher error when the average production time is increased [Hv], because of the increase of its variance. This error affects the system stochasticity leading to longer waiting times and therefore higher average resupply times, as demonstrated in Table

7.5 (416.91h for the exponential model versus 301.85h for the detailed model). Even if the production volume is low (M1 model) and therefore the average production time is reduced, the increase of the annual demand leads to an increase of the machine utilization, causing the propagation of the exponential error on all the production line, with longer buffer queues and consequently longer average resupply times. In fact, the exponential model leads to an average resupply time equal to 317.13h, while the detailed model to 128.33h for the same (r,Q) parameters (Table 7.5). The combination of both a high production volume and a high annual demand (Hv-Hd) shows the highest annual inventory cost difference, equal to 63.04%, being the exponential model resupply time more than double of the detailed model one (Table 7.5).

From an inventory management point of view, using an exponential model means to select higher reorder point. This can be observed by looking at Table 7.5, that collects the optimal (r,Q) parameters for the analysed scenarios, while the complete optimal parameters results are available in the Appendix A.5.1. As expected, the Lv-Ld case leads to the same optimal parameters selection. For the other cases, it can be noted that the lot size choice in both the models is equal or almost equal. The most sensible discrepancy is found by looking at the reorder point *r*, being the exponential one always higher. This is a direct consequence of the higher average resupply time pointed out before. In fact, the exponential model requires waiting longer before the orders come to replenish the stock. In order to be able to satisfy the customer demand avoiding to incur in high backorder cost, it is therefore better to stock more and have a higher safety stock. The direct consequence is an increase of the on-hand cost, leading to higher total annual inventory cost difference.

Service level analysis

One of the often studied inventory performances is the service level, defined as the ratio between the demand actually met and the overall demand. With this metric it is possible to measure the ability to promptly respond to customers' demand, not incurring in backorder. It is considered interesting to analyse the service level behaviour in case of the SLM process modelling with a detailed or approximated model.

Figure 7.3 shows the service level results for the different test cases. In particular, the green surface represents the detailed model, while the meshed white one the exponential model. The results are compared with a target service level of 95%, described by the red surface in figure. By looking at Figure 7.3, it is possible to observe how the exponential model's service level is always lower with respect to the detailed one, especially for the high demand test cases (Figure 7.3 c and d). These results are in accordance with the analysis already done for the two models comparison. In fact, having the exponential model a higher variability, the average resupply time is much longer with respect to the detailed one, especially when the

SLM machine is highly utilized. This result in a lower system capability to respond just in time to clients' requests, leading to higher average backorders cost and lower service level.

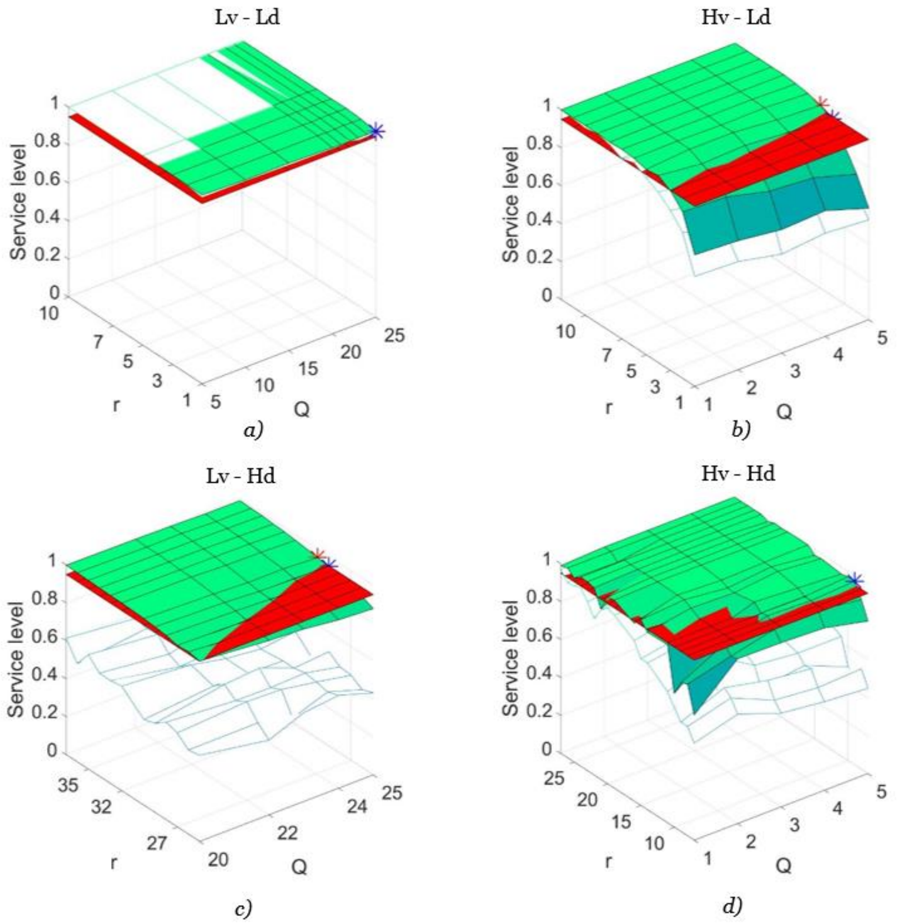


Figure 7.3: Service level for the different test cases: a) Lv-Ld, b) Hv-Ld, c) Lv-hd, d) Hv-Hd. The green graph represents the detailed model, while the meshed one the exponential model. The red plane is the target service level, equal to 0.95. The blue and the pink stars are the service level for the optimal r,Q parameters for the detailed and exponential model respectively, while the red and the black ones are the new optimal points for respecting the service level constraint. CI: 3% of the mean value for the detailed model and a,b exponential model; 8% of the mean for the exponential c,d cases.

To avoid this disservice, managers often define a target service level to meet. In this work it is selected equal to 95% and represented by the red surface in Figure 7.3. It is possible to appreciate that the detailed model is able to satisfy the service level constraint in most of the cases for the already selected optimal (r,Q) parameters (blue stars in Figure 7.3), or eventually it is possible to improve the result to meet the constrain (red star in Figure 7.3). This is not the case of the exponential model: the obtained service level for the high demand cases is not sufficient to meet the requirements, meaning that higher reorder points must be simulated and selected to satisfy the target, further increasing the annual inventory holding cost and finally the total annual inventory cost difference between the two models.

7.2.3 Alternative SLM modelling: the normal distribution test

For sake of completeness, statistical distributions other than the exponential family ones are tested. In particular, it is interesting to appreciate the consequences on the inventory costs when a distribution with lower variance with respect to the exponential one is applied to model the service times, evaluating its comparison with the selected detailed model.

One of the most used statistical distribution to describe operations as sum of different tasks is the normal distribution. To maintain a low variability, this distribution is created so that its coefficient of variation (CV), which is the ratio of the standard deviation to the mean, is equal to 0.1. This distribution type is applied to all the process stages of the AM production. In particular, having pointed out the relevance of a detailed description of the SLM printing steps and willing to provide a sufficiently accurate comparison with the already developed detailed model, the AM technology is again modelled as the sum of tree steps: set up, printing (layer scanning and powder spreading) and cool down. All these steps and also the post-processing operations are described by a normal distribution having mean equal to the detailed model one, and CV = 0.1. Since the normal distribution has an infinite domain, the left tail that goes to minus infinity has been cut, selecting a new interval domain (0, ∞) for generating the random variables representing the production time durations. The demand process is still considered Poisson, being such a distribution appropriate for describing the erratic demand arrivals.

The test cases studied are the same as reported in Section 7.1.2 and the same analysis is followed. The data are collected at the steady state. For this reason, a warm-up period of 800,000h is selected. Results are related to 30 simulation runs of one year length (1year = 8760h). Confidence intervals are with a confidence level of 95%.

Table 7.6 reports the maximum total annual inventory cost difference evaluated considering all the cost scenarios studied. The results comparison is performed following the same procedure presented in the exponential and detailed model analysis. The complete results for all the scenarios tested with the respective optimal (r,Q) parameters can be found in Appendix A.5.2.

Total annual inventory cost difference	**M1 [Lv]**	**M3 [Hv]**
Low annual demand [Ld]	0.00%	3.35%
High annual demand [Hd]	1.10%	5.32%

Table 7.6: Maximum annual total inventory cost difference between the detailed and normal models.

By looking at the results collected in Table 7.6 it is possible to appreciate that the two models total annual cost differences are significantly lower with respect to the same comparison performed with the exponential model (Table 7.4). In particular, the percentage difference reaches its maximum value again in the Hv-Hd case, but with a value equal to just 5.32%. Overall, it can be said that the annual cost difference between the two models is statistically negligible.

The motivation beside these results can be found observing the SLM distribution coefficient of variation. In particular, this analysis is mainly focused on the SLM machine and not on the post processing operations because the SLM machine has been recognized to be the bottleneck of the line, and therefore it is the main responsible for the resupply time behaviour and finally of the inventory dynamics and costs. The post processing operations servers, having an almost null queue for the selected optimal parameters, provide an operation time equal to the mean value for both the models, not impacting the difference in total resupply time.

