From the Foundational Theories to the Prospects of the DDMRP: A Review and Research Agenda

Olufemi Adetunji
Department of Industrial and Systems Engineering
University of Pretoria, Pretoria 0002, South Africa
olufemi.adetunji@up.ac.za

Abstract

A review of the principles and techniques for Manufacturing Flow was made from the earlier Reorder Point (RoP) to the Demand Driven Materials Requirement Planning (DDMRP). The hermeneutics double loop review procedure was used to progressively analyse and build the body of knowledge, following a chronology of key events and challenges leading from ROP to DDMRP. Some key principles of the DDMRP were presented in response to some of these challenges, followed by a discussion of opportunities for research that exists in the implementation of DDMRP. Findings show that there are opportunities to strengthen the theoretical underpinning of this technique rather than depending mainly on the experiential judgment of those implementing it. Also, more efforts need to be invested in how to understand the similarities of the different MFC techniques so as to create a more integrated and responsive approach, rather than focusing on their differences as seems to be currently done in most research. In addition, making inventory decisions independent of extant capacity considerations (and vice versa) is unlikely to produce an optimal result, hence, the need for a more integrated approach to the planning and control of the manufacturing system. The main limitation to this research is that not all existing MFC have been reviewed, but only those that may be considered as foundational to the evolution of the DDMRP logic, hence, the motivation for the selection of the review approach as well. This can be expanded in subsequent studies, where more comparative studies can be done.

Keywords
Manufacturing Flow Control, Reorder Point, Material Requirement Planning, Theory of Constraints, Demand Driven MRP, Lean Manufacturing

1. Introduction
Manufacturing Flow Control (MFC) is a difficult decision. Often, process managers have to make trade off in many ways because there is no easy solution to the complicated system this problem portends with its numerous interacting issues: e.g. ensuring product availability while saving costs; utilising capacity while limiting explosive inventory growth; responding quickly to customers while managing long manufacturing lead time; making products to stock or commencing manufacturing only when demand is known; and many other related problems. Many techniques have been proposed to manage the flow of materials to achieve organisational goals while balancing these multi-faceted objectives in magical manner. One of the most recent techniques that is getting popular is the Demand Driven Material Requirement Planning (DDMRP). The testament to its acceptance includes the facts that respected organisations like the Association of Supply Chain Management (ASCM, formed from many popular organisations like the American Production and Inventory Management Society, APICS; Supply Chain Council, SCC; and some others) have embraced it, and a number of leading Enterprise Systems vendors (e.g. SAP) are beginning to implement it in their offerings.

Every MFC (e.g. Lean Manufacturing (LM), Theory of Constraints (TOC), Constant Work in Process (CONWIP), Paired-Cell Overlapping Loop with Card Authorisation (POLCA), Control of Balance by Card Based Navigation (COBACABANA), and even the Re-Order Point principle, RoP, and DDMRP) has some built in philosophies that guide its operations, such as the treatment of capacity and stock, financial accounting principles, making to stock or to order, planning with forecast or actual demand (or some mix), and many other issues. There have always been arguments about which technique works best under different conditions. The question, probably, should rather change to how do we exploit the strengths to mitigate the weaknesses of each of these techniques, and how do we build a resilient system based on this hybridisation. This literature study seeks to trace the evolution of many MFC techniques, starting from the foundational RoP principles as an extension of the Economic Order Quantity (EOQ) model, to the more recent DDMRP. In doing so, some commonalities of objectives and particular differences are identified, after which possible research agenda was proposed.

2. Foundational thought

The evolution of diverse MFC techniques can be construed as the continual search for a solution that could address the logic of planning and control of the activities of organisations involved in the production of goods and/or services. The solutions proposed at each stage of the evolution, as is typical of most scientific enquiry, usually lead to more questions; but in many such instances, also provides pathways for continual improvement of the understanding of the problem of interest, and hence, the inherent opportunities for solutions. This paper is an attempt to review the on-going effort in the quest to find a solution that addresses the planning and control of operational systems by both the industry and academia, so as to deliver superior value to the system stakeholders, which may include customers, shareholders and producers of goods and services. As a background for this review, it may still be imperative to provide concrete definitions or contexts for key terms in order to ensure commonality of opinions. This article is, thus, starts by discussing what a philosophy is, then a makes quick review of the meaning and scope of OM, after which the idea of an OM philosophy is formalised, and its importance is highlighted.

A philosophy can be considered as a way of thinking. It is a lens through which actions can be evaluated and conflicts resolved. It guides preferences of the operations manager in situations where objectives might conflict. In many situations where there may not be simple right or wrong options, philosophies become the filter through which choices may be made. While a philosophy may lead to superior thinking, it needs not necessarily be. It is like a culture that may not necessarily be better than another, but provides a framework through which a group of people may be able to determine what is considered acceptable or not, especially when they need to make complex decisions. Herein lies its importance, because it enables decisions be logical trade-offs when there many possible choices and consequences of such, hence, different individuals may be able to arrive at similar conclusions.

It may be pertinent to discuss organisational flows and the main business functional area before discussing the scope of OM. The activities of an enterprise are built around three main flows: information, material and funds. The business wants to understand what its strategic priorities are as this guides the choices. It, therefore, needs to understand the customer's needs, its own goals, and how to reconcile the two. While there are other information flows in the enterprise, demand is the main information the organisation needs as signals for production, and the functional area in charge of this is usually sales and marketing. The organisational response leads to the flow of materials, and this is the main duty of the operations units of the business and, hence, the OM requirements of the enterprise. The main custodian of material flow in an enterprise is therefore the operations unit, and to manage this flow, there is the planning and control (P&C) function at its heart. The third flow relates to financial flow and the financial units are in charge of this. They fund operations, sales and marketing functions, and are the custodians of the costing information that is necessary to optimise the diverse processes of the organisation.

The main function of P&C is the creation and protection of material flows through effective planning, execution and control mechanisms, subject to organisational objectives. There is the need to also formalise the concept organisational practice so as to firmly discuss the scope of operations management and decision making. Operations are decomposed to activities and it is practices that guides how exactly activities are implemented. For example, in the procurement process, there may be the need to communicate a purchase order, which may be sent vial email, a formal letter or even a simple WhatsApp message. The decision of how exactly this is done is the practice of that activity, and this is governed by the philosophy. Practice goes a long way in influencing the effectiveness and efficiency of an organisational process, and hence, may be common, good, bad, obsolete, leading or emerging.

The main purpose of planning and control is the determination of the level (quantity) and frequency (timing) of flow of materials, and the protection of the flow in response to organisational objectives. Many processes have implications for this flow and be impacted by it. Order generation influences decision about the quantity and timing. Capacity influences production level attainable. The level and timing of production influences inventory build-up. Stock level influences the timing of sales and delivery. Facility location and sizes have implications on the level and speed of flow. Product quality affects the planned level and timing of flow. Operations activities manages the flow of materials as it gets transformed through organisational processes, material build-up and customer response speed.

OM uses framework consisting of inputs, transformation, output and control structures, but our preferred framework is that by Pycraft et al (2010), which separates the input resources into two categories: the transformed inputs and the transforming inputs. This separation is important as is provides management views for the resources. While the main concern in the management of transformed resources is inventory, transforming resources are capacity driven. These two views are not independent of each other, and that leads to the concept of operational slacks.

Webster (2008) stated that system slacks exist in the form of any operational resources that are idle, underutilised or non-value adding. Slacks are present due to the lack of alignment between demand and supply, due to demand uncertainty, process variability, time lag between demand and supply, economy of scale opportunities, or operational pressure due to strategic choices, e.g. response time. Slacks provide system stability due to their ability to absorb shocks, however, this comes at cost. The slacks may be in the form of extra inventory, capacity or time. All these are convertible between each other, and are all reconcilable by cost equivalence, and hence, the possibility of substituting one by the other. In instances where there needs to be strategic trade-off between slacks, which one is permissible or forbidden is usually predicated on the philosophy of the organisation. The slack invariance principle implies that operational slacks can be measured in three basic units and they can be converted from one form to the other: material, time and monetary units. Consequently, inventory is material units, has an equivalent number of days of demand, equivalent machine hours (time units), or monetary value. This is similar for machines, warehouses, trucks etc. This is an interesting abstraction that helps a manager to reconcile resources in the organisation with the demand signal and make decisions about which resource is needed and in what forms and quantities operational slacks should exist in order to achieve the organisational objectives in the most efficient manner.

All enterprise activities are to add some value to the material that flows through them. The value added by an activity can be perceived from the concept of utility. Marketing and sales add value by creating utility of possession, which expands the market reach. Manufacturing adds value by creating the utility of form (changing the product into the form it is needed) and time (through the regulation of speed of production). Transportation adds utility of space (getting the product into spaces beyond the manufacturing location) and time (through the regulation of the speed of delivery), while inventory mainly adds utility of time (making the material available when needed (Coyle et al, 2003). From the transformation framework's perspective, it is apparent that while transforming resources may contribute utilities of form, space or time, transformed resources contribute utilities of time. The operations manager may then decide whether to use extra capacity, inventory, or time through discounting to meet a demand spike; hence, what type and quantity of slack to deploy (Webster, 2008). An interesting principle in OM is that the supply response pattern may be completely different from the demand signal received. The important thing is the search for the most efficient supply plan (Chase et al 2011). This is where operations research interfaces with operations management.

