Managing Inventory in Supply Chains with Nonstationary Demand

John J. Neale, Sean P. Willems
School of Management, Boston University, Boston, Massachusetts 02215
{jneale@bu.edu, willems@bu.edu}

Many companies experience nonstationary demand because of short product life cycles, seasonality, customer buying patterns, or other factors. We present a practical model for managing inventory in a supply chain facing stochastic, nonstationary demand. Our model is based on the guaranteed service modeling framework. We first describe how inventory levels should adapt to changes in demand at a single stage. We then show how nonstationary demand propagates in a supply chain, allowing us to link stages and apply a multiechelon optimization algorithm designed originally for stationary demand. We describe two successful applications of this model. The first is a tactical implementation to support monthly safety stock planning at Microsoft. The second is a strategic project to evaluate the benefits of using an inventory pool at Case New Holland.

Key words: multiechelon inventory optimization; stochastic, nonstationary demand; base-stock policy; supply chain application.
History: Published online in *Articles in Advance* June 10, 2009.

Nonstationary (i.e., changing with time) demand is the rule rather than the exception in most industries today. Companies in all markets are introducing new products at a higher frequency with increasingly shorter life cycles. For example, at Hewlett-Packard (HP), the product life cycle for a personal computer is often only three months. HP digital cameras have an average life cycle of less than 12 months, with some as short as 6 months. Demand is never really stationary as these products move through the launch, ramp, peak, and end-of-life phases of their life cycles (Figure 1). Not only does the rate of demand change over the life cycle, but the uncertainty does also. Johnson and Anderson (2000) document the case of an HP printer with a forecast error (as measured by the monthly coefficient of variation) of 135 percent during its ramp phase, 90 percent during its end-of-life phase, and only 15 percent during its peak phase.

Even within phases of the product life cycle, nonstationary demand is common. The demand for many products has a seasonal component. Microsoft experiences about two-thirds of the annual demand for its Xbox video game consoles in the 13 weeks before Christmas (Figure 2). Kraft Foods sells approximately twice as many hot dogs per week during the summer than it does during the rest of the year. Elmer's Products does almost 75 percent of its annual school-glue volume during the four-month back-to-school season from May to August. Again, demand uncertainty might also change by season. For example, because Elmer's is in more frequent contact with its retail channel during the busy back-to-school season, its forecast error drops from 53 percent in the low season to 40 percent during this high season.

When product demand is not related to the seasons of the year, many companies still experience monthly or quarterly "hockey stick" patterns (i.e., sales spikes at the end of the period) because of sales-force incentives and customer buying behavior. Dell, for example, experiences end-of-month peaks in demand for its enterprise products because of corporate buying patterns (Figure 3). Dell also experiences a demand spike in July because of school and government buying behavior. All these factors combine such that it is rare to find a company whose forecasts do not change depending on the month, if not week, of the year.

Managing inventory in the face of nonstationary, stochastic (i.e., uncertain) demand can be challenging. Many companies separate the problem into its strategic and tactical components. Strategic decisions
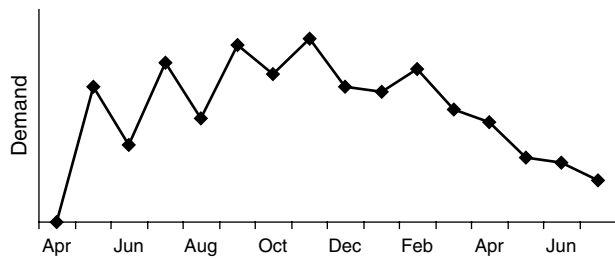
Figure 1: The graph illustrates nonstationary demand over the life cycle of a typical HP inkjet printer. Adapted from Johnson and Anderson (2000).
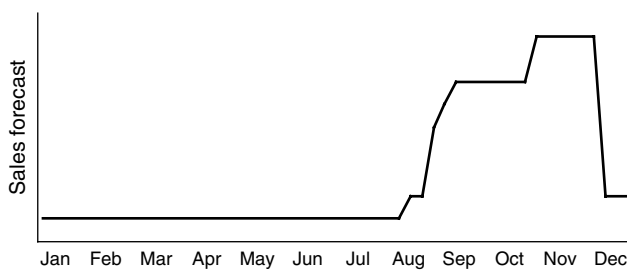