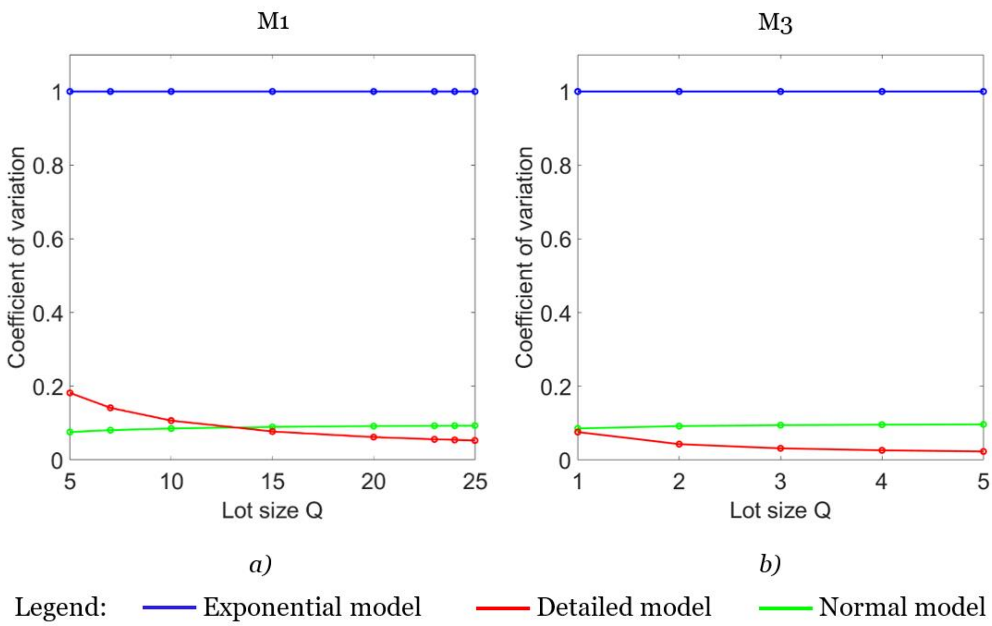


Figure 7.4: Coefficient of variation for the different SLM distribution studied for a) M1 production, b) M3 production.

Figure 7.4 shows the three distributions applied for the SLM production time modelling coefficient of variations, for M1 (Figure 7.4a) and M3 (Figure 7.4b) manufacturing. In particular, the CVs are calculated considering the whole SLM busy time, starting to set up and finishing with cool down phase. It is reminded that the mean production time varies with the lot size, and so does the CV. From Figure 7.4, it is possible to note that the detailed model's CV is always the lowest one considering the optimal lot size selected (Q = 25 for M1 and Q = 5 for M3), showing a value around or less than 0.1. This value is really similar to the normal distribution's CV, which is around 0.1 as expected. On the contrary, the exponential's CV, which is by definition equal to 1, is always sensibly higher with respect to the others two models. The CVs behaviour is reflected on the average resupply time evaluation. In fact, the exponential model leads to always higher average resupply time (Table 7.5) with respect to the detailed model, while the normal one provides very similar results (Table 7.7).

	Lv		Hv	
	Detailed	Normal	Detailed	Normal
Ld	126.39h	125.84h	301.85h	301.33h
Hd	128.33h	128.82h	394.23h	397.64h

Table 7.7: Detailed and normal models resupply time comparison calculated considering the optimal normal (r,Q) parameters in the scenarios of analysis. CI: 3% of the mean value.

7.2.4 Analysis conclusions

The analysis carried out demonstrated the importance of an accurate description of the SLM production times, defining in a detailed way all the time steps that this technology requires. It is possible to note that modelling the SLM manufacturing time by means of a generic statistical distribution as the exponential one can lead to sensible errors in resupply time estimation, especially when the production volume is high or the annual demand is conspicuous. In these scenarios, in fact, the SLM machine is mostly utilized, causing orders to wait in queue before being manufactured. The higher variability provided by a comprehensive statistical distribution as in the exponential case propagates in all the system dynamics, having as consequence a higher average resupply time. This result has repercussion on the total annual inventory cost and service level. In fact, being the average resupply time longer, the orders need more time to replenish the stock: to not incur in high annual backorders cost, it is preferable to select higher reorder point and therefore have a higher safety stock and finally a higher annual holding cost. Furthermore, for the same reasons, the service level metrics is often unsatisfactory in case of the exponential model application, requesting the selection of higher safety stock to improve the inventory performance, leading again to an inventory cost increase. Therefore, the use of a generic statistical distribution for modelling the production ex-ante would drive the decision makers to stock more than what should be necessary, incurring in not justified annual inventory cost.

In order to provide a more complete analysis, the detailed model's exponential family distribution has been substituted with normal ones having a low coefficient of variation. Despite this, no statistically relevant differences have been noted between the two models, neither in term of total average resupply time not on total annual inventory costs. This sensitivity analysis demonstrated that an accurate description on the SLM production process, which has been recognized to be the bottleneck of the AM manufacturing system, is the main focal point to derive accurate inventory management decisions, rather than the selection of the statistical distribution to model each one of the single process phases.

8 Conclusions

This work has posed as aim an in-depth analysis of the Additive Manufacturing (AM) production with specific focus on Selective Laser Melting (SLM), to evaluate the impact of this technology on Supply Chain (SC) and, more specifically, on inventory management. In particular, one of the objectives was to further investigate the identified literature gap between the studies evaluating possible AM consequences on SC and the ones focused on AM production efficiency. In fact, the firsts advocate AM as a make to order production, meaning that a job is run every time a demand occurs, while the latter suggests accumulating requests to fill up the AM machine building chamber to minimize the unitary production time, energy and material costs. Furthermore, researches that concern AM impact on SC often assert many assumptions, as infinite AM machine capacity, negligible set up times, not contemplated post processing operations and constant or generically distributed resupply times. It is interesting to appreciate if the removal of these hypotheses, providing instead an accurate SLM description as done in publications regarding AM production, can have relevant impacts also on inventory management.

Data and information on the SLM process have been gathered. SLM production is made by the sum of different steps: set up for machine pre-heating and input parameters loading, powder spreading and laser scanning for 3D objects manufacturing, and cool down for temperature reduction and parts extraction. These manufacturing stages have been modelled in a production contest, considering limited AM machine capacity and therefore possible queues formation and waiting times, by means of specifically designed Markov Chains. In this respect, different statistical distributions have been tested to model the manufacturing times. It has been noted how the Hypoexponential distribution, which allows a detailed description of the process assigning a rate for each of the technology's time phases, can lead to more accurate resupply time estimation. In fact, it has been proved that this model allows significant resupply time error reduction with respect to more comprehensive and general models that do not distinguish the different process time steps, as the use of an overall Exponential distribution. Furthermore, an accurate SLM machine analysis conducted with the developed model has pointed out how a reduction of the cool down time, possible thanks to the selection of high-performance machines, can positively impact the total resupply time, reducing this parameter significantly. This is an important outcome from an industrial scenario point of view, where the seek for efficiency and better performing systems is often desired.

Having underlined the importance of an appropriate description of the SLM production process, the impact of this technology on the inventory management has been analysed. In this respect, two different inventory policies have been selected to provide an analysis on both the production strategies identified in literature: the

(S-1,S), in which orders are issued every time a demand arrives, leading to a make to order or piece-by piece production, and the (r,Q), that instead accumulates demands placing orders with a batch size equal to Q. The scenario depicted consists in a single-item, single-location inventory managed by the AM manufacturer. Simulation tools have been selected to describe the complex system dynamics. In this way, it has been possible to consider the SLM limited capacity and therefore its waiting times and queues, post processing activities as thermal treatments and finishing operations, SLM machine set up and cool down times and finally SLM production times that vary with respect to the number of pieces allocated in the building chamber. An appropriate cost model has been defined to evaluate the SLM inventory costs, also taking care that a fraction of the unmelted raw powder is wasted, being this amount inversely proportional to the number of pieces in the building chamber.

To provide an analysis as close to reality as possible, a case study has been identified to test the developed models. In particular, two gear hobs tools manufactured by Fubri Company have been selected as target components. Production times and costs data were collected both interviewing experts at Politecnico di Milano Mechanical Engineering laboratories and considering a previously conducted experimental campaign.

The developed simulation models, applied to Fubri tools production and inventory management, are used to compare the (S-1,S) and (r,Q) inventory policies in terms of total annual inventory costs. The analysis revealed that the (r,Q) policy leads to lower annual inventory costs for both the components studied. This result is mainly due to the high impact of the order costs: with the (S-1,S) model, an order is placed every time a demand occurs, having as consequence an elevated incurrence of the order cost, that is not negligible with respect to the other inventory costs. Furthermore, following a make to order strategy would mean to not exploit the SLM machine capacity, requesting to wait the fixed time durations as set up, cool down and powder spreading for every order placement. Nesting many pieces in one job, following a (r,Q) policy, allows instead to optimize the fixed time durations, allocating them on all the parts to be printed. It has been in fact proved that the unitary resupply time sensibly decreases when the lot size is increased, thus improving the whole system efficiency. Finally, it is important to consider that locating many pieces in a single build job permits the reduction of the total raw material scrapped, leading to a decrease of the unitary production cost and consequently of the on-hand and backorder costs.

It was finally considered interesting to evaluate the impact of an accurate SLM modelling on the inventory costs estimation. In this respect, the more detailed Hypoexponential distribution and the comprehensive Exponential distribution modelling analysed with the Markov Chain, have been employed to represent the SLM production process in the wider scenario identified by the inventory replenishment. In particular, the focus on the SLM machine modelling has been further justified because this server has been observed to be the bottleneck of the AM manufacturing process, determining the system performance. The simulation results showed how an approximated modelling would lead to higher average resupply time and longer waiting times, especially when the production volume or the annual demand are high. This outcome has as a consequence an increase of the reorder point selection, causing higher and not justified annual inventory costs. Furthermore, the service level obtained with the approximated model is sensibly lower: to satisfy a target service level, the Exponential model would suggest stocking more units than the ones effectively needed. Finally, a conclusive analysis was performed replacing the Exponential family distributions with Normal distributions having the same mean and low variability, keeping the accurate description level given by the Hypoexponential model. Despite this, no statistically relevant differences were noted with respect to the detailed model in the optimal parameters estimation and in the total annual inventory cost computation. These results proved how a careful SLM production modelling, being the machine the line bottleneck, allows an improved evaluation of the optimal inventory policies parameters, providing decision makers an accurate tool to better define the inventory strategy.