The main purpose of the information flow is to enable the operations function to make appropriate material flow plan in response to demand information, hence, the focus of all OM processes and philosophies is the utilisation of the enterprise resources to generate and protect this flow in the most efficient manner. Planning, execution and control activities of OM is the main tool at the heart of generating and protecting this flow, and hence, the focus of this review. A common pattern can be seen in how this is achieved by all OM philosophies, and by understanding this commonality, it should be easier to synchronise the techniques generate and protect material flow. Each OM philosophy has a view about how each resource, and hence, how the slacks are traded-off against each other. The main challenge of P&C is reconciling demand with supply, and in cases where alignment is difficult, this guides which slacks are preferred. For instance, while Lean Manufacturing (LM) will have pockets of excess capacity to handle demand uncertainty, Material Requirement Planning (MRP), will likely support production plan with extra inventory. While LM typically responds to actual demand signal, MRP typically forecasts demand and produces ahead. It is important to state that one philosophy is not necessarily better than the other, it is the context that matters. Even more important is the understanding that no philosophy seeks to be wasteful of capacity or inventory, but rather create flow in response to demand, and to protect the flow through their slacks preference.

Consequently, an operations management philosophy may be defined as a framework used to guide decisions of an enterprise about creating material flow in response to demand information, and to protect the flow from failure due to demand signal distortions and supply variabilities, using preferable system slacks, in consonance with the philosophical view, while utilising the transforming and transformed resources in the most efficient manner.

Thus far, the first section presents research background, while the second section discusses the foundational thought by reviewing the core concepts. Section 3 presents the methodology that is adopted for the literature review. It provides a motivation for the choice of the methodology and how it helps to achieve the review purpose of section 4. Section 4 reviews the relevant articles in details. Section 5 is a collation of ideas of research opportunities derived from section 3, after which the study is concluded in section 6.

3. Review methodology

Given that the objective of this review was mainly to trace the evolution of popular MFCs and get to the DDMRP, the hermeneutics approach was selected. The Systematic Literature Review (SLR) approach is more linear and rigid. Hermeneutics proposes that the interpretation of the literature should be done in the in the context of its text. Moreover, it is an iterative process in which the scope and content of the review is not determined a priori, but progressively done as the review progresses. It makes use of iterative reading and understanding cycles, through which the scholar engages the text to understand the literature within the context of its writing, and continually expands the scope of the literature until saturation of the information is relatively attained. Also, unlike SLR where the goal is high repeatability of the search and exhaustiveness of the content, hermeneutics seeks out the relevant literature, filters the core, and from these, builds up the body of knowledge by progressively increasing the content through their connection to the selected core.

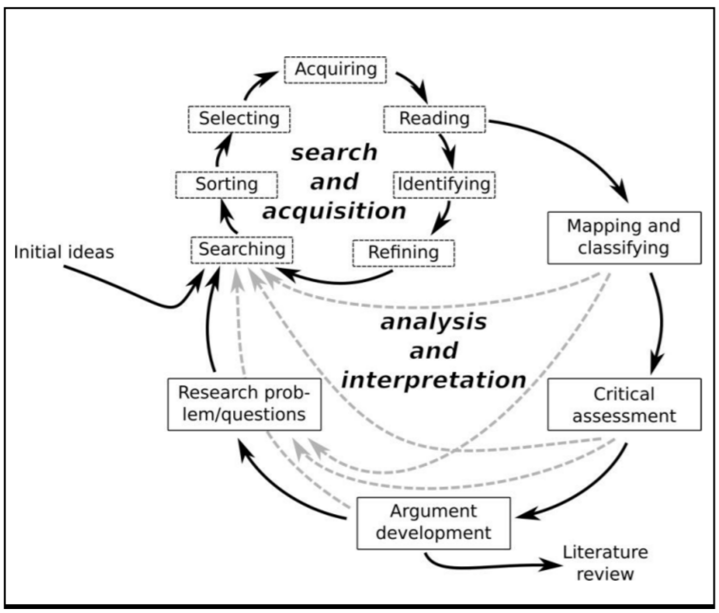

Figure 1. Hermeneutics Loop, source: Boell and Cecez-Kecmanovic (2014)

Hermeneutics is exploratory in nature. It makes use of the double discovery loop (Figure 1). The scholar engages the literature in a dialog using the analysis and interpretation loop (clarification and insight) to understand the literal and the contextual meaning of the texts of literature, and then uses the search and acquisition loop (information requirement) to expand the scope of the content, building up the whole picture from a deep understanding of the nature and context of the textual contents. It argues that there must be a convergence of the projection of the horizon of the scholar's perception and that of the text that is engaged, such that the range of vision of the context of the study aligns with the range of the content of the text, thereby constructing the big picture from an intimate understanding of the lower textual contexts.

This approach seems more ideal for this literature review for many reasons. Firstly, the study is exploratory, and there is the need to understand not only the terms, ideologies, and jargons, but also the contexts of their creation. Secondly, the ideologies evolved over time, and hence, the need to understand the historical context, chronology and evolution. Lastly, the ideologies were driven by factors that also drove the evolution of the social, business and technological environment, hence, it is difficult to understand the relevance and appeal of the philosophies without understanding the factors that incubated their evolution and acceptance, thus, the need to integrate multiple projections of realities.

3.1 Methodology application

When using hermeneutics, it is advisable to start from well-known and understood concepts. The most relevant materials are advised, and from this, other areas are identified through the iterative process of the hermeneutics loops. In contrast to SLR, having a well-focused starting point with few study materials is preferred to reading many materials at the beginning. In line with this, some well-known articles of OM philosophies and their history and evolution in the special issue of the Journal of Operations Management, Volume 25, Issue 2 of 2007 was used. All articles in this issue were skimmed for content as advocated in the hermeneutics approach and seven considered to be most relevant were identified as the core from which the hermeneutics iteration started. These papers were read in detail and more keywords were identified. These were used to find papers for the subsequent iterative cycle. In addition, reference tracking and snowballing were done to identify relevant literature for the next cycle of search and acquisition. Mapping was done by clustering the keywords, hence, new domains of related study were identified.

Skimming (or glancing) is the process of having a quick scan of the document to identify the most relevant ones for deeper study. This includes study of the abstract, the findings and conclusions, and a quick look at the methodology. It helps to identify which articles may be worth considering for the analysis and interpretation cycle. Reference tracking is the process of analysing the list of references of the core articles with the intention of finding common citations and classic articles from which the concepts were developed. This helps to have a better context of the terms and provides backward integration of the research in time. Citation tracking is the process of tracking the citations of the articles of interest such that the more recent articles that have cited the core articles are identified, and helps with forward integration of the core articles, hence, a basis for chronological alignment. Skimming was then repeated, and the cycle of analysis and interpretation was repeated again. The cycle was repeated until a reasonable level of information saturation was attained. As the number of keywords get bigger, affinity diagram and word cloud were used to create clusters and isolate new domains.

The four main philosophies that were extracted from the first loop were Reorder Point (RoP), Material Requirement Planning (MRP), Lean Manufacturing (LM) and Demand Driven Material Requirement Planning (DDMRP). For mapping of keywords and content, Boell and Cecez-Kecmanovic (2014) provided nine possible means of mapping and classifying the literature and keywords in the study. Of these, three were considered most relevant to this study: major concept, theoretical lens, and historical development.

4. Content Analysis

In order to properly analyse the texts of the articles, the scope of the study was limited to those related to the role of inventory management in the protection of material flow within the P&C framework. This is because most efforts publications on managing variability in material flow seems to have focussed on inventory. This helps the reading, interpretation and analysis the body of work thematically, after which the system slack perspective is discussed. The themes derived from the analysis of the texts starts from the foundational work in inventory management, and progresses to the different eras of the evolution of OM philosophies as discussed in the texts. The discussion then cycles back to general buffering using organisational resources and the need for an integrated view to take advantage of their substitutability when providing protecting material flow.

4.1 On Lot Sizing and Reorder point

The Economic Order Quantity (EOQ) model of Harris (1914) is the work-horse that laid the foundation for research in inventory modelling and has many assumptions. Demand determinism and instantaneous replenishment assumptions make the model to have simple solution while providing powerful insight inventory management. The corollary is not to order until just when the system runs out of stock, known as the zero inventory property of the optimal policy (Zipkin, 2000). The fact that these two assumptions are not always realistic led to the introduction of two other concepts: the reorder point and safety stock. It is important to know that even when there is Lead Time (LT), an optimal system still needs no safety stock provided both demand and lead time are deterministic, although there may now be a positive quantity reorder point. The presence of variation in demand rate and/or the LT time is what necessitates the need for safety stock, which quantity is generally incorporated into the reorder point. Some of the important initial extension of include Taft's (1918) Economic Production Quantity (EPQ), and Wagner and Whitin's (1958) Dynamic Economic Lot (DEL) models. The DEL model is particularly important because it is the most applicable model when the demand pattern can be approximated by a fairly deterministic but time varying demand as seen in cyclic products. It makes use of mathematical programming model and solution.