Figure 2: The graph illustrates the highly seasonal demand for Microsoft's Xbox video game consoles. Sales peak in the weeks leading up to Christmas.
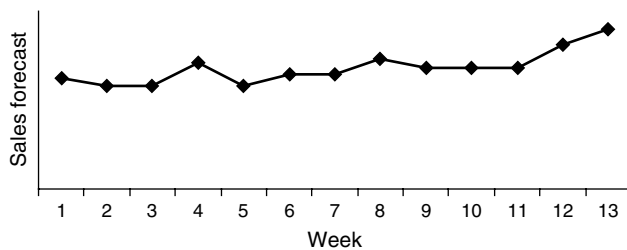

Figure 3: The graph illustrates an example of end-of-month and end-of-quarter peaks for a Dell projector. The timeline covers one quarter with months ending in weeks 4, 8, and 13.

include where to locate inventories in the supply chain, what levels of service to provide, and whether these choices should change over time. Tactical decisions primarily boil down to how to calculate time-phased inventory targets that change with demand from week to week, month to month, or season to season. In this paper, we present a model that many companies have applied successfully to help them make these strategic and tactical inventory decisions.

We organized the remainder of the paper as follows. In the next section, *Literature Review*, we position our approach relative to previous work on stochastic,

nonstationary inventory problems. In *A Nonstationary Supply Chain Inventory Model*, we present our model. We start by explaining the logic for a single stage and then show how stages can be linked and optimized in a complete supply chain model. We also discuss common extensions to the model. In the *Application Examples* section, we describe two successful implementations of our model. The first is a tactical project at Microsoft and the second is a strategic initiative at Case New Holland. For both, we highlight the need for a nonstationary solution and the insights and results achieved. Finally, we conclude the paper by summarizing the key learnings from our work and suggesting useful paths for future investigation.

Literature Review

Relative to stationary demand inventory models, much less work on nonstationary demand exists. Nonstationarity typically complicates the analysis and limits the results that can be obtained. Still, a number of authors have recognized the importance of this problem. We characterize this work by its focus on optimality versus performance evaluation and its assumptions about the underlying demand process.

The majority of work on stochastic, nonstationary demand inventory models focuses on characterizing the form of the optimal policy. These authors typically assume proportional holding and backorder costs. For a single stage without fixed order costs, a time-varying base-stock policy is optimal (Karlin 1960). When fixed order costs apply, Scarf's (1960) proof of the optimality of (s, S) policies extends immediately to nonstationary demand. Because the time-varying parameters for these optimal policies are difficult to compute, more recent work focuses on heuristic algorithms to calculate the parameters (Morton and Pentico 1995, Bollapragada and Morton 1999).

Some authors make specific assumptions about the nonstationary demand process to demonstrate results. Time-series demand models and Markov-modulated demand processes are most common. Graves (1999) uses an autoregressive, integrated moving-average (ARIMA) demand process to make observations about the level of inventory, amplification of demand variability, and value of information sharing in single and multistage systems with nonstationary demand.

Chen and Song (2001) show that an echelon base-stock policy with state-dependent order-up-to levels is optimal for a serial supply chain with Markov-modulated demand. Treharne and Sox (2002) consider the case in which demand is not only nonstationary but also partially observed (i.e., the decision maker does not know the true distribution of demand). They show that a state-dependent base-stock policy is optimal for a single stage with Markov-modulated demand, and test the performance of a number of heuristics.