This work made possible to further investigate the AM application in inventory management. In particular, the accurate technology modelling, removing assumptions as infinite machine capacity or negligible set up times, revealed that the reorder point, lot size inventory policy can reduce the annual inventory costs and allowed, in addition, a better optimal policy parameters estimation. It is known that the work focus is directed on SLM technology only, but it is believed that the developed models can be easily adapted to others AM technologies working similarly to SLM as SLS or STL, and additional insights can be gained. Furthermore, it would be interesting to broader the selected scenario to a multi-inventory system, introducing in this way the scheduling problem and the product mix management, but also fully exploiting the AM flexibility, technology able to produce many different parts and shapes in just one job without requiring specific tools or dies to be continuously changed and set up.

References

Abdulhameed, O. *et al.* (2019) 'Additive manufacturing: Challenges, trends, and applications', *Advances in Mechanical Engineering*, 11(2), pp. 1–27. doi: 10.1177/1687814018822880.

Adan, I. and Resing, J. (2015) *Queueing systems*, *Automatica*. Eindhoven, The Netherlands: Department of Mathematics and Computing Science Eindhoven University of Technology. doi: 10.1016/0005-1098(66)90015-X.

Di Angelo, L. and Di Stefano, P. (2011) 'A neural network-based build time estimator for layer manufactured objects', *International Journal of Advanced Manufacturing Technology*, 57(1–4), pp. 215–224. doi: 10.1007/s00170-011-3284-8.

Di Angelo, L., Di Stefano, P. and Guardiani, E. (2019) 'A build time estimator for Additive Manufacturing', *2019 IEEE International Workshop on Metrology for Industry 4.0 and IoT, MetroInd 4.0 and IoT 2019 - Proceedings*, pp. 344–349. doi: 10.1109/METROI4.2019.8792907.

Ashour Pour, M. *et al.* (2017) 'Additive manufacturing impacts on a two-level supply chain', *International Journal of Systems Science: Operations and Logistics*, 6(1), pp. 1–14. doi: 10.1080/23302674.2017.1340985.

Axsater, S. (2000) *Inventory Control*, *International series in operations research & management science*. Springer Science+Business Media New York.

Ayathamaraju, S. and Demir, A. G. (2020) *Study of Fabrication and Economic Modelling of Gear Hob with Selective Laser Melting Process*. Master Thesis, Politecnico di Milano.

Banoriya, D., Purohit, R. and Dwivedi, R. K. (2015) 'Modern Trends in Rapid Prototyping for Biomedical Applications', *Materials Today: Proceedings*, 2(4–5), pp. 3409–3418. doi: 10.1016/j.matpr.2015.07.316.

Baumers, M. *et al.* (2011) 'Energy inputs to additive manufacturing: Does capacity utilization matter?', *22nd Annual International Solid Freeform Fabrication Symposium - An Additive Manufacturing Conference, SFF 2011*, (January 2015), pp. 30–40.

Baumers, M. and Holweg, M. (2019) 'On the economics of additive manufacturing: Experimental findings', *Journal of Operations Management*, 65(8), pp. 794–809. doi: 10.1002/joom.1053.

Benita M. Beamon (1999) 'Measuring Supply Chain Performance', *International Journal of Operations & Production Management*, 19(3), p. 65.

Blos, M. F., da Silva, R. M. and Wee, H. M. (2018) 'A framework for designing supply chain disruptions management considering productive systems and carrier viewpoints', *International Journal of Production Research*, 56(15), pp. 5045–5061. doi: 10.1080/00207543.2018.1442943.

Breuer, L. and Baum, D. (2005) *An introduction to queueing theory and matrix-analytic methods, An Introduction to Queueing Theory and Matrix-Analytic Methods*. doi: 10.1007/1-4020-3631-0.

Brika, S. *et al*. (2017) 'Multi-Objective Build Orientation Optimization for Powder Bed Fusion by Laser', *Industrial Engineering & Management*, 6(4). doi: 10.4172/2169-0316.1000236.

De Brito, F. M. *et al*. (2019) 'An optimization model for the design of additive manufacturing supply chains', *IEEE International Conference on Industrial Informatics (INDIN)*, 2019-July, pp. 881–885. doi: 10.1109/INDIN41052.2019.8972028.

Buchholz, P., Kriege, J. and Felko, I. (2014) 'Phase-Type Distributions', in *Input Modeling with Phse-ype distributions and Markov Models*. Springer. doi: 10.1007/978-3-319-06674-5.

Çakanyildirim, M., Bookbinder, J. H. and Gerchak, Y. (2000) 'Continuous review inventory models where random lead time depends on lot size and reserved capacity', *International Journal of Production Economics*, 68(3), pp. 217–228. doi: 10.1016/S0925-5273(99)00112-7.

Carvalho, H. *et al*. (2012) 'Supply chain redesign for resilience using simulation', *Computers and Industrial Engineering*, 62(1), pp. 329–341. doi: 10.1016/j.cie.2011.10.003.

Cestana, A. *et al*. (2019) 'Reducing resupply time with additive manufacturing in spare part supply chain', *IFAC-PapersOnLine*, 52(13), pp. 577–582. doi: 10.1016/j.ifacol.2019.11.220.

Chekurov, S. *et al*. (2018) 'The perceived value of additively manufactured digital spare parts in industry: An empirical investigation', *International Journal of Production Economics*, 205(July 2017), pp. 87–97. doi: 10.1016/j.ijpe.2018.09.008.

Chen, J. *et al*. (2019) 'Maintenance, repair, and operations parts inventory management in the era of industry 4.0', *IFAC-PapersOnLine*, 52(13), pp. 171–176. doi: 10.1016/j.ifacol.2019.11.171.

Chergui, A., Hadj-Hamou, K. and Vignat, F. (2018) 'Production scheduling and nesting in additive manufacturing', *Computers and Industrial Engineering*, 126(September), pp. 292–301. doi: 10.1016/j.cie.2018.09.048.

Cockx, R., Armbruster, D. and Bendul, J. C. (2019) 'Resource sharing as supply chain disruption risk management measure', *IFAC-PapersOnLine*, 52(13), pp. 802–807. doi: 10.1016/j.ifacol.2019.11.228.

Costabile, G. *et al*. (2016) 'Cost models of additive manufacturing: A literature review', *International Journal of Industrial Engineering Computations*, 8(2), pp. 263–282. doi: 10.5267/j.ijiec.2016.9.001.

Dolgui, A., Ivanov, D. and Sokolov, B. (2018) 'Ripple effect in the supply chain: an analysis and recent literature', *International Journal of Production Research*, 56(1–2), pp. 414–430. doi: 10.1080/00207543.2017.1387680.

Durach, C. F., Kurpjuweit, S. and Wagner, S. M. (2017) 'The impact of additive manufacturing on supply chains', *International Journal of Physical Distribution and Logistics Management*, 47(10), pp. 954–971. doi: 10.1108/IJPDLM-11-2016-0332.

Fera, M. *et al.* (2018) 'A modified genetic algorithm for time and cost optimization of an additive manufacturing single-machine scheduling', *International Journal of Industrial Engineering Computations*, 9(4), pp. 423–438. doi: 10.5267/j.ijiec.2018.1.001.

GE Additive (2020) *3D Printing in Aerospace & Aviation | GE Additive*. Available at: https://www.ge.com/additive/additive-manufacturing/industries/aviation-aerospace (Accessed: 28 December 2020).

Gertler, J. (2010) 'V-22 Osprey Tilt-Rotor Aircraft: Background and Issues for Congress', in *Library of Congress Congressional Research Service*.

Ghadge, A. *et al.* (2018) 'Impact of additive manufacturing on aircraft supply chain performance: A system dynamics approach', *Journal of Manufacturing Technology Management*, 29(5), pp. 846–865. doi: 10.1108/JMTM-07-2017-0143.

Gisario, A. *et al.* (2019) 'Metal additive manufacturing in the commercial aviation industry: A review', *Journal of Manufacturing Systems*, 53(June), pp. 124–149. doi: 10.1016/j.jmsy.2019.08.005.

Glock, C. H. (2012) 'Lead time reduction strategies in a single-vendor single-buyer integrated inventory model with lot size-dependent lead times and stochastic demand', *International Journal of Production Economics*, 136(1), pp. 37–44. doi: 10.1016/j.ijpe.2011.09.007.

Hadley, G. and Whitin, T. M. (1963) *Analysis of Inventory Systems*. Englewood Cliffs: Prentice-Hall.

Heinen, J. J. and Hoberg, K. (2019) 'Assessing the potential of additive manufacturing for the provision of spare parts', *Journal of Operations Management*, 65(8), pp. 810–826. doi: 10.1002/joom.1054.

Holmström, J. *et al.* (2010) 'Rapid manufacturing in the spare parts supply chain: Alternative approaches to capacity deployment', *Journal of Manufacturing Technology Management*, 21(6), pp. 687–697. doi: 10.1108/17410381011063996.

Hopkinson, N. and Dickens, P. (2003) 'Analysis of rapid manufacturing - Using layer manufacturing processes for production', *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 217(1), pp. 31–40. doi: 10.1243/095440603762554596.

Hosseini, S. *et al.* (2019) 'Resilient supplier selection and optimal order allocation under disruption risks', *International Journal of Production Economics*, 213(August 2018), pp. 124–137. doi: 10.1016/j.ijpe.2019.03.018.

Hosseini, S., Ivanov, D. and Dolgui, A. (2020) 'Ripple effect modelling of supplier disruption: integrated Markov chain and dynamic Bayesian network approach', *International Journal of Production Research*, 58(11), pp. 3284–3303. doi: 10.1080/00207543.2019.1661538.

van Houtum, G.-J. and Kranenburg, B. (2015) *Spare parts inventory control under system availability constraints*. 1st edn, *International Series in Operations Research & Management Science*. 1st edn. New York: Springer. doi: 10.1007/978-1-4899-7609-3.

Ivanov, D. *et al.* (2016) 'Dynamic recovery policies for time-critical supply chains under conditions of ripple effect', *International Journal of Production Research*, 54(23), pp. 7245–7258. doi: 10.1080/00207543.2016.1161253.