Research in inventory modelling accelerated in the 1950s. One of the seminal works in this regards is Arrow, Harris and Marschack (AHM, 1951), in which a number of postulations were made. It argued that the optimal policy in an inventory management system is an (s,S) policy, where the small s is the RoP, and the big S is the order-up-to level, and that there are unique values that optimise these variables. In particular, when demand and lead-time are deterministic, the optimal policy is to order the EOQ at a regular interval, while setting the reorder point to zero. Dvoretsky, Keifer and Wolfowitz (1952) proceeded from here to show that the (s,S) policy may not always be optimal in every system, but also showed the conditions under which the policy is optimal. This generated a flurry of works in the determination of the (s,S) values, the estimate of which has been noted to be difficult. Igelhart (1963) derived the optimal inventory (s,S) policy for the DEL problem with infinite horizon. Motivated by AHW, researchers worked on finding the optimal values for the RoP and the order-up-to level numerically. Igelhart (1963) and Veinott and Wagner (1965) may be considered trailblazers in this area and against which many subsequent solutions were benchmarked. The two main approaches adopted are the dynamic programming approach and the stationary distribution approach. In general, finding s and S involves iterative procedures that seek to progressively find better bounds until the solution is deemed to have converged. Generally, two measures are used to evaluate the performance of the system: the long run average cost per time or the total discounted cost. Renewal theorem was also used to find the limiting distribution of the LT demand, and hence, the values of (s,S) that minimise these costs as the best for the inventory system. Deriving the optimal values is not an easy problem to solve, and the behaviour of the cost function surface which is not unimodal has been said to be the main problem (Sahin, 1982).

Determining the joint distribution of the lead time and the demand during the lead time is not trivial. Problems include the computational requirement from data and the computational demand from the computing system. Specify the distributions of both the lead time and the demand properly and accurately and to the estimate of all their parameters are necessary. Bellman, Glicksberg and Gross, (1955) noted that it is often true that the determination of the optimal policy may depend on far less information than would seem to be necessary and the elegance of a good model may help to obtain good approximate to solutions. Many algorithms have, therefore, exploited some inherent structures of the model or the distribution to find good solutions. In fact, many scholars have tried to simplify this by using some fairly well behaved functions to approximate these distributions. For instance, using the normal distribution approximation implies that there may be no need to use moments of the distribution higher than two. Consequently, simplifications by using normal approximation (Maccormick, 1977, Ehrhardt, 1979), power function approximation (Ehrhardt, 1979), Poisson, exponential, beta, and gamma (Sivaslian 1971, Archibald and Silver, 1978) amongst others are reported. The problem reported with the approximation is that the resulting joint distribution may sometime be different from the individual distributions result in distribution function not seen or known (Danish 1972). This has made the problem of finding algorithms to compute (s,S) an ongoing area of research, but has not detracted from both the usefulness and importance of its logic.

4.2 Policies other than (s,S) and their similarities

Some author use the RoP model to refer to both the model of the upper and the lower limits. This is a misnomer as the lower limit may be easily estimated without recourse to the upper limit, while the upper limit can hardly be estimated without the lower limit. The lower limit was set because of the non-deterministic nature of demand and/or LT. In this paper, we will stick to the lower return limit as the true definition of RoP.

Many other policies have been adopted for managing the inventory system other than the (s,S) but these are interestingly similar to (s,S) or may be considered its instances. For instance, the (r,Q) policy is similar to the (s,S) in that r is simply the RoP. Q, however, indicates that a constant quantity be procured, meaning the masximum stock level may not always be the order-up-to level. This policy reverts precisely to the (s,S) when stock goes down at unit rate. There (S-1,S) policy implies that replenishment should be done once an item leaves the inventory, and this simply implies an (s,S) system in which the RoP is (S-1). Drawing this parallel is important because it creates a mechanism for comparing the policy of protecting material flow in many environments operating diverse OM paradigms. For instance, LM ideally operates the (S-1,S) policy, which may be impractical in most real life scenarios, hence, the Kanban size is a compromise between (s,S) and (S-1,S). Another possible decision is to have no RoP, but this is like a perpetual motion machine and is impossible as long as we cannot have instantaneous replenishment in the presence of variabilities.

4.3 Rehashing the RoP, its meaning and components

Most industry practitioners are said to adopt a normal approximation for the distribution of demand and that of the lead time (Danish, 1972). This makes the distribution of demand during lead time to also be assumed normal. While this approximation is not always realistic, it has been shown that there are a number of conditions under which this approximation may be suitable, for instance, when the demand is Poisson and the lead time is approximately normal. Moreover, when the lead time can be assumed to be constant and the expected demand is large enough, this approximation may be tenable. When such approximation holds, it yields a form that makes the calculation of RoP simpler because it can be assumed to be made up two main components: the expected demand during the lead time and the safety stock upon which the cycle stock rests (Figure 2). The total stock may then be conceived to consist of the cycle stock and safety stock, with the reorder point being somewhere between these two values.

While the individual normality assumption may not always be true, the decomposition logic of the RoP remains a useful simplifying logic, particularly because although the manager may not have much control over demand variability, it may be possible to exercise reasonable control over supply lead time variability to make it close to constant. This perspective makes it possible to optimise the RoP and the lot size separately, using different techniques or models, as opposed to the numerical approach where they both have to be determined simultaneously. This is an interesting simplification that makes it possible to use a lot more array of optimisation models that would have otherwise been difficult to apply, and is a very useful abstraction. The overarching logic of the RoP model is that the cycle stock balances ordering cost against holding cost or constrained by a minimum order requirement, while the safety stock balances holding cost against shortage cost or may be estimated given an allowable shortage limit. The cycle stock determines the replenishment interval while the safety stock is determined by the lead time and these two time lengths do not have to be equal, except when so designed or constrained. How the model is implemented is another issue which should be separate from the logic itself as this becomes subject to issues of probability distribution assumptions and how replenishment orders are placed, but as mentioned, the model of RoP can be readily decoupled from the order-up-to model, which is where the entire logic is vilified.

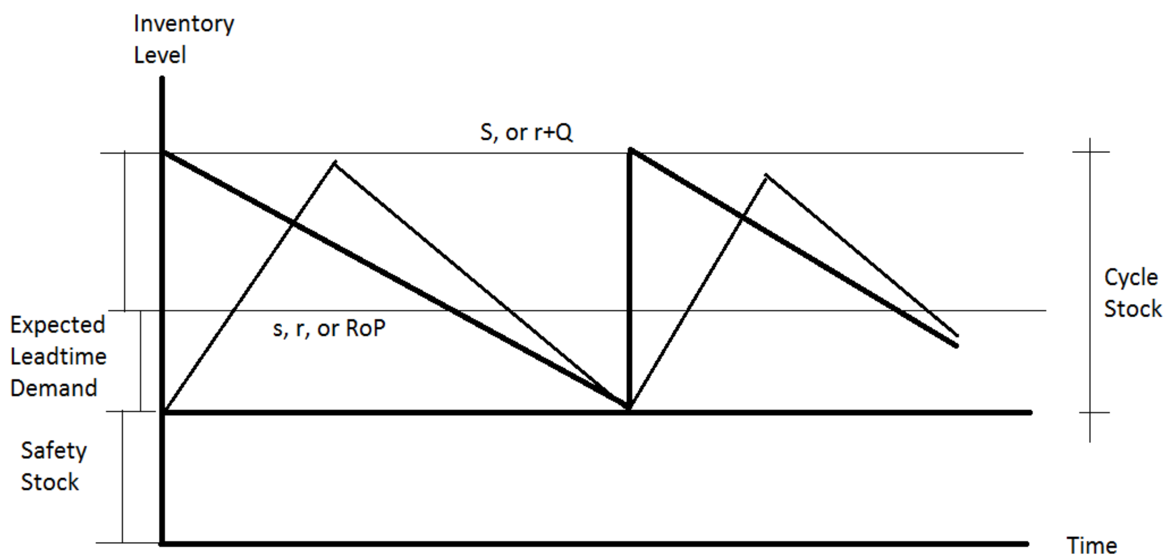

Figure 2. Representation of the components of the total stock

4.4 The RoP problem and the MRP crusades

Sequel to the introduction of the Order Quantity and RoP models, many companies reported significant savings in their inventory management cost (Whitin, 1954). It was stated that the classic RoP was well suited for the immediate post World War II (WWII) era in which it evolved as the economy was booming with stable outlook in America, which was the manufacturing centre of the world the, and the main focus was cost reduction based on economy of scale and division of labour (Mabert, 2007; Rondeau and Litteral, 2001; Jacobs and Weston Jr, 2007). This, together with the introduction of the mainframe computer that enables the automation of the lot sizing and reorder system, leads to an industry wide spread of the logic until the movement for MRP began to gain momentum. Part of the reason for the growth of the MRP may be adduced to the perceived weaknesses of the RoP model that the crusaders for the MRP actually propagated.

The leading discontent with the EOQ/RoP system was the inability to consider dependency in the demand structure of many manufactured products, hence, managing every input item as an island. Other concerns include its entire backward looking nature, not anticipatory of future changes in order patterns, and not providing sufficient mechanism for meaningful feedback for shop floor control. This led to the advancement of the MRP system, led by the trio of Joseph Orlicky, Oliver Wight and George Plossl, a movement termed the MRP crusade, whose goal was to get MRP thinking into the industry. They likened the MRP revolution to the heliocentric proposition of Nicolas Copernicus, against the popularly held geocentric belief of the time, to which they likened the RoP. Another crusade sought to get MRP thinking into the academia, and this also was achieved by the movement (Mabert, 2007). The new area of study in manufacturing strategy led by the Harvard team that puts focus on market stratification and manufacturing strategy and the role of APICS were believed to have contributed to the massive penetration of MRP (Chase and Jacobs, 2011; Rondeau and Litteral, 2001). In another part of the world, however, Japan that has been badly damaged after the WWII was also trying to bootstrap itself. The government prioritised the manufacturing industry, particularly auto-manufacturing, for massive support, and an ideology that will challenge the MRP dominance was being birthed.