We take a more practice-oriented approach to the stochastic, nonstationary demand inventory problem. We do not attempt to establish the optimality of our inventory policy. Instead, we select a policy to match the typical planning systems used in industry. Because backorder costs are difficult to quantify in practice, we use service level inputs to ensure sufficient inventory. We adopt a general demand model that matches the forecasting practices of the many companies with whom we have worked. In this sense, our approach is most similar to the supply chain inventory model developed at IBM by Ettl et al. (2000). These authors also construct a practical model to optimize supply chain inventories under the assumptions of base-stock policies, service level constraints, and nonstationary demand. However, their modeling approach and a number of key assumptions differ. Specifically, they use a queueing model with compound Poisson demand to model single-stage inventories and focus on approximating the replenishment lead times in the supply chain, assuming no expediting or lost sales.

A Nonstationary Supply Chain Inventory Model

The intent of our model is to address the problem of determining inventory locations and levels in a supply chain facing stochastic, nonstationary demand. The goal is to decide which locations should hold safety stock and exactly how much they should hold in each period. Given the inherent complexity of modeling nonstationary demand, we seek a pragmatic approach that requires approximations and compromises to get results that might apply in practice. We adopt the guaranteed service modeling framework, as

described in Graves and Willems (2000), and extend it to the case of nonstationary demand. This framework has been successfully applied at HP (Billington et al. 2004), among other places. Our work is closely related to Graves and Willems (2008), which also extends Graves and Willems (2000) to permit nonstationary demand. However, our paper is more practitioner focused, with new results, observations, and application examples.

In this section, we describe the key assumptions of our model and explain its logic, starting with a single stage and progressing to the complete supply chain network. We also describe common extensions to the basic model. The appendix includes supporting mathematical details.

Assumptions

We begin by highlighting the key assumptions of the guaranteed service (GS) modeling framework. A supply chain is modeled as a network in which each node, or stage, represents a major processing function and each arc denotes a supply relationship. Every stage operates under a base-stock inventory policy and is a potential location for holding safety stock. In other words, there is a common planning period at which time each stage observes demand and places a replenishment order so as to reach an inventory target. We assume that the lead time to replenish inventory is known and does not depend on the size of the order (i.e., there are no capacity constraints). Each stage promises to meet its customers' demand within a certain service time. Inventory levels are calculated so that every stage provides 100 percent service for all demand within a reasonable limit or bound. We refer the reader to Graves and Willems (2000) for a more complete discussion of the practicality of these assumptions.

We extend the GS framework to include nonstationary demand by adding a practical yet flexible demand process. In particular, we assume that the planning horizon is divided into different phases. Each phase can represent a stage of the product life cycle, season of the year, or a company-defined planning interval (e.g., week, month, etc.). We assume the demand within each phase is stationary with a known mean and standard deviation. However, we make no assumptions about the number of phases or the length

of each phase. The time phases can be as granular and uneven as necessary to capture the nonstationary nature of the demand. A common practice is to define the phases so that the demand forecasts within each phase are (relatively) level. One can then calculate demand uncertainty by comparing historical demands to forecasts for that same window of time in the past.

Extending the GS model to include this nonstationary demand process entails primarily two challenges. The first is generalizing the inventory equations of the single-stage model to account for the necessary changes in inventory levels over time. The second is characterizing how the nonstationary demand is propagated to upstream stages in the supply chain. We cover each of these in turn in the next two subsections.

Single-Stage Inventory Model

In this subsection, we present the single-stage model that serves as the building block for the multistage supply chain. Recall that each stage operates under a base-stock policy with a common review period of one time unit. Each period, a stage observes its demand and places an order to replenish its supply. In the appendix, we demonstrate that this order should make the total inventory position (on-hand plus on-order) large enough to cover demand over the next $SI + T - S$ time periods, where SI represents the inbound service time, T represents the lead time, and S represents the outbound service time. This order-up-to level is the base-stock level, $B(t)$, for the stage at time t. Because demand can differ across periods, the base-stock level will need to change or adapt over time along with demand. In all periods, it must be large enough to cover demand over the upcoming $