Ivanov, D. (2019) 'Disruption tails and revival policies: A simulation analysis of supply chain design and production-ordering systems in the recovery and post-disruption periods', *Computers and Industrial Engineering*, 127, pp. 558–570. doi: 10.1016/j.cie.2018.10.043.

Ivanov, D. (2020) 'Predicting the impacts of epidemic outbreaks on global supply chains: A simulation-based analysis on the coronavirus outbreak (COVID-19/SARS-CoV-2) case', *Transportation Research Part E: Logistics and Transportation Review*, 136. doi: 10.1016/j.tre.2020.101922.

Ivanov, D., Das, A. and Choi, T. M. (2018) 'New flexibility drivers for manufacturing, supply chain and service operations', *International Journal of Production Research*, 56(10), pp. 3359–3368. doi: 10.1080/00207543.2018.1457813.

Ivanov, D., Dolgui, A. and Sokolov, B. (2019) 'The impact of digital technology and Industry 4.0 on the ripple effect and supply chain risk analytics', *International Journal of Production Research*, 57(3), pp. 829–846. doi: 10.1080/00207543.2018.1488086.

Janssen, R. *et al.* (2014) 'TNO: The impact of 3-D printing on supply chain management'.

Jia, F. *et al.* (2016) 'Investigating the feasibility of supply chain-centric business models in 3D chocolate printing: A simulation study', *Technological Forecasting and Social Change*, 102, pp. 202–213. doi: 10.1016/j.techfore.2015.07.026.

Jiang, J., Xu, X. and Stringer, J. (2018) 'Support Structures for Additive Manufacturing: A Review', *Journal of Manufacturing and Materials Processing*, 2(4), p. 64. doi: 10.3390/jmmp2040064.

Karmarkar, U. S. . (1987) 'Lot Sizes , Lead Times and In-Process Inventories', *Management Science*, 33(3), pp. 409–418. Available at: http://www.jstor.org/stable/2631860.

Khajavi, S. H., Partanen, J. and Holmström, J. (2014) 'Additive manufacturing in the spare parts supply chain', *Computers in Industry*, 65(1), pp. 50–63. doi: 10.1016/j.compind.2013.07.008.

Kim, J. S. and Benton, W. C. (1995) 'Lot size dependent lead times in a Q, R inventory system', *International Journal of Production Research*, 33(1), pp. 41–58. doi: 10.1080/00207549508930136.

King, W. E. *et al.* (2015) 'Laser powder bed fusion additive manufacturing of metals; physics, computational, and materials challenges', *Applied Physics Reviews*, 2(4). doi: 10.1063/1.4937809.

Knofius, N., van der Heijden, M. C. and Zijm, W. H. M. (2019) 'Moving to additive manufacturing for spare parts supply', *Computers in Industry*, 113. doi: 10.1016/j.compind.2019.103134.

Kucukkoc, I. (2019) 'MILP models to minimise makespan in additive manufacturing machine scheduling problems', *Computers and Operations Research*, 105, pp. 58–67. doi: 10.1016/j.cor.2019.01.006.

Kunovjanek, M. and Reiner, G. (2020) 'How will the diffusion of additive manufacturing impact the raw material supply chain process?', *International Journal of Production Research*, 58(5), pp. 1540–1554. doi: 10.1080/00207543.2019.1661537.

Lancaster, R. *et al.* (2016) 'Structural integrity of an electron beam melted titanium alloy', *Materials*, 9(6). doi: 10.3390/ma9060470.

Law, A. M. (2007) *Simulation Modeling and Analysis*. McGraw-Hill, Series in Industrial Engineering and Management.

Li, X. and Zhang, K. (2018) 'Single batch processing machine scheduling with two-dimensional bin packing constraints', *International Journal of Production Economics*, 196(November 2017), pp. 113–121. doi: 10.1016/j.ijpe.2017.11.015.

Li, Y. *et al.* (2017) 'Additive manufacturing technology in spare parts supply chain: a comparative study', *International Journal of Production Research*, 55(5), pp. 1498–1515. doi: 10.1080/00207543.2016.1231433.

Liu, P. *et al.* (2014) 'The impact of additive manufacturing in the aircraft spare parts supply chain: Supply chain operation reference (scor) model based analysis', *Production Planning and Control*, 25, pp. 1169–1181. doi: 10.1080/09537287.2013.808835.

Lücker, F. (2019) 'Using inventory to mitigate the Ripple effect', *IFAC-PapersOnLine*, 52(13), pp. 1272–1276. doi: 10.1016/j.ifacol.2019.11.373.

Marchese, K., Crane, J. and Haley, C. (2015) '3D opportunity for the supply chain About', *Deloitte University Press*, pp. 1–16.

Marrey, M. *et al.* (2019) 'A framework for optimizing process parameters in powder bed fusion (PBF) process using artificial neural network (ANN)', *Procedia Manufacturing*, 34(June), pp. 505–515. doi: 10.1016/j.promfg.2019.06.214.

Mashhadi, A. R., Esmaeilian, B. and Behdad, S. (2015) 'Impact of Additive manufacturing adoption on future Supply Chains', in *International Manufacturing Science and Engineering Conference*. doi: 10.13140/RG.2.1.4187.5048.

Materialise (2013) *Iris van Herpen Debuts Wearable 3D Printed Pieces at Paris Fashion Week*. Available at: https://www.materialise.com/en/cases/iris-van-herpen-debuts-wearable-3d-printed-pieces-at-paris-fashion-week (Accessed: 28 December 2020).

Materialise (2015) *adidas Futurecraft: The Ultimate 3D-Printed Personalized Shoe | Materialise - Innovators you can count on*. Available at: https://www.materialise.com/en/cases/adidas-futurecraft-ultimate-3d-printed-personalized-shoe (Accessed: 28 December 2020).

Mentzer, J. T. *et al.* (2001) 'defining Supply Chain Management', *Journal of Business Logistic*, 22(2), pp. 1–25.

Muckstadt, J. A. and Sapra, A. (2010) *Principles of Inventory Management When You Are Down to Four, Order More, Analysis*. Edited by T. V. Mikosch, S. I. Resnick, M., and S. M. Robinson. Springer.

NASA (2020) *Marshall Space Flight Center*. Available at: https://www.nasa.gov/centers/marshall/home/index.html (Accessed: 28 December 2020).

Noblesse, A. M. *et al.* (2014) 'Lot sizing and lead time decisions in production/inventory systems', *International Journal of Production Economics*, 155, pp. 351–360. doi: 10.1016/j.ijpe.2014.04.027.

Nyamekye, P. *et al.* (2017) 'Preliminary Investigation on Life Cycle Inventory of Powder Bed Fusion of Stainless Steel', *Physics Procedia*, 89, pp. 108–121. doi: 10.1016/j.phpro.2017.08.017.

Pavlov, A. *et al.* (2019) 'Integrated detection of disruption scenarios, the ripple effect dispersal and recovery paths in supply chains', *Annals of Operations Research*. doi: 10.1007/s10479-019-03454-1.

Piili, H. *et al.* (2015) 'Cost Estimation of Laser Additive Manufacturing of Stainless Steel', *Physics Procedia*, 78, pp. 388–396. doi: 10.1016/j.phpro.2015.11.053.

Previtali, B. *et al.* (2017) 'Comparative Costs f Additive Manufacturing vs. Machining: the Case Study of the Production of Forming Dies for Tube', in *28th Annual International Solid Freeform Fabrication Symposium – An Additive Manufacturing Conference*, pp. 2816–2834.

Rickenbacher, L., Spierings, A. and Wegener, K. (2013) 'An integrated cost-model for selective laser melting (SLM)', *Rapid Prototyping Journal*, 19(3), pp. 208–214. doi: 10.1108/13552541311312201.

Ridinger, S. (2018) *NASA Advances Additive Manufacturing For Rocket Propulsion | NASA*. Available at: https://www.nasa.gov/centers/marshall/news/nasa-advances-additive-manufacturing-for-rocket-propulsion.html (Accessed: 28 December 2020).

Roca, B. J. *et al.* (2019) 'Technology cost drivers for a potential transition to decentralized manufacturing', *Additive Manufacturing*, 28, pp. 136–151. doi: 10.1016/j.addma.2019.04.010.

Rogers, H., Baricz, N. and Pawar, K. S. (2016) '3D printing services: classification, supply chain implications and research agenda', *International Journal of Physical Distribution and Logistics Management*, 46(10), pp. 886–907. doi: 10.1108/IJPDLM-07-2016-0210.

Ruffo, M., Tuck, C. and Hague, R. (2006) 'Cost estimation for rapid manufacturing - Laser sintering production for low to medium volumes', *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, 220(9), pp. 1417–1427. doi: 10.1243/09544054JEM517.

Ryan, M. J. *et al.* (2017) '3D printing the future: scenarios for supply chains reviewed', *International Journal of Physical Distribution and Logistics Management*, 47(10), pp. 992–1014. doi: 10.1108/IJPDLM-12-2016-0359.

Schultz, C. R. (1990) 'On the optimality of the (S — 1,S) policy', *Naval Research Logistic*, 37(5), pp. 715–723.

Sirichakwal, I. and Conner, B. (2016) 'Implications of additive manufacturing for spare parts inventory', *3D Printing and Additive Manufacturing*, 3(1), pp. 56–63. doi: 10.1089/3dp.2015.0035.

Stewart, W. J. (2009) 'Chapter 12. Queues with Phase-Type Laws: Neuts' Matrix-Geometric Method', in *Probability, Markov Chains, Queues, and Simulation*. Princeton University Press, pp. 444–474. doi: 10.1515/9781400832811-013.

Thompson, M. K. *et al.* (2016) 'Design for Additive Manufacturing: Trends, opportunities, considerations, and constraints', *CIRP Annals - Manufacturing Technology*, 65(2), pp. 737–760. doi: 10.1016/j.cirp.2016.05.004.

Tijms, H. C. (2003) *A First Course in Stochastic Models*, *Technometrics*. Wiley. doi: 10.1198/tech.2005.s293.