4.5 The shortcomings of MRP and the growth of Lean Manufacturing

MRP, on the back of massive economic boom, had powered the US economy for some time, but the Achilles' heels were beginning to show. It was said that MRP does not put customers at the centre of all its operations, overproduces, and does not pay enough attention to quality. These issues were the main focus of the Lean Manufacturing movement as an emerging management philosophy. Operationally, the exploitation of the dependent demand structure the Bill of Material (BOM) had been the strength of MRP over the classic EOQ/RoP, such that the end product and its components are easily planned through the netting and explosion logic (Ptak and Smith, 2011). This same structure became the albatross of the MRP as the dependence creates the phenomenon of nervousness in an MRP system, such that a small change at the highest level of the BOM can create a massive ripple down the components, leading to massive delays, order cancellations and expediting. All these lead to unnecessary cost and compromises the production system (Jacobs et al, 2011). MRP also depends heavily on forecasting, and became more difficult in a world with short product lifespan and little customer tolerance for delay and failure (Kumar and Meade, 2002). There were usually many instances of production stoppage for lack of the needed materials while there are large stock of materials not needed (Ptak and Smith, 2019). This is in addition to the complexity of the implementation of the MRP system as it is usually deployed within the ERP system that is expensive and complex to implement and manage. MRP was also considered inflexible for implementation in the rather agile industry that evolved with the global inter-connectedness birthed by the internet and globalisation era and is not implementable by businesses that need to produce in frequent small volumes (Mabert, 2007). All these sped up the popularity of LM and its industry penetration.

The typical MRP system is designed to operate with zero buffer using lot-for-lot replenishment, but this is hardly so (Lagodimos and Anderson, 1993) as it was predicated on the assumption of deterministic demand, reliable lead time and lack of product defect. Graves (1988) actually called it the MRP folklore. This necessitates having safety stock, and implicitly the RoP into MRP. In addition, there was debate about where to place stock in the product hierarchical and how much. Moreover, most MRP inventory analysis is said to be based on simulation modelling and lack rigorous theoretical underpinning (Lagodimos and Anderson, 1993). A few analytical approach for safety stock modelling developed also uses probability distributions for both the lead time and demand (Grubbstrom and Tang, 1999), and most foundational models for safety stock calculation in MRP have the basic form of that in the RoP (see Wacker, 1985).

Some other MRP issues in addition to demand and supply variations include the considerations of what exactly a planning horizon should be. Theoretically, it should be at least the longest path in the BOM structure, but for optimisation purposes, the model may require much more. Dellaert and Jeunet (2005) discussed the horizon effect, the rolling schedule problem, and the process of horizon extension. Changing the horizon may mean changing what is optimal and this particularly has great impact on the first lot size. Bater and Peterson (1979) remarked that longer horizons improve performance but with decreasing returns when demand is stochastic, while Satdtler (2000) proposed the window look beyond model be applied to the Wagner-Whiten solution, but discouraged after the horizon has been supposedly found. Also, most simulation solutions are difficult to extend to multi-level products with commonality of components and lead time effects worsen the horizon problem in MRP lot sizing (Dellaert and Jeunet 2005). Inderfurth (2007) stated that not only must safety stock in MRP respond to yield and scrap issues, it also needs to make allowance for the interaction between the different factors of uncertainty. MRP implementation usually implies planning for the horizon, implementing the first time bucket and rescheduling for some subsequent periods. Yano and Carlson (1987) noted that frequent rescheduling may obfuscate the effects or degrade the performance of increased safety stock.

Schmidt and Nahmias (1985) observed that the optimal Policy for a two-stage assembly system under random demand is actually identical to the (s,S) policy, and even more so when positive setup cost is considered (Lambrecht et al, 1984).

Strategies proposed to combat the challenges of MRP implementation include use of safety stock, safety lead time, over-planning, hedging, freezing of horizon and incorporating change cost (Murthy and Ma, 1991; Blackburn et al, 1986). Of all these, more effort has gone into comparing the effectiveness of safety LT and safety stock. Jacobs and Whyback (1992) advocated use of safety LT to combat LT uncertainty and safety stock for demand uncertainty, but this seems unsupported. For instance, Buzacott and Shantikumar (1994) wrote that safety lead time is desirable only when it is possible to make accurate forecast of delivery LT and that safety stock is more robust in coping with fluctuations in LT and demand. Moliner (2007) seems not to have noticed a difference in implications on performance by using safety lead time while Alves et al (2004) also advocates using safety stock. The implication of this is that using time in response to a time variation and stock in response to demand variation should rather use a more system approach like the resource slack theory. The weaknesses of MRP also propelled the evolution of planning paradigms like Theory of Constraints (TOC), CONWIP, Demand Flow Technology, COBACABANA, and POLCA amongst others. Bagni et al (2018) did a review of P&C methodologies available between 1999 and 2018 and puts the number at thirteen, stating that seven of the new ones are similar to Kanban, POLCA or CONWIP. For the purpose of this review only LM, TOC and DDMRP will be considered in addition to MRP and RoP.

LM has both industry and academic leads. The industry leg was led by Taiichi Ohno, supported by others, while the academic support came from some American scholars (Deming and Juran) that found great acceptance in Japan, who were pushing the quality ideology that eventually became the complementary ideology of lean, and which would revolutionise the management of production system and propelled Japan to a level challenging the American dominance. Ohno had studied the moving line of Henry Ford and believes it uses too much inventory, such that the solution to every process error seems to be found in carrying more inventory, thereby elevating dead stock and became part of the albatross of MRP. While Deming and Juran seem to have made some contribution to improving the quality and productivity of the US manufacturing industries, much of their ideologies seem not yet acceptable in the US, unlike Japan. Ohno's concept of eliminating hindrances to the rapid flow of materials, and the total quality management concepts of the quality gurus, formed, respectively, the complementary go and stop columns on which LM principles were built. The third inherent principle of LM is people involvement (Schoenberger, 2007), and this is sometime argued to have been rooted in the culture of the eastern countries.

LM is quite difficult to define because it is a mix of many practices that have been learnt, adapted, and perfected from different places and is implemented within the context of the local imperatives and/or constraints. This made it quite difficult for many western companies and countries to understand what it was about initially, other than that the effect of LM on the competitiveness of Japan was quite obvious to them. For instance, it has been said that the idea of moving line by Ford was adopted but modified for small batch manufacturing. This was motivated by the economic downturn in Japan post WWII era and lack of finance to construct large-scale plant as obtained in the west, but turned out to be a blessing. Line-stopping idea came from the experience of Ohno in the textile industry, but was adapted for auto-manufacturing. The idea of the takt-time was said to have come from the German plane manufacturing system, but adapted for lean. The Kanban idea was said to have originated from a visit to some auto manufacturers in the US, but adapted for lean. Work standardisation and measurement principles were rooted in scientific management, and the keiretsu principles from social structure Japan. Consequently LM was said to have organically grown over decades, but became an integrated entity dubbed as the Japanese (or Toyota) way of management (Holweg, 2006). The role of the International Motor Vehicle Program (IMVP) based in the Massachusetts Institute of Technology (MIT) in disseminating the broad philosophical concepts of LM has been reported (ibid).

The original LM frowns at excessive manufacturing and produces one when one is used up, but this is unrealistic, hence, there is always some stock and batching. Stocks are maintained in the First-In-First-Out (FIFO) queues and the supermarkets maintained in the system, but the question is how the quantity of these stocks are determined. In addition, LM does not really manufacture in batches of single items, but the quantity that a Kanban contains, hence another question is what guides the size of this Kanban. While the number of Kanban cards needed in a LM environment is readily determined once the total stock level and the Kanban size have been determined, the decisions about these two prior parameters are not trivial, and much has not been reported about how optimal they are. Since LM does not simply make all items ready for the customer, the Kanban becomes its heart to limit inventory, regulate flow, determine batch sizes and prioritise jobs.

It is important to note that one reason why LM is able to make to demand is because there are pockets of slack capacities within the system. Also, rather than deploy such capacity in a lump sum, LM works with manufacturing cells, thereby having many cells that can be opened or shut down on demand, hence the idea that LM utilises plants within a plant structure (Chase and Jacobs, 2011). This has a number of implications. Firstly is that excess capacity compensates for reduced inventory, otherwise, variability makes LM implementation infeasible. Secondly, having small cellular system doing the same thing may imply replicated setup for the same type of production, hence, without aggressive setup effort and cost reduction, a lean system may not favourably compete with a large plant, especially in an environment with large demand. In addition, LM demands very strict quality management everywhere in the system, otherwise the line always stops. Quality comes at a cost, hence, the cost of strict quality control and assurance including all the poka-yoke and autonomation mechanisms and devices become relevant.

Wang and Wang (1990) noted that most people have an erroneous assumption that lean has the capability to make products without keeping stock and with a lot size of one. Such a system is attainable only when there is no economy of scale due to lot sizing and when there is no inherent variability in the demand and supply processes, which is practically impossible in any operational system. There is always, therefore, the need to still lot size in lean environment. The next natural questions are what would be the practical quantity in an LM batch, what would be the total stock level to maintain, and is there an appropriate safety stock level to maintain. Just as in MRP, the only time this question is circumvented is when the practitioners or analysts have to make assumption that there is no setup cost or there is no variability in the system. The Kanban question immediately reverts to whether one is applying an (s,S), (r,Q) or (S-1,S) policy, and hence, the application of the same mathematical techniques used in RoP optimisation.

Monden (2012) discussed the basic operations of Kanban which may be operated as a dual-card or a single-card system. These cards are called production and withdrawal/transport Kanbans respectively. The production system should consider the relative implications of machine setup, stock holding and shortages, while the transport/withdrawal should balance material handling/transportation against material availability costs and implications. Without any loss of generality, our discussion about the management of LM inventory systems will focus on the withdrawal system, which directly extends to the production system. There are two main approaches to withdrawal: instantaneous or periodic review of inventory position. When reviewed periodically, the material handler may cycle in regular intervals (pitch) but withdraw variable quantity or withdraw fixed quantity and non-constant intervals. Sometimes, Kanbans may be accumulated until a certain number is attained before dispatch. It may be apparent that this draws similarities to the classic reorder principles, and the basic formula used is the demand during the replenishment lead time plus the safety stock.

Arktyuk and Erhun (1999) provides a succinct summary of research into models for optimising LM Kanban parameters. Similar to the RoP system, stochastic, mathematical programming and simulation models have been used to seek optimal stock levels for LM systems as well. Some models consider balancing fixed costs (setup, order or transport) and variable costs (holding and shortage) while some ignored these costs altogether and simply seek to measure some other service parameters, like order fill rate, against the level of stock maintained. In cases where an analytical solution is desired, some form of probability distribution is needed for the demand and/or lead time, and Poisson/exponential combinations are also common (e.g. Di Mascolo, Frein and Dallery, 1996; Mausui, 2002; Matzka, Di Mascolo and Furmans, 2012). Mathematical programming formulations are similar to those of the DEL model used in RoP and MRP systems, apart from most such models having their Kanban size specified ahead, hence, the number of Kanbans is usually the important variable (Bitran and Chang 1987, Wang and Wang, 1990; Widyadana, Wee and Chang 2010; Yang, Zhang and Jiang, 2010; Aghajani et al 2016). Simulation is particularly adopted in cases where less restrictive distributions or when some more adaptive systems are desirable (e.g. Gupta, Al-Turki and Perri, 1999; Tardif and Maaseidvaag 2001; Köchel and Nieländer, 2010), but also suffers from the problem of generalisation as in MRP. There have been other techniques like regression analysis (Chapman, 1989; Takashi and Nakamura, 2002), but these are few and far between.

The main difference between LM optimisation models and those of RoP is that the aggregate inventory, which is essentially cycle plus safety stock, is divided into smaller Kanban sizes, with most researchers not stating how they decided what the Kanban size should be, but simply sought to optimise the number of Kanban cards needed. Depending on the implementation of the Kanban system, the Kanban is simply RoP system with Kanban size or its multiple as the lower limit. When optimisation is applied to multiple levels of inventory is like optimising a multi-echelon inventory system.

LM has been said to be with its own weaknesses as well, and this has led to the proposition of the DDMRP, which shall be discussed later. Overall, the principles of LM has been summarised by Rother and Shook (1999) as allowing materials to flow through the production system whenever possible, and when this is impossible, to pull in response to demand from the strategic storage points using the Kanban, which may be seen as an adapted (s,S) system. We discuss the TOC that also seeks to address similar MRP problem like LM.

4.6 TOC Philosophies

TOC is believed to have its roots in the Optimised Production Technology (OPT), an application developed by Dr Goldratt in response to a client's scheduling need. The solution metamorphosed from an application into a philosophy over five distinct eras. The first is the OPT era, when the idea was packaged as a secret application, which eventually led to a court case; the second was the goal era, during which the Drum-Buffer-Rope (DBR) technique and the five focusing steps for continuous improvement were introduced; the third was the haystack syndrome era, during which issues related to performance measurement system in organisations was addressed, with particular focus on the cost accounting system that was branded as a public enemy. This led to the development of a process focussed system called throughput accounting; the fourth was the era of "It is not luck" where organisational success was said not to be due to mere luck, but based on tested principles of thinking for continuous improvement. This includes concepts such as the Current Reality Tree CRT), Future Reality Tree (CRT), Transition Tree (TT), Prerequisite Tree (PT) Evaporating Cloud (EC), and the Core Conflict Cloud (CCC); the fifth was the critical chain era, during which the ideology of TOC got integrated into project management as the Critical Chain Project Management (CCPM) (Watson, Blackstone and Gardiner, 2007).

To many, the DBR is considered as the heart of TOC as it helps in the management of inventory and job release. It regulates the rate of flow of material through the system and needs to be protected by buffer. There are three main types of buffers: the constraint buffer protecting the drum; the assembly buffer to protect the assembly points after to the constraint; and the market buffer, to manage variability from demand sources. It is from these buffers that jobs are released into the system based on their level, and the communication system from the buffer to the input release mechanism is the rope (Rahman, 1998). TOC drastically reduces the necessary control points in a system and advocates the use of colour codes (red, yellow and green) to signify the buffer positions, and this guides the reorder point. It can be imagined that this is like the three layers of stock in the classic RoP model, where there is the cycle stock, demand during the lead time and safety stock layers. There aren't, however, clearly stated models to determine these three levels. The practitioner is advised to simply fix the maximum level, divide it to three and iterate with the processes by reducing the maximum level if the stock is mostly in the green zone and increasing it if it is mostly in the red zone. The question remains what the definition of mostly in a zone is and how one determines decision's optimality. In addition, this model simply ignores the cost of keeping or running out of stock.

TOC limits the quantity of stocks maintained in the system because it has pockets of extra capacity all over the system. TOC uses this capacity in many ways. Firstly, in cases where there is quality problem with any item produced, this is used to quickly rework the item or make a new one altogether. In addition, TOC must manage not only the aggregate demand, but in cases where the same resource is used to make more than an item, TOC must manage the mix. One use of extra capacity in TOC is to break down the batch sizes, especially at the Non-Bottleneck (NBN) resources. In addition, because the production and transfer batches are usually not the same, the small batches made by the NBN need to be frequently transferred to where they are needed (Chase and Jacobs, 2011). It is simply assumed that these multiple setup and transfer costs are negligible. If this is not so, it has to be traded off against the stock reduction to know how profitable the system will be (Adetunji and Yadavalli, 2012).

Researchers have attempted to model buffering TOC as well. Radolvisky (1998) is about the classic work that proffers analytical solution to optimising the maximum buffer size for a TOC. He assumed a Poisson arrival and an exponentially distributed service time to determine the minimum stock level to provide a given level of protection of flow, which mirrors the RoP model. Zhao and Hou (2014) extended this model to a case of general distribution of arrival and service patterns but could only solve the model numerically. Ye and Han (2008) used reliability analysis to study buffer sizing in a TOC network structure. Adetunji et al (2011) studied the impact of utilisation on the buffer level in a TOC environment. AlGhamdi and Weheba (2016) used the Taguchi's quadratic loss function approach. They balanced the loss of machine idleness due to flow synchronisation against the delay of materials queuing ahead of the drum, and derived an EOQ-like model.

3.7 DDMRP Thinking and motivation

DDMRP has been said to become necessary as a result of the gaps inherent in each of the OM philosophies in themselves. While the challenges of MRP have been highlighted, advocates of DDMRP have also discussed the numerous challenges of LM. In particular, it was stated that LM lacks the capacity for the integrated planning of MRP which exploits the demand dependency through its BOM structure and material explosion. This makes LM to become more reactive and restricted control of materials to what has actually happened and not what is anticipated to happen. Moreover, it brings the control structure back to where the advocates of MRP stated was the weakness of the traditional EOQ/RoP system in that each stock point in the organisation is managed on its own without considerations for what may happen in the other parts of the system on which it depends or that depend on it. This makes LM more short-sighted than MRP. Consequently, while MRP may be overly complicated in certain instances, LM itself may also become oversimplified, and both situations are not good enough (Ptak and Smith, 2011).

The philosophy of DDMRP is built around three principles of Position, Protect and Pull (PPP). This means appropriate positions for material buffer must be identified, which should help to protect the flow of materials through appropriate buffer profiles and levels, and through which material is pulled to enable flow using Demand Driven Planning. They proposed the use of the Decoupled Lead Time (DLT) as a compromise between component LT, which may be too short for stability of the system, and the aggregate LT, which may be too long for responsiveness. The material explosion process then utilises the DLTs for decoupled explosion based on a netting formula that is a modification of the traditional projected balance calculation of MRP. This formula uses concepts called the Order Spike Horizon (OSH) and Qualifying Spikes (QS) in the netting calculations to manage variability in demand, and hence, determine when order is placed and how much quantity it is. Demand, zone and LT adjustment factors were proposed to guide the timing and quantity of replenishment. To determine the buffer levels, they also proposed the three colour coding system like TOC, but chose the factors depending on the material type (Ptak and Smith, 2019).