Togwe, T., Eveleigh, T. J. and Tanju, B. (2019) 'An Additive Manufacturing Spare Parts Inventory Model for an Aviation Use Case', *EMJ - Engineering Management Journal*, 31(1), pp. 69–80. doi: 10.1080/10429247.2019.1565618.

Urban, W. and Rogowska, P. (2020) 'Methodology for bottleneck identification in a production system when implementing TOC', *Engineering Management in Production and Services*, 12(2), pp. 74–82. doi: 10.2478/emj-2020-0012.

White, G. and Lynskey, D. (2013) 'Economic Analysis of Additive Manufacturing for Final Products: an Industrial Approach'. University of Pittsburgh, pp. 1–11. Available at: http://136.142.82.187/eng12/history/spring2013/pdf/3039.pdf.

Wohlers Report (2012) *Additive manufacturing and 3D printing, State of the industry*.

Zhang, Y. *et al.* (2015) 'Fast adaptive modeling method for build time estimation in Additive Manufacturing', *CIRP Journal of Manufacturing Science and Technology*, 10, pp. 49–60. doi: 10.1016/j.cirpj.2015.05.003.

Zhang, Y. *et al.* (2019) 'Modeling and analysis of the on-demand spare parts supply using additive manufacturing', *Rapid Prototyping Journal*, 25(3), pp. 473–487. doi: 10.1108/RPJ-01-2018-0027.

Zhang, Y. and Bernard, A. (2013) 'Generic build time estimation model for parts produced by SLS', in *High Value Manufacturing: Advanced Research in Virtual and Rapid Prototyping - Proceedings of the 6th International Conference on Advanced Research and Rapid Prototyping, VR@P 2013*, pp. 43–48. doi: 10.1201/b15961-10.

Appendix

A.1 Markov Chain Matlab Codes

This section reports the code for the Markov Chain analytical model computation, developer in Matlab and used for solving the models presented in Section 3.2.3. The implementation approach followed is the on suggested by (Stewart, 2009). It is presented the code formulated for the Hypoexponential distribution. The code for the Erlang model is really similar, having as only difference that the production rate would be equal for all the stages, as computed with the Equation (3.13) in Section 3.3.2.

```
clc
clear all

v = [20,100,400,700,1000,1300,1600,1900,2200,2500];
K = [20*ones(1,3),12*ones(1,8)];

%%
for j=v
%% variables definition
mu = 1/(0.003);
mu_su = 1/0.92;
mu_cd = 1/4.5;
lambda = 1/100;
r=j+2; %number of states

%% check stability
production_time_avg = 1/(mu_su) +1/(mu_cd) +(r-2)/mu;
mean_hypoexpo = production_time_avg;
var_hypoexpo = 1/(mu_su)^2 +1/(mu_cd)^2 +(r-2)/mu^2;
dev_standard_hypoexpo = sqrt(var_hypoexpo);
cv_hypoexpo = dev_standard_hypoexpo/mean_hypoexpo;
mu_avg = 1/production_time_avg;
rho((find(v==j))) = lambda/mu_avg;

if rho((find(v==j)))>=1
    error('The MC is not stable')
end

% % sub-matrix construction to obtain the transition rate matrix Q

% A0: sub-matrix that describes the transition from (k+1,n) state to
(k,1). It is the cool down stage.

 A_0(r,1)=mu_cd;

 %A1: it is the Q diagonal.
 diag_A1 = -(lambda+mu)*ones(1,r);
 diag_A1(1) = -(lambda+mu_su);
 diag_A1(end) = -(lambda+mu_cd);
 A1_diag = diag(diag_A1);
```

```
 diag_sup_A1 = mu*ones(1,r-1);
 diag_sup_A1(1) = mu_su;
 A1_diag_sup = diag(diag_sup_A1,1);
 A_1 = A1_diag+ A1_diag_sup;

 %A2: sub-matrix that described the transition from(k,i) to (k+1,i),
so clients arrivals.
 diag_A2 = lambda*ones(1,r);
 A_2 = diag(diag_A2);

 %B00: Q diagonal for the first arrival.
B_00 = -lambda;

 %B01: sub-matrix for the first arrival. Transition from (0,0) to
(1,1)
 B_01 = zeros(1,r);
 B_01(1) = lambda;


 %B10: sub-matrix for transition from (1,r) to (0,0)
 B_10(end) =mu_cd;

 %% Q building

 %k: parameter for Q dimensioning
k =K(find(v==j));

 %fist arrival

 Q_it = [B_00,B_01,zeros(1,(k+1)*r);
         B_10,A_1,A_2,zeros(r,(k)*r);
         zeros(r,1),A_0,A_1,A_2,zeros(r,r*(k-1))];

 %iteration for all the other arrivals

 for i = 1:k-1
     Q_it_new= [zeros(r,1),zeros(r,r*i),A_0,A_1,A_2,zeros(r,r*(k-
(i+1)))];
     Q_it= [Q_it; Q_it_new];
 end

%add last line for Q to be square
Q_end = [zeros(r,1),zeros(r,r*k),A_0,A_1];
Q = [Q_it; Q_end];

%% system solving
Q=Q';
[n,m]= size(Q);
if n~=m
    error ('Q is not square')
end

%vector of the known terms
b = zeros(n,1);
b(end) = 1;
Q(end,:)= 1;

pi = Q\b;
```

```
%% average clients' number
clients = [0];
for i = 1: (n-1)/r
clients = [clients,i*ones(1,r)];
end

avg_clients = clients*pi;

%Little's law and average resupply time computation
tao_avg_exact_model(find(v==j)) = avg_clients/lambda
end
```

A.2 Inventory policy formulas derivation

For sake of completeness, the formulas derivation for the (r,Q) model are reported in this section. The procedure suggested is the one of Muckstadt and Sapra, (2006).

A.2.1 (S-1,S) model formula derivation

This section provides the detailed (S-1,S) inventory model optimization problem formulation, following the Muckstadt and Sapra, (2010) approach.

It is reminded that the model formulation assumes that the Palm Theorem (Section 4.1.1, Equation (4.2)) holds. In particular, this remarkable result, stating that the resupply times are independent and identically distributed, allows the calculation of the steady state probability that x units are in resupply as follow

$$P\{X = x\} = p(x|\lambda\bar{\tau}) = e^{-\lambda\bar{\tau}} * \frac{(\lambda\bar{\tau})^x}{x!}$$

Being X the random variable defining the number of units in resupply and $\bar{\tau}$, the average resupply time. Thus, the probability that there are x unit in resupply is Poisson distributed with mean $\lambda\bar{\tau}$, i.e. it is not need to know the exact distribution of the resupply or procurement time, but just its mean value.

On the basis of this important result, it is possible to define the probability expressions of the backorders and on-hand inventory. At the steady state, it holds that:

$$B = (X - S)^+ = max\{0, \mathrm{X} - \mathrm{S}\} \qquad OH = (S - X)^+ = max\{0, \mathrm{S} - \mathrm{X}\}$$

In particular, backorders verify if and only if the unit in resupply x, are more than the one in stock S, i.e. $x > S$: this lead to unsatisfied demand. Otherwise, zero backers are expected. Consequently, the expected number of units backordered in steady state condition is:

$$B(S) = \sum_{x>S}(x - S) * p(x|\lambda\bar{\tau}) = \sum_{x>S}(x - S) * e^{-\lambda\bar{\tau}} * \frac{(\lambda\bar{\tau})^x}{x!}$$

After having defined the backorders expression, it is possible to describe the formula for the on-hand inventory. Recall that the inventory position S is the sum of the on-hand inventory (OH) plus on order (O) minus backorders (B). Therefore

$$OH(S) = S - O + B(S)$$

The on order inventory can be described by the expected value of the number of parts in resuplly, E[X]

$$E[X] = \int_0^{\infty} x\, e^{-\lambda\bar{\tau}} * \frac{(\lambda\bar{\tau})^x}{x!}\,\mathrm{dx}$$

All the needed information are provided in to define the total cost function, which should be minimized:

$$C_{tot} = C(OH) + C(B) = h[\,S - E[X] + B(S)] + bB(S)$$

Where h is the holding cost per year associated to every single unit stock in the inventory and b is the backorder cost associated to every unit backordered.

Considering this type of inventory model, an important measure which is often used is the fill rate, defined as follow (Muckstadt and Sapra, 2010):

Given a stock level of S, the fill rate, F(S), is the expected fraction of demands that can be satisfied immediately from on-hand stock.

As is intuitively clear, as s increases the fill rate will increase. Generally, in an optimization problem, one of the condition is to consider the fill rate to fulfill a certain level α, in order to be able to meet demands and satisfy customer requirements.

The above description provides all the elements useful to define the optimization problem in order to find the optimal inventory position S*

$$\min C_{tot}(s) = C_{OH}(S) + C_B(S) = h[\,S - E[X] + B(S)] + bB(S)$$
$$s.t.$$
$$F(S) > \alpha$$
$$S \geq 0 \;\; \forall\, S \in N$$

A.2.2 An approximated (r,Q) model when backordering is permitted: formulas derivation

In this Appendix section, the inventory policy formula of the first (r,Q) model presented in Section 4.1.2 are derived. The Muckstadt and Sapra, (2010) approach is followed. This model is easy to understand, the resolution methods are quite simple, and can be suitable for very low demands. Nevertheless, it is based on many assumptions: one must be aware of their impact on the policy parameters resulting values.

The key assumption in this model is that there is never more than one single order of size Q outstanding in any point in time. This implies that, whenever the reorder point r is reached, there are no orders outstanding, or, in other words, that the demand over the resupply time never exceed Q. This hypothesis has as consequence that, at the reorder point, the inventory position (I), which is the sum of the on-hand inventory (OH), the units ordered (O) minus backorders (B), is equal to the net inventory N, so OH minus B.

$$I = N = OH - B$$

Another assumption is that the reorder point r is non-negative. This is normally true in practice, because one will not normally wait until there are backorders to place orders. This assumption, together with the previous one, lead to affirm that, at the reorder point, the inventory position is exactly equal to the on-hand stock. This can be appreciated by looking at Figure A.1, that describes behaviour of systems inventory position and net inventory under this type of hypotheses.