There have not really been works using analytical models for optimising buffers from the DDMRP perspective, probably because it is a more recent approach. Most works publish are simulation based, and they suffer from the same generalisability issues mentioned earlier. Lee and Rim (2019) provides an analytical solution that uses the coefficient of variation of demand and LT instead of their variances and tested the model using simulation, but the validity and generalisability of the model is not proven either. It is, however, reasonable to assume that the analytical models proposed for TOC systems should be equally applicable to the to the DDMRP system. Most papers seem to be comparing the performance of DDMRP to the other approaches.

3.8 Comparative performance analysis of philosophies

Many authors have compared the operational performance of the OM philosophies in order to ascertain which one seems best. This is reviewed briefly here, but the results will summarised using the push/pull context for better insight. Many authors have written in this regards, but only those involving the philosophies of interest and considered relevant are discussed. Jacobs and Whyback (1992) was one of the first, and concluded that MRP performs better than RoP, except when a deliberately perverse situation is created. Puche-Regaliza (2021) concluded that TOC works better than LM in situations where there is higher volatility, and this seems to agree with the positions of Chakravorty and Atwater (1996) and Watson and Patti (2008), who believed this is due to the strategic inventory placement of TOC. Steel et al (2005) stated that TOC definitely performed better than MRP and believed it is mainly due to TOC's lot splitting and ordering characteristics. They stated that MRP matches production to requirements better, but crudely schedules the MPS, while TOC better controls the bottleneck operation flow time, but moves rigid common lot sizes all the way from components through to finished goods inventory and loses some control of flow time in the non-constraint operations. Chakravorty and Atwater (1996), however, believe LM works better when variability is low, and Gupta and Snyder (2009) also seem to agree with this in a repetitive production environment, stating that it might be better for the manager to apply TOC until variability in the system is brought under control. Krajewski et al (1987) also stated that LM is better than MRP, but there are scenarios where both seem to perform poorly. He advocated the identification of appropriate factors to guide the use of each approach. Thürer et al (2020) studied MRP, Kanban, TOC and DMRP systems together and concluded that MRP seems to have generally performed most poorly, attributing this to its enforcement of predetermined start dates. They believe the Kanban and DDMRP system fared better in cases where there are no bottlenecks, while TOC and DDMRP worked well with bottlenecks.

MRP utilises the Input-Output (I-O) chart for shop floor control, and it is apparent from most reports that this seems not to work well when compared to LM's Kanban and TOC's DBR. A probable reason is the data intensity upon which I-O depends and the complexity of ERP system that implements it, while Kanban is very simple to implement, but may also become quite weak when the number of Kanban cards moving along the production system increase significantly and when the variability in the system becomes quite high, in which case TOC may seem to fare better than LM. On the other hand, MRP has the most forward looking perspective with highly connected structure for all stock items due to its BOM and dependent demand structure. Consequently, the question should not be about which system performs best, but how each should be used where it works best. This is probably where the move towards DDMRP is a welcome development, however, a lot has not been studied about this technique, especially through the use of analytical model.

MRP is usually considered a push production system while LM and TOC are basically considered to be pull systems, which Michel et al (2006) has pointed out as a misnomer because all these approaches are planning paradigms and the reality is that most practical systems are a mixture of both. Push system are usually associated with high utilisation, but also high stock level, and pull systems are believed to operate with low inventory level. Consequently, poor performance in pull systems are typically assumed to be due to failure of implementation, while for a push system, it is expected (Bonney et al, 1999). The question, in fact, is if these tacit stereotypes are valid? The first issue worth examining is what constitutes a push system. Many authors have pointed out that it actually depends how the systems are characterised. Some of the logics used include how demand was observed (forecast vs order), how replenishment order was triggered (request vs plan), the planning paradigm (anticipation or actual occurrence), direction of logic of planning information (same as material flow or opposite), estimation technique (look ahead vs look back), and many others. It was stated that using these approaches, a system classified as push by a logic may appear like a pull relative to another (Michel et al, 2006; Bonney et al, 1999). Also, it has been argued that the superiority of performance reported by many authors may depend on the measures used. For instance, Masuchun et al (2004) reported in their research that push outperformed pull in all instances and scenarios in terms of customer service and throughput performances, while pull outperforms push in terms of level of stock maintained. In some instances where pull was said to perform better based on stock level, it was argued that many such models assumed there is no cost for multiple setups or ordering through which the lower stock level was attained. Spearman and Zazanis (1992) argued that while there has been evidence for better result of pull over push, evidence also suggests that systems including both have performed better. Some researchers have stated that the focus should rather be on alignment of processes to the operational environment and market rather than focusing on the push/pull dichotomy and some possible models of alignment may include the product-process matix Hayes and Wheelright (1979) product-supply chain matrix and Fisher (1997).

5. Take-away from the philosophies

The Kanban logic of LM is a powerful technique to control and connect workstations. A more interesting benefit it has is the capacity to make obvious the velocity of materials, especially when utilised with LM's pitch. Unlike the classic (s,S) system that may not sense changes in material velocity quickly enough, breaking the lump of material between the lower and upper limits into smaller Kanban units makes it possible for the manager to be able to sense the velocity of work on the shop floor and take proactive measures to anticipate the possibility of shortages or material build up.

The DBR principle of TOC helps to massively reduce the number of control points in any operational system and makes it possible to control an entire system by focusing on a few points. This is a powerful technique that helps the manager to focus on the important issues and operations.

The material dependency structure of the MRP provides a powerful approach to anticipate the implications of an action in any point of the production system on all other areas. This power is particularly exploited in the pegging and where used reports, and in the various modifications of the BOM, like super bill and matrix bill, that helps to manage products in more logical groups. This helps to anticipate the impact of actions and decisions made about material volumes and mixes from a system wide perspective.

The idea of the decoupled lead time in the DDMRP is a powerful logic that helps the manager to choose their battle points to their advantage. It seems simple, but may later prove to be one of the most revolutionary thoughts of operations management in the recent era. It also creates a midway between managing nervousness of MRP in a manner better than time freezing, and simultaneously managing aggregate stocks and customer facing response time.
The Reorder point logic is so basic but indispensable in managing variations in demand and lead time, but its relevance is seen only when two separations are made: the logic from the implementation; and the decisions about the upper from that about the lower limits.

5.1 Important summary of discussion

The foregoing may be summarised in the following points:

1. The classic RoP model may have some problems, but this is not necessarily due to its logic, but the implementation assumptions (e.g. the probability distribution), hence, discarding it is throwing away the baby with the water. The assumption of some simplified distribution is due mainly to the tractability of the solution, and there is always the need to strike a compromise between having a perfect representation and having a tractable model as obtains in the representation of every system.
2. There has not been any analytical solution proposed by any of the OM philosophies that does not make use of the reorder point logic. Simulation has probably been used as an alternative to analytical models in order to use some more realistic distributions, but in many such instances, the results may not be readily generalizable. While some OM philosophies might have left the judgment of deciding the buffer zone parameters and levels to the opinions and experience of the shop floor managers, the quality of their solution are difficult to analyse of prove.
3. In reality, there is no zero inventory system as is usually touted. Every manufacturing planning system that has non-deterministic parameters like demand and/or LT uses some form of RoP mechanism to control their inventory level. Such mechanism may be a two-parameter structure having lower and upper return limits like (s,S), (r,Q) or (S-1,S), as also adapted in lean, or it may be explicitly broken into a three-zone structure of safety stock, lead time demand, and cycle stock, as seen in TOC and DDMRP. Also, by adding the LT demand and the safety stock together, the three-zone structure becomes a two-parameter structure.
4. No production paradigm, push or pull is better than the other in all ways, but any superior result depends on the context of the production. Once variability increases, the only way out is to have sufficient slack in the system, either in the form of capacity, inventory, or some of combination of these, and which option is better depends on the goal of the organisation. Moreover, both inventory, capacity and time are reasonably substitutable and reducible to some monetary equivalent. Good operations management practice demands determining which of the slacks to carry, to what extent, and in what area in the system.
5. Determining optimum inventory level without considering the nature and level of capacity in a production system may not make sense. Both extra stock and extra capacity are meant to protect material flow and both cost the organisation, hence, two organisations having similar market structure and inventory profiles but different capacity profiles cannot guarantee the same response and service level. Also, modelling optimal stock level without clearly relating it to the buffer capacity structure or the organisation’s control over customer facing LT may not paint the complete picture.
6. When an organisation decides to reduce their stock level and still maintain the level of service demanded, there is also usually the need for stricter quality control of items. This may involve using the buffer capacity to rework or remake the items, and this may sometime also imply extra setup or changeover. In a number of models evaluating performance of models based on the different philosophies, this is not taken into account and these costs of more stringent quality control or assurance are ignored. Such evaluation may not be realistic, even if it seems to have produced better theoretical results.
7. OM philosophies have more in common that their differences, hence the goal of research should rather shift to how they can work in complementary roles than their differences. It should be possible to transfer practices from one philosophy to another to have better operations than seeking to show how an approach works better than the other. Actually, no OM philosophy works best in any system, but in context. The practices of a philosophy can be adapted to another to make the system better. For instance, nothing stops TOC or DDMRP from using the Kanban principle to manage its buffer levels in place of the three zones approach.
8. While LM and TOC seem quite effective for the controlling aspect of P&C due to their transparency, responsiveness and simplicity, they do not seem ideal for planning purposes as they usually reduce the system to islands and only responds to signals from the upstream end of the supply chain rather than proactively anticipating them. MRP seems to be more anticipatory due to the dependent demand nature captured in its BOM structure and should provide some advantage for planning aspect.