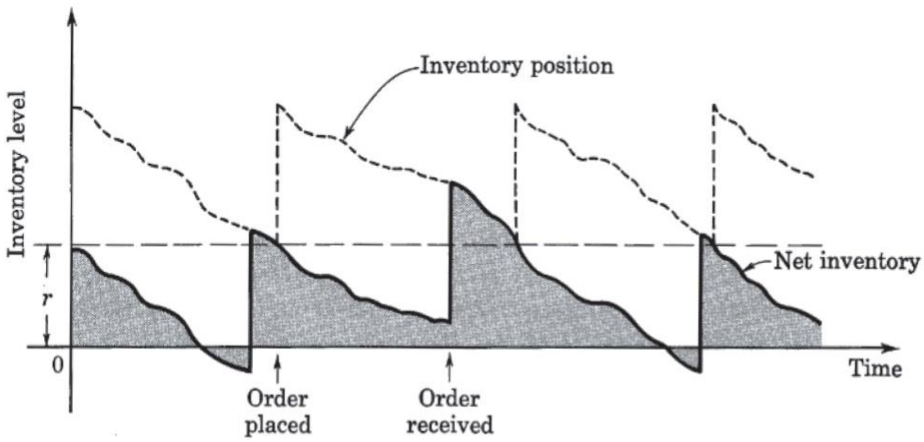


Figure A.1: Inventory dynamics (Hadley and Whitin, 1963).

To simplify the subsequent calculations, another assumption is made. Backorders, especially if expensive, are undesirable and generally occur only at the end of a cycle and in a very small amount, as described by Figure A.1. Therefore, the average number of backorders at random point in time is very small compared to the average amount of stock on-hand. For this reason, in the computation of the expected annual net inventory, the expected number of annual backorders is considered negligible:

$$E[Net\ inventory] = E[On\ hand] - E[Backorders]$$

But

$$E[Backorders] \approx 0$$

So

$$E[Net\ inventory] \approx E[On\ hand]$$

After having clarified the model assumptions, it is possible to derive the inventory costs formulations. It is reminded that, for this case, the resupply time τ is considered constant. Three different costs are considered: the order cost, the backorder cost and the holding cost.

The order cost is the annual cost associated to the total number of orders placed. Being the average annual demand rate λ and since an order is placed after Q demands, the average number of orders placed per year is $\frac{\lambda}{Q}$. Therefore, the average annual cost for placing orders C_O is:

$$C_o = K\frac{\lambda}{Q}.$$

To calculate the expected number of shortage cost per year, it is necessary to calculate the expected number of backorders per cycle, and multiply it by the number of cycles per year, which is again $\frac{\lambda}{Q}$ and by the backorder unit cost b.

A backorder verifies in case the on-hand inventory is not sufficient to meet the demand during the resupply time, so before the order already placed arrives. Let x the demand over the resupply time. A backorder therefore occurs if and only if $x > r$ during the resupply time and $x - r$ determines the number of backordered units. Hence

$$\int_r^{\infty} (x-r)f(x)dx$$

represents the number of backorders per cycle, where *f(x)* is the demand probability mass function. In case of discrete demand, it can be also written as

$$\sum_{x>r} (x-r)f(x)$$

Therefore, the total annual backorders cost is given by

$$C_B = b * \frac{\lambda}{Q}\int_r^{\infty} (x-r)f(x)dx$$

Finally, the holding costs is computed. It is reminded that, because the average number of outstanding backorders in a random point in time is considered negligible, the following relation holds:

$$E[Net\ inventory] \approx E[On\ hand]$$

Furthermore, the expected net inventory at the time of the order arrival is the safety stock, S, and it assume the value of $S + Q$ immediately after the order arrival. Because of the above considerations, these values are also assumed by the on-hand inventory. Having defined as *cycle* the time that elapses between the arrivals of two orders, and knowing that the mean demand rate μ in a cycle is constant, it is possible to calculate the average on-hand inventory during a cycle as:

$$\frac{S + (S + Q)}{2}$$

To completely formulate the optimization problem only in the r and Q decision variables, the relation that holds between r and S is applied to the previous formula. The safety stock by the time of the order arrival, considered the mentioned assumption, is equal to $r - x$, with x the demand occurred in the resupply time. The expected value of the safety stock, averaged on all the x is therefore:

$$S = \int_0^\infty (r - x) f(x) dx = r - \mu$$

Being μ the expected resupply time demand.

Consequently, the average annual cost of carrying inventory is:

$$C_{OH} = h * \left[\frac{Q}{2} + r - \mu\right]$$

By combining the above expressions, it is possible to obtain the average annual total inventory cost:

$$C_{tot}(Q, r) = \frac{K\lambda}{Q} + h\left[\frac{Q}{2} + r - \mu\right] + b\,\frac{\lambda}{Q}\int_r^\infty (x - r) f(x) dx$$

<u>Stochastic resupply time</u>

The above results were obtained considering a constant resupply time. It is possible that the replenishment time τ is stochastic, hence some modifications of the already described model are necessary. Considering the *resupply time independent from the demand,* the marginal distribution of the resupply time demand is:

$$h(x) = \int_0^\infty f(x, t) g(t) dt$$

Being *g(t)* the density function of the stochastic resupply time.

Therefore, the expected number of backorders would be:

$$\int_r^\infty (x - r) h(x) dx$$

And the safety stock

$$s = \int_0^\infty (r - x) h(x) dx = r - \mu^*$$

With μ^* the expected resupply time demand i.e. $\mu^* = \int_0^\infty x h(x) dx$

In the same way, it is possible to obtain the on-hand inventory expression C_{OH}:

$$C_{OH} = h * \left[\frac{Q}{2} + r - \mu^*\right]$$

with μ^* the expected resupply time demand i. e. $\mu^* = \int_0^\infty x h(x) dx$

The total annual average cost C_{tot}, considering stochastic resupply time, is calculated as:

$$C_{tot}(r,Q) = \frac{K\lambda}{Q} + h\left[\frac{Q}{2} + r - \mu^*\right] + b\,\frac{\lambda}{Q}\int_r^{\infty}(x-r)h(x)dx$$

A.2.3 The exact (r,Q) model: formulas derivation

The derivations that follows aim to provide an exact formulation of the total cost function of (r,Q) policy described in Section 4.1.2, removing the assumptions done in the approximated model.

Because of the loss of the approximated model assumption, it is necessary to determine the stationary distribution of the inventory position to determine the probability distribution of the inventory position and on-hand and backorders, which follow.

First of all, it is possible to note that the inventory position I random variable is always a value between $r+1$ and $r+Q$. In fact, when demand arises and lead the inventory position to reach r, an order of size Q is immediately placed, and therefore I does not assume the value of r for a positive amount of time.

Because the demand follows a Poisson process, the time between two demand is exponentially distributed with mean 1/λ. In addition, the time until the next demand is independent of the state of the system is in, generally called $r+j$.

Considering the above system properties, it is possible to state that the stationary distribution of the inventory position random variable is:

$$P[I = r + j] = \frac{1}{Q}$$

which is a uniform distribution between the values $r+1$ and $r+Q$.

The second random variable distribution that should be defined is the one of the net inventory N.

Suppose first that $n \le r$, being n the number of units in the net inventory. The inventory position at time $t-\tau$ can assume any value in $\{r+1, \dots, r+Q\}$ and have $N = n$ at time t. The probability that $N = n$ at time t, given that $I = r+j$ at time $t-\tau$, is $p(r+j-n;\ \lambda\tau)$, which is independent of t. Furthermore, $P[I = r+j] = 1/Q$ and therefore

$$P[N = n] = \sum_{j=1}^{Q} p(r+j-n;\ \lambda\tau)P[I = r+j]$$

$$= \frac{1}{Q}[\wp(r+1-n;\ \lambda\tau) - \wp(r+Q+1-n;\ \lambda\tau)]$$

Suppose now that $n > r$. Than, I cannot be less than n at time $t - \tau$, knowing that I = { n, n+1, ..., r+Q}. Consequently,

$$P[N = n] = \sum_{j=n-r}^{Q} p(r + j - n;\ \lambda\tau)$$

$$= \frac{1}{Q}[1 - \wp(r + Q - n + 1;\ \lambda\tau)]$$

When $r+1 \leq n \leq r+Q$.

Having described the stationary probability distribution of both the net inventory and the inventory position, it is possible to obtain the performance measure and formulate the total cost objective function, to be minimized.

The average annual cost for placing orders C_O is expressed by:

$$C_o = K\frac{\lambda}{Q}$$

Where the average number of orders placed per year is $\frac{\lambda}{Q}$.

Some different considerations should be taken into account for calculating the average annual backorder and holding costs.

Starting from backorders, the PASTA (Poisson Arrival See Time Averages) is recalled:

Let the probability that the system is out of stock at a random point in time be denoted by P_{out}. This probability is the same as the one that an arriving customer will find no stock on the shelf.

When N = −n, n = 0,1, ... , there is no stock on the shelf and therefore

$$P_{out} = \sum_{n=0}^{\infty} P[N = -n]$$

$$= \frac{1}{Q}\sum_{n=0}^{\infty}\{\wp(n + r + 1;\ \lambda\tau) - \wp(n + r + Q + 1;\ \lambda\tau\}$$

$$= \frac{1}{Q}[\sum_{u=r+1}^{\infty} \wp(u;\ \lambda\tau) - \sum_{u=r+Q+1}^{\infty} \wp(u;\ \lambda\tau)]$$

$$= \frac{1}{Q}[g(r) - g(r + Q)]$$

Where g(j) = $\sum_{u=j+1}^{\infty} \wp(u;\ \lambda\tau)$. Hence, *g(r)* measures the expected demand in excess of r over a period of length τ. Likewise, *g(r+Q)* equals the expected demand in excess of *r*+*Q* over a period of length τ.