5.2 Revisiting DDMRP

The DDMRM is progress in a good direction and is worth considering in some more details because it is a deliberate effort at integrating these philosophies and creating the logic to fill the gaps from such integration. Since our objective is to consider how material flow is protected in order to meet demands, the main discussion will be around the buffer management thinking and the identified gaps, thereby setting the direction for possible future research. The DDMRP model will be compared to the classic three zone model.

DDMRP classifies all parts as either replenished parts, replenish override parts, or min-max parts, with different management approaches. The first two are strategic in nature and managed using the three zone logic (and would be the main focus of our discussion), while the third is managed using the two parameter (s,S) logic. For the three zone approach, there is the green zone, the purpose of which is to meet the normal usage (similar to the cycle stock). There is the yellow zone which is to cover the expected demand during the lead time, and the red zone which is the safety stock (same as in the classic RoP model). To determine the values of these zones, the parameters needed are the average daily usage (ADU, i.e. the daily demand), the part's decoupled lead time, the type of item (purchased, manufactured, distributed or intermediate), and the variabilities of demand and lead time. The size of the green zone is determined as the product of the ADU, LT, and the lead time factor (LTF). When LT is long, the LTF should be small so that there would be small frequent orders to minimise inventory. For long LTs, LTF is suggested to be 20% to 40% of ADU, 41% to 60% for medium, and 61% to 100% for short. Demand and LT variability are also classified as high medium or low. The part type, lead time and variability values are used to create what is called buffer profiles. A full factorial design leads to 36 profiles. The buffer profiles are used for setting the rules for estimating and adjusting the values for each of the buffer zones.

The gaps identified in the current estimate procedure are presented next. Firstly, there was no clear definition of what long or short lead time is. The implementation team has to decide that. Also, how the demand factor ranges were fixed was not stated, and how to choose values within the zone is also left to the team. There was also no clear definition of high or low LT or demand variation, and it was not clear how to decide what the variation factors should be. In addition, in creating the buffer profiles, the demand and LT variabilities seemed to have been combined into a single variation factor, and it was not clear how this combination was done. For instance, if demand variation is low and lead timer variation is high, would that be interpreted as low, medium or high variation? This needs to be done before the factors can even be allocated. Another important question relates to the management of the red zone. The DLT is used for the lead time, but DLT is usually for aggregate items. How exactly should the safety stock for the aggregate items be determined? Should it be based on the relative cost of each component, relative proportion based on the BOM, or some other parameters or combinations? This even gets more complicated if the items in the aggregate are also used in some other end products. Should the impacts of these other items be considered in the safety stock determination? All these are left to the implementation team to simply choose the values, yet there is no theoretical model to guide such choice. Finding optimal values of a two parameter (s,S) is quite some task. How do we guarantee optimality when so many factors have no guiding model, yet the team must select all of them? There are some other issues not discussed yet, like managing order spikes for the dynamic adjustment of the buffer for which there are no analytical predicates for making these decisions. Understanding the factors that impact buffer positions and their relative importance still need more considerations. To establish DDMRP as a philosophy, all these issues and possibly others need to be addressed.

4.3 Suggested research agenda

There are a few areas on which research effort might be focussed in order to expedite the process of integration of these OM philosophies, and some of these are discussed. There should be efforts to develop techniques for the management of items that need to be managed together like those with integrated DLT. Such may include joint, dependent and/or multi-echelon items that would need to be controlled together, reordered together etc. Interestingly, there are such research trends in a number of areas of operations research, but an interesting area that seems not to have enjoyed much attention is the joint management of safety stocks. The selection of the best decoupling points in the system also needs further research. In particular, there are a number of factors that have been proposed to be necessary to consider in making this decision, but not much seems to have been considered in validating these factors and the relative importance of each of them and the dynamics of interrelationship among the factors. Also, decisions about a number of scaling factors for some proposed buffer management factors have been left to the practitioners with little guidance on the choice of their values. If such areas are to be managed dynamically as proposed, there would be the particular need to consider how Artificial Intelligent (AI) techniques may be applied to make them self-adjusting. This will also enhance their implementation on ERPs. The usefulness of an integrated pitch and Kanban design in the study of Kanban velocity was mentioned, but interestingly, there seems to be a dearth of study on even the design of pitch, not to talk of an integrated system. This would be important in creating responsive control mechanism that may be generalised to all OM systems. It was also mentioned that the RoP logic is a good idea that is pervasive in all OM ideologies and systems, but still has some implementation limitations, sometimes because of the assumptions made for tractability. We suggest that work on a full documentation of the resulting joint distribution from the pairwise combination of all possible multi-item and multi-level demand and LT distributions, and a mechanism for extracting all their relevant quantiles in the form of online functions and their inverses (or even tables) would be very valuable to advance the efforts of jointly managing safety stock and reordering of stocks for decoupling the various integrated systems based on the DLT. Such efforts may later advance to giving consideration to several shock events that may occur within the DLT considerations.

6. Conclusion

This research traces the evolution of the major philosophies of operations management from RoP to DDMRP using the hermeneutics approach for literature review. It clarifies that the RoP logic is still very relevant as the reliable model for the estimate of Safety Stock and LT demand in all inventory management environments. It discusses organisational slacks and their relevance for stability due to demand and process variations, and their interchangeability when buffering to support material flow in a production planning environment. It emphasises that research emphasis should be on integration of the philosophies rather than differences in performance that seems to be the current focus with the superiority being flaunted based on possibly selective performance measures. It, therefore, considers the DDMRP as good progress due to its integrated approach and identifies the current gaps in its implementation. It identifies key characteristics of the philosophies that may foster their integration and suggests which areas researchers may consider to provide theoretical bases for such integration.

References

Adetunji O., K. Adendorff and V.S.S. Yadavalli., Dynamic Buffering of a Capacity Constrained Resource via the Theory of Constraints. Proceedings of the 2011 International Conference on Industrial Engineering and Operations Management Kuala Lumpur, Malaysia, January 22 – 24, 2011.

Adetunji OAB. and VSS Yadavalli, An integrated utilization, scheduling and lot sizing algorithm for pull production, International Journal of Industrial Engineering, Theory, Application and Practice, 19(3), 171-180.2012.

Aghajani, M., A. Keramati, R.T. Moghadam and Mirjavadi S.S, A mathematical programming model for cellular manufacturing system controlled by kanban with rework consideration, Int J Adv Manuf Technol, 83:1377–1394. 2016

AlGhamdi FS, and G. S. Weheba. Economic Approach for Determining the Time Buffer in Synchronous Manufacturing, The Journal of Management and Engineering Integration, 9(2), 40 – 50. 2016

Alves T., V.H. Machado, V.C. Machado, Modelling MRP Systems under Uncertainty: Safety Stock versus Safety Time, IIE Annual Conference. Proceedings, Norcross, 1-6. 2004

Archibald B.C. and E.A. Silver Policies under continuous review and discrete compound poisson demand, Management Science, 24(9), 899 – 909.1978.

Arrow K.J., T. Harris and J. Marschak Optimal Inventory Policy, Econometrica, 19(3), 250-272, The Econometric Society. 1951

Bagni G., M. G. Filho, M. Thürer & M. Stevenson, Systematic review and discussion of production control systems that emerged between 1999 and 2018, Production Planning & Control, 32(7), 511-525, 2021

Bellman R., I. Glicksberg and O. Gross, On the Optimal Inventory Equation, Management Science, 2(1) 83-104, INFORMS, 1955.

Bitran, G.R. and L. Chang, A Mathematical Programming Approach to a Deterministic Kanban System, Management Science, 33(4), ProQuest pg. 427, 1987.

Blackburn, J.D., D.M. Kropp and R.A. Millen, A comparison of strategies to dampen nervousness in MRP Systems, Management Science, 32(4), pg. 413, 1986.

Boell, SK. and D Cecez-Kecmanovic , A Hermeneutic Approach for Conducting Literature Reviews and Literature Searches, Communications of the Association for Information Systems, 34(12), 257 – 286.2014.

Boell SK & D Cecez-Kecmanovic , Literature Reviews and the Hermeneutic Circle, Australian Academic & Research Libraries, 41(2), 129-144. 2010.

Bonney MC, Z Zhang, M.A. Head, C.C. Tien, and R.J. Barson, Are push and pull systems really so different?, Int. J. Production Economics 59, 53-64.1999.

Bryant M., J. Vistad, and L. Fredendall, In Lean Production: Semantics Matters, Journal of Business & Leadership: Research, Practice, and Teaching (2005-2012), 2(2) , 2006.

Buzacott, J.A. and J.G. Shanthikumar, Safety stock versus safety time in MRP controlled production, Management Science, 40(12), 1678 – 1689.1994.

Chakravorty, S.S. and B.J. Atwater, A comparative study of line design approaches for serial production systems, International Journal of Operations & Production Management, 16(6), 91-108, 1996.

Chapman S.N., Just-In-Time supplier inventory: an empirical implementation model, The international Journal of Production Research, 27(12), 1993-2007., 1989.