Therefore, the expected annual number of backorders can be calculated as

$$E(r, Q) = \lambda * P_{Out}$$

In the same way, the average number of backorders outstanding at a random point in time is

$$B(r, Q) = \sum_{n=1}^{\infty} n * P[N = -n]$$

$$= \frac{1}{Q} * [G(r) - G(r + Q)]$$

where $G(j) = \sum_{u=j+1}^{\infty}(u - (j + 1))\wp(u; \lambda\tau)$. $G(j)$ measures the expected time-weighted demand in excess of j over a period of length τ.

Finally, the last performance measure to be computed is the on-hand inventory. By definition,

$$E[I] = OH(r, Q) + E[on\ order] - B(r, Q)$$

For Little's law, the annual expected value of the inventory on order can be described as the product between the annual average demand λ and the resupply τ, which is the average resupply time demand μ

$$E[on\ order] = \lambda\tau = \mu$$

The average expected value for the inventory position can be calculated knowing its stationary probability distribution:

$$E[I] = \sum_{j=r+1}^{Q} (r + j) * P[\, I = r + j] = \frac{1}{Q} \sum_{j=r+1}^{Q} (r + j) = \frac{1}{Q} * \left[Qr + Q * \frac{Q + 1}{2}\right] = \frac{Q + 1}{2} + r$$

From the above formulas, it is possible to obtain the on-hand inventory at random point in time:

$$OH(r, Q) = \frac{Q + 1}{2} + r - \mu + B(r, Q)$$

The computation of all the inventory performances has been exactly formulated. Hence, it is possible to define the final average total cost function, that in the optimization problem would be minimized:

$$C_{tot}(r, Q) = \frac{\lambda K}{Q} + h\left[\frac{Q + 1}{2} + r - \mu + B(r, Q)\right] + b * E(r, Q) + \hat{b} * B(r, Q)$$

A.3 Complete simulation output experiments for model validation

This Appendix section reports the complete simulation models runs results that have been employed in the validation phase described in Section 4.2.4.

Validation of the SLM production stream

Simulation models resupply time replications results [h]	Test case 1	Test case 2	Test case 3	Test case 4
	7.636323	2.111745	9.174763	45.0769
	7.577715	2.085308	8.893261	44.840518
	7.755062	2.127225	8.783352	46.788435
	7.700529	2.119320	8.942468	44.539702
	7.688830	2.103038	8.965659	44.851956
	7.699845	2.118615	8.758597	44.924959
	7.713898	2.092738	9.167568	46.866257
	7.818487	2.060994	9.264734	46.00563
	7.822045	2.102723	8.983220	46.024815
	7.598624	2.081675	9.016923	44.840834

Table A.1: Resupply time validation: simulation runs result of the different test cases.

Validation of the (S-1,S) inventory model

Simulation models Backorders replications results [parts]	Test case 1	Test case 2	Test case 3	Test case 4
	0.010656	0.000068	0.001060	0.000147
	0.036399	0.000054	0.000456	0.002191
	0.051919	0.000088	0.000936	0.005070
	0.008915	0.000127	0.000000	0.000000
	0.022894	0.000625	0.000000	0.003195
	0.019787	0.000150	0.000000	0.001083
	0.033014	0.000350	0.000000	0.000000
	0.030343	0.000389	0.000112	0.004377
	0.006852	0.000549	0.000777	0.000000
	0.024066	0.000210	0.000000	0.014911

Table A.2: (S-1,S) model validation: simulation runs backorders results of the different test cases.

Simulation models On-hand replications results [parts]

Test case 1	Test case 2	Test case 3	Test case 4
2.295311	1.893502	1.889423	2.605895
2.227857	1.883880	1.906325	2.597179
2.184047	1.896137	1.906489	2.514159
2.274004	1.886423	1.926663	2.534855
2.261280	1.887805	1.923549	2.589909
2.235021	1.896552	1.912352	2.607032
2.254071	1.895882	1.880724	2.711796
2.310388	1.896204	1.876739	2.393025
2.313879	1.874848	1.910650	2.576043
2.262665	1.894773	1.907950	2.511473

Table A.3: (S-1,S) model validation: simulation runs on-hand results of the different test cases.

Validation of the (r,Q) model

Simulation models Backorders replications results [parts]

Test case 1	Test case 2	Test case 3	Test case 4
0.521195	3.022794	0.100492	0.016389
0.619626	2.660226	0.096908	0.003142
0.466862	2.788286	0.124225	0.016689
0.483154	2.577331	0.115176	0.010463
0.579405	2.966326	0.084299	0.019993
0.477611	2.880315	0.12025	0.025276
0.546826	2.520177	0.101688	0.018145
0.508274	2.440996	0.151905	0.020724
0.557266	2.416517	0.095661	0.002748
0.452928	3.007472	0.100707	0.012505

Table A.4: (r,Q) model validation: simulation runs backorders results of the different test cases.

Simulation models On-hand replications results [parts]	Test case 1	Test case 2	Test case 3	Test case 4
	5.0401	3.0472	9.6593	7.56668
	4.8811	3.3835	9.7593	7.57061
	5.0538	3.0841	9.5401	7.58581
	5.2187	3.2573	9.3362	7.60892
	4.8928	2.9008	9.4642	7.31599
	5.2458	2.9627	9.5116	7.53234
	4.9843	3.3247	9.8144	7.34948
	5.0185	3.4875	9.2908	7.50396
	4.9481	3.5802	9.6713	7.80663
	5.2529	2.9314	9.8988	7.61838

Table A.5: (r,Q) model validation: simulation runs on hand results of the different test cases.

Simulation models Orders replications results [orders]	Test case 1	Test case 2	Test case 3	Test case 4
	144	149	86	87
	150	146	88	85
	145	148	88	87
	145	144	89	86
	151	149	87	87
	141	149	90	89
	146	144	84	89
	145	143	91	91
	147	141	85	82
	144	150	86	86

Table A.6: (r,Q) model validation: simulation runs orders results of the different test cases.

A.4 Inventory policies cost analysis

This Appendix section reports the complete backorder, on-hand and order percentage on the total annual inventory cost for both the (r,Q) and (S-1,S) model applied on M1 and M3 Fubri gear hobs production. The results are reported for all the eight scenarios tested, evaluated in the optimal inventory parameters.

M1 model

(r,Q) policy			
Scenario [b,h,o]	OH	B	O
1 - LLL	46.58%	0.00%	53.42%
2 - HLL	46.58%	0.00%	53.42%
3 - HHL	63.70%	0.01%	36.29%
4 - LHL	63.70%	0.01%	36.29%
5 - LLH	36.76%	0.00%	63.24%
6 - LHH	59.23%	0.00%	40.77%
7 - HLH	36.76%	0.00%	63.24%
8 - HHH	59.23%	0.00%	40.77%

(S-1,S) policy			
Scenario [b,h,o]	OH	B	O
1 - LLL	0.84%	0.01%	99.14%
2 - HLL	0.84%	0.02%	99.14%
3 - HHL	2.08%	0.02%	97.90%
4 - LHL	2.08%	0.01%	97.90%
5 - LLH	0.56%	0.01%	99.43%
6 - LHH	1.40%	0.01%	98.59%
7 - HLH	0.56%	0.01%	99.42%
8 - HHH	1.40%	0.01%	98.59%

Table A.7: (r,Q) and (S-1,S) on-hand (OH), backorder (B) and order (O) percentage of the total annual inventory cost evaluated in the optimal inventory parameters for the M1 production.

M3 model

(r,Q) policy			
Scenario [b,h,o]	OH	B	O
1 - LLL	9.45%	3.31%	87.25%
2 - HLL	9.34%	4.36%	86.30%
3 - HHL	20.48%	3.82%	75.69%
4 - LHL	20.68%	2.90%	76.42%
5 - LLH	6.58%	2.30%	91.12%
6 - LHH	14.96%	2.10%	82.94%
7 - HLH	6.53%	3.05%	90.43%
8 - HHH	14.86%	2.77%	82.37%

(S-1,S) policy			
Scenario [b,h,o]	OH	B	O
1 - LLL	3.36%	1.18%	95.47%
2 - HLL	4.11%	0.58%	95.31%
3 - HHL	4.45%	3.84%	91.72%
4 - LHL	4.49%	2.91%	92.60%
5 - LLH	2.27%	0.80%	96.93%
6 - LHH	3.07%	1.99%	94.94%
7 - HLH	2.27%	1.06%	96.67%
8 - HHH	3.05%	2.63%	94.32%

Table A.8: (r,Q) and (S-1,S) on-hand (OH), backorder (B) and order (O) percentage of the total annual inventory cost evaluated in the optimal inventory parameters for the M3 production.

A.5 (r,Q) resupply time modelling

This Appendix section reports the simulation model results for all the defined scenarios and test cases of Section 7.1.2 in their entireness. In particular, the optimal (r,Q) inventory parameters for both the exponential (and normal) and detailed models are collected, and the total annual inventory cost is computed. Subsequently, the total annual inventory cost of the detailed model evaluated in the optimal exponential (normal) parameters is reported, and the total cost difference between this value and the detailed one is computed. The scenario showing the highest delta costs, whose results are reported in Table 7.4 and Table 7.6 is highlighted in light blue.

A.5.1 Detailed and Exponential model

M1 production with low annual demand [Lv – Ld]

Scenario [b,h,o]	**Detailed**			**Exponential**			**Det. total cost for optimal Exp. (r,Q) [€]**	**Total cost delta %**
	r	Q	Total cost [€]	r	Q	Total cost [€]		
1 - LLL	1	25	216.06	1	25	223.84	216.06	0
2 - HLL	1	25	216.25	1	25	224.21	216.25	0
3 - HHL	1	25	279.28	1	25	287.90	279.28	0
4 - LHL	1	25	279.09	1	25	287.54	279.09	0
5 - LLH	1	25	302.80	1	25	313.99	302.80	0
6 - LHH	1	25	365.83	1	25	377.68	365.83	0
7 - HLH	1	25	302.99	1	25	314.35	302.99	0
8 - HHH	1	25	366.02	1	25	378.04	366.02	0

Table A.9: (r,Q) optimal parameters and respective total annual inventory cost for all the different scenarios, calculated with the detailed and exponential model for M1 production with low annual demand. The 8th column reports the detailed annual inventory cost evaluated in the exponential optimal (r,Q) parameters, while the 9th column is the total annual inventory cost percentage difference between the value obtained in the 8th column and the optimal detailed model one (4th column). (CI inventory KPIs: 3% of the mean value).

M1 production with high annual demand [Lv – Hd]

Scenario [b,h,o]	Detailed			Exponential			Det. total cost for optimal Exp. (r,Q) [€]	Total cost delta %
	r	Q	Total cost [€]	r	Q	Total cost [€]		
1 - LLL	29	25	3614.26	35	24	4742.42	3808.76	5.38
2 - HLL	29	25	3619.12	35	24	5065.72	3809.86	5.27
3 - HHL	29	25	3701.96	35	24	5139.94	3923.26	5.98
4 - LHL	29	25	3697.10	35	24	4816.64	3922.16	6.09
5 - LLH	29	25	5386.50	35	24	6603.95	5673.69	5.33
6 - LHH	29	25	5469.33	35	24	6678.17	5787.09	5.81
7 - HLH	29	25	5391.35	35	24	6927.25	5674.79	5.26
8 - HHH	29	25	5474.19	35	24	7001.47	5788.19	5.74

Table A.10: (r,Q) optimal parameters and respective total annual inventory cost for all the different scenarios, calculated with the detailed and exponential model for M1 production with high annual demand. The 8th column reports the detailed annual inventory cost evaluated in the exponential optimal (r,Q) parameters, while the 9th column is the total annual inventory cost percentage difference between the value obtained in the 8th column and the optimal detailed model one (4th column). (CI inventory KPIs: 3% of the mean value).

M3 production with low annual demand [Hv – Ld]

Scenario [b,h,o]	Detailed			Exponential			Det. total cost for optimal Exp. (r,Q) [€]	Total cost delta %
	r	Q	Total cost [€]	r	Q	Total cost [€]		
1 - LLL	4	5	1027.71	9	5	1173.10	1165.73	13.43
2 - HLL	4	5	1034.69	9	5	1186.95	1165.73	12.66
3 - HHL	4	5	1197.80	5	5	1474.38	1263.33	5.47
4 - LHL	4	5	1190.82	5	5	1404.35	1259.56	5.77
5 - LLH	4	5	1476.72	9	5	1625.51	1630.90	10.44
6 - LHH	4	5	1639.83	5	5	1835.50	1719.63	4.87
7 - HLH	4	5	1483.70	9	5	1639.37	1630.90	9.92
8 - HHH	4	5	1646.81	5	5	1905.53	1723.40	4.65

Table A.11: (r,Q) optimal parameters and respective total annual inventory cost for all the different scenarios, calculated with the detailed and exponential model for M3 production with low annual demand. The 8th column reports the detailed annual inventory cost evaluated in the exponential optimal (r,Q) parameters, while the 9th column is the total annual inventory cost percentage difference between the value obtained in the 8th column and the optimal detailed model one (4th column). (CI inventory KPIs: 3% of the mean value).

M3 production with high annual demand [Hv – Hd]

Scenario [b,h,o]	Detailed			Exponential			Det. total cost for optimal Exp. (r,Q) [€]	Total cost delta %
	r	Q	Total cost [€]	r	Q	Total cost [€]		
1 - LLL	8	5	1495.97	25	5	2037.60	1966.28	31.44
2 - HLL	8	5	1505.72	25	5	2096.64	1966.28	30.59
3 - HHL	8	5	1738.83	25	5	2813.19	2849.22	63.86
4 - LHL	7	5	1726.34	25	5	2754.14	2849.22	65.04
5 - LLH	8	5	2151.63	25	5	2728.97	2655.10	23.40
6 - LHH	8	5	2384.74	25	5	3445.52	3538.04	48.36
7 - HLH	8	5	2161.38	25	5	2788.02	2655.10	22.84
8 - HHH	8	5	2394.49	25	5	3504.57	3538.04	47.76

Table A.12: (r,Q) optimal parameters and respective total annual inventory cost for all the different scenarios, calculated with the detailed and exponential model for M3 production with high annual demand. The 8th column reports the detailed annual inventory cost evaluated in the exponential optimal (r,Q) parameters, while the 9th column is the total annual inventory cost percentage difference between the value obtained in the 8th column and the optimal detailed model one (4th column). (CI inventory KPIs: 3% of the mean value).

A.5.2 Detailed and Normal model

M1 production with low annual demand [Lv – Ld]

Scenario [b,h,o]	Detailed			Normal			Det. total cost for optimal Norm. (r,Q) [€]	Total cost delta %
	r	Q	Total cost [€]	r	Q	Total cost [€]		
1 - LLL	1	25	216.06	1	25	215.59	216.06	0.00
2 - HLL	1	25	216.25	1	25	215.80	216.25	0.00
3 - HHL	1	25	279.28	1	25	278.03	279.28	0.00
4 - LHL	1	25	279.09	1	25	277.82	279.09	0.00
5 - LLH	1	25	302.80	1	25	302.33	302.80	0.00
6 - LHH	1	25	365.83	1	25	364.56	365.83	0.00
7 - HLH	1	25	302.99	1	25	302.54	302.99	0.00
8 - HHH	1	25	366.02	1	25	364.77	366.02	0.00

Table A.13: (r,Q) optimal parameters and respective total annual inventory cost for all the different scenarios, calculate with the detailed and normal model for M1 production with low annual demand. The 8th column reports the detailed annual inventory cost evaluated in the normal optimal (r,Q) parameters, while the 9th column is the total annual inventory cost percentage difference between the value obtained in the 8th column and the optimal detailed model one (4th column). (CI inventory KPIs: 3% of the mean value).

M1 production with high annual demand [Lv – Hd]

Scenario [b,h,o]	Detailed			Normal			Det. total cost for optimal Norm. (r,Q) [€]	Total cost delta %
	r	Q	Total cost [€]	r	Q	Total cost [€]		
1 - LLL	29	25	3614.26	28	25	3617.18	3622.93	0.24
2 - HLL	29	25	3619.12	28	25	3620.87	3629.06	0.27
3 - HHL	29	25	3701.96	28	25	3698.06	3706.42	0.12
4 - LHL	29	25	3697.10	25	25	3693.04	3737.87	1.10
5 - LLH	29	25	5386.50	28	25	5394.52	5399.42	0.24
6 - LHH	29	25	5469.33	28	25	5471.72	5476.78	0.14
7 - HLH	29	25	5391.35	28	25	5398.20	5405.55	0.26
8 - HHH	29	25	5474.19	28	25	5475.40	5482.91	0.16

Table A.14: (r,Q) optimal parameters and respective total annual inventory cost for all the different scenarios, calculate with the detailed and normal model for M1 production with high annual demand. The 8th column reports the detailed annual inventory cost evaluated in the normal optimal (r,Q) parameters, while the 9th column is the total annual inventory cost percentage difference between the value obtained in the 8th column and the optimal detailed model one (4th column). (CI inventory KPIs: 3% of the mean value).

M3 production with low annual demand [Hv – Ld]

Scenario [b,h,o]	Detailed			Normal			Det. total cost for optimal Norm. (r,Q) [€]	Total cost delta %
	r	Q	Total cost [€]	r	Q	Total cost [€]		
1 - LLL	4	5	1027.71	4	5	1051.00	1027.71	0.00
2 - HLL	4	5	1034.69	4	5	1059.73	1034.69	0.00
3 - HHL	4	5	1197.80	3	5	1200.37	1234.04	3.03
4 - LHL	4	5	1190.82	3	5	1178.11	1208.68	1.50
5 - LLH	4	5	1476.72	7	5	1497.43	1526.21	3.35
6 - LHH	4	5	1639.83	3	5	1631.37	1675.55	2.18
7 - HLH	4	5	1483.70	7	5	1498.03	1526.96	2.92
8 - HHH	4	5	1646.81	3	5	1653.64	1700.91	3.29

Table A.15: (r,Q) optial parameters and respective total annual inventory cost for all the different scenarios, calculate with the detailed and normal model for M3 production with low annual demand. The 8th column reports the detailed annual inventory cost evaluated in the normal optimal (r,Q) parameters, while the 9th column is the total annual inventory cost percentage difference between the value obtained in the 8th column and the optimal detailed model one (4th column). (CI inventory KPIs: 3% of the mean value).

M3 production with high annual demand [Hv – Hd]

Scenario [b,h,o]	Detailed			Normal			Det. total cost for optimal Norm. (r,Q) [€]	Total cost delta %
	r	Q	Total cost [€]	r	Q	Total cost [€]		
1 - LLL	8	5	1495.97	9	5	1573.39	1569.47	4.91
2 - HLL	8	5	1505.72	9	5	1586.27	1583.43	5.16
3 - HHL	8	5	1738.83	9	5	1839.78	1831.30	5.32
4 - LHL	7	5	1726.34	7	5	1807.20	1726.34	0.00
5 - LLH	8	5	2151.63	9	5	2256.26	2250.64	4.60
6 - LHH	8	5	2384.74	7	5	2479.87	2400.71	0.67
7 - HLH	8	5	2161.38	9	5	2269.14	2264.60	4.78
8 - HHH	8	5	2394.49	9	5	2522.65	2512.47	4.93

Table A.16: (r,Q) optimal parameters and respective total annual inventory cost for all the different scenarios, calculate with the detailed and normal model for M3 production with high annual demand. The 8th column reports the detailed annual inventory cost evaluated in the normal optimal (r,Q) parameters, while the 9th column is the total annual inventory cost percentage difference between the value obtained in the 8th column and the optimal detailed model one (4th column). (CI inventory KPIs: 3% of the mean value).