Chase R.B. and Jacobs F.R., Operations and Supply Chain Management, Mc-Graw Hill Irwin, 2011.

Coyle J.J., Bardi E.J., Langley Jr C.J., The management of Business Logistics: A supply Chain Approach, South Western Thomson Learning, 2003.

Dellaert N, J. Jeunet., An alternative to safety stock policies for multi-level rolling schedule MRP problems, European Journal of Operational Research, 163, 751–768, 2005.

Di Mascolo M., Y. Frein and Y. Dallery, An Analytical Method for Performance Evaluation of Kanban Controlled Production Systems, Operations Research, 44(1), Special Issue on New Directions in Operations Management, 50-64, INFORMS, 1996.

Dvoretzky A., J. Kiefer and J. Wolfowitz, The Inventory Problem: I. Case of Known Distributions of Demand, Econometrica, 20(2) 187-222, 1925.

Ehrhardt R., The power approximation for computing (s, S) inventory policies, Management Science, 25(8), Pg. 777 – 786, 1979.

Fisher M.L., What is the right Supply Chain for your product, Havard Business Review, March-April 1997

Grubbstrom R.W. and O. Tang, Further developments on safety stocks in an MRP system applying Laplace transforms and input—output analysis, Int. J. Production Economics, 60—61, 381—387, 1999.

Gupta M. and D. Snyder , Comparing TOC with MRP and JIT: a literature review, International Journal of Production Research, 47(13), 3705-3739, 2009.

Gupta S.M., Y.A.Y. Yousef, Al-Turki, and R.F. Perry, Flexible kanban system, International Journal of Operations & Production Management, 19(10), 1999, 1065-1093. 1999.

Harris, F.W., How Many Parts to Make at Once Factory, The Magazine of Management ,10, 135-136, 152, 1913.

Hayes R.H. and S.C. Wheelwright, Link Manufacturing Process and Product Life Cycles, Havard Busines Review, January 1997.

Hopp W.J., Supply Chain Science, Mc-Graw Hill, 2008.

Igelhart D., Dynamic Programming and Stationary Analysis in Inventory Problems, In Multi-stage Inventory Models and Techniques, chap. 1, H. Scarf, D. Guilford and M. Shelly (eds.). Stanford University Press, Stanford, Calif, 1963.

Inderfurth K., How to protect against demand and yield risks in MRP systems, Int. J. Production Economics, 121, 474–481, 2009.

Jacobs F.R., and F.C. 'Ted' Weston Jr, Enterprise resource planning (ERP)—A brief history, Journal of Operations Management, 25, 357–363, 2007.

Jacobs F.R. and D.C. Whyback , A comparison of RoP and MRP inventory logic, Decision Sciences, 23(2), ProQuest Central, 332 – 342, 1992.

Köchel P. and U. Nieländer , Kanban optimization by simulation and evolution, Production Planning & Control, 13(8), 725-734, 2002.

Kumar S., and D. Meade, Has MRP run its course? A review of contemporary developments in planning systems, Industrial Management and Data systems, 102(8), 453 – 462. 2002.

Lambrecht M.R., J.A. Muckstadt and R. Luyten, Protective stocks in multi-stage production systems, The International Journal of Production Research, 22(6), 1001-1025.1984.

Lagodimos A.G. & E. J. Anderson, Optimal positioning of safety stocks in MRP, The International Journal of Production Research, 31(8), 1797-1813.1993.

Lee K.J.,B.E. King, L.P. Ritzman, D.S. Wong, Kanban, MRP and shaping the future of manufacturing, Management Science, 33(1), ProQuest, 39-57, 1987.

Mabert V.A., Early road to material requirements planning, Journal of Operations Management, 25, 346–356, 2007.

Masuchun W., S. Davis and J. W. Patterson , Comparison of push and pull control strategies for supply network management in a make-to-stock environment, Int. J. Prod. Res., 42(20), 4401–4419, 2004.

Matzka J , M Di Mascolo, K Furmans , Buffer sizing of a Heijunka Kanban system, Journal of Intell Manuf, 23, 49–60, 2012.

Monden Y., Toyota Production System : an integrated approach to Just-In-Time, CRC Press, 2012.

Murthy D.N.P. and L. Ma., MRP with uncertainty: a review and some extensions. International Journal of Production Economics, 25, 51-64, 1991.

Puche-Regaliza J.C., B. Ponte, L. Costas, R. Pino, D. de la Fuente, The Behavior of Lean and the Theory of Constraints in the Wider Supply Chain: A Simulation-Based Comparative Study Delving Deeper into the Impact of Noise. In: De la Fuente D., Pino R., Ponte B., Rosillo R. (eds) Organizational Engineering in Industry 4.0. ICIEOM 2018. Lecture Notes in Management and Industrial Engineering, 149-159, 2021.

Ptak C. and C. Smith, Orlicky's Material Requirement Planning, 3rd Ed., Mc-Graw Hills, 2011.
Ptak C. and Smitc C., Demand Driven Materials Requirement Planning, Industrial Press Inc), 2019.
Pycraft M., Singh H., Philela K., Slack N., Chambers S., and Johnston R., Operations Management, 2nd Ed, Pearson, 2010.
Rahman S., Theory of constraints A review of the philosophy and its applications, International Journal of Operations & Production Management, 18(4), 336-355, 1998.
Rondeau, P. and Litteral, L.A., The evolution of manufacturing planning and control systems: From reorder point to enterprise resource planning, Scholarship and Professional Work – Business, 41, 2001.
Rother M., J. Shook, Learning to See: Value Stream Mapping to Add Value and Eliminate Muda, 1st Ed., Lean Management Institute , 1999.
Sahin I., On the Objective Function Behavior in (s, S) Inventory Models, Operations Research, 30(4), 709 –724, 1982.
Schmidt C.P. and S. Nahmias, Optimal Policy for a Two-Stage Assembly System under Random Demand, Operations Research, 33(5), 1985.
Spearman M.L., M. Zazanis , Push and Pull Production Systems: Issues and Comparisons, Operations Research, 40(3), ProQuest Central, pg 521, 1992.
Spina G., Manufacturing paradigms versus strategic approaches: a misleading contrast, International Journal of Operations & Production Management, 18(8), 684-709, 1998.
Steele D.C., P.R. Philipoom , M.K. Malhotra and T.D. Fry, Comparisons between drum–buffer–rope and material requirements planning: a case study, International Journal of Production Research, 43(15), 3181-3208, 2005.
Taft E.W., The Most Economical Production Lot, The Iron Age, 101, 1410-1412, 1918.
Takahashi K., K. Morikawa & Y.-C. Chen, Comparing Kanban control with the theory of constraints using Markov chains, International Journal of Production Research, 45(16), 3599-3617, 2007.
Tardif V. and L. Maaseidvaag, An adaptive approach to controlling Kanban Systems, theory and methodology, European Journal of Operations Research, 132 411-424, 2001.
Thürer M., O.N. Fernandes and M. Stevenson, Production planning and control in multi-stage assembly systems: an assessment of Kanban, MRP, OPT (DBR) and DDMRP by simulation, International Journal of Production Research, available online, 2020.
Veinot Jr, A.F., and H.M. Wagner., Computing Optimal (s, S) Inventory Policies, Management Science, 11(5), Series A, Sciences (Mar., 1965), 525-552, INFORMS, 1965.
Wacker J.G., A theory of material requirements planning (MRP): an empirical methodology to reduce uncertainty in MRP systems, International Journal of Production Research, 23(4), 807-824, 1985.
Wagner H.M.and Whitin T.M., Dynamic version of the economic lot size model, Management Science, 5(1) 89-96, 1958.
Wang H. & H-P. Wang, Determining the number of kanbans: a step toward non-stock-production, The international journal of Production Research, 28(11), 2101-2115, 1990.
Watson K.J., J.H. Blackstone and S.C. Gardiner, The evolution of a management philosophy: The theory of constraints, Journal of Operations Management, 25, 387–402, 2007.
Watson K.J., and A. Patti, A comparison of JIT and TOC buffering philosophies on system performance with unplanned machine downtime, International Journal of Production Research, 46(7), 1869–1885, 2008.
Webster S.C., Principles and tools for supply chain management, Mc-Graw Hill Irwin, 2008.
Whitin T.M. , Inventory Control Research: A Survey, Management Science, 1,(1), 32-40, INFORMS, 1954.
Widyadana G.A., H.M. Wee and J-Y Chang, Determining the optimal number of Kanban in multi-products supply chain system, International Journal of Systems Science, 41(2), 189-201, 2010.
Yang L., X. Zhang and M. Jiang, An optimal Kanban system in a multi-stage, mixed-model assembly line, J Syst Sci Syst Eng, 19(1), 036-049, 2010.
Yano C.A. and R.C. Carlson, Interaction between frequency of rescheduling and the role of safety stock in material requirements planning systems, International Journal of Production Research, 25(2), 221-232, 1987.
Ye T. & W. Han, Determination of buffer sizes for drum–buffer–rope (DBR)-controlled production systems, International Journal of Production Research, 46(10), 2827-2844, 2008.
Zhao X. and J. Hou, Analyzing the time buffer in the Theory of Constraints based lean operations, Journal of Management Analytics, 3(1), 185-199, 2014.
Zipkins P. Foundations of Inventory management, Mc_Graw Hill Irwin, 2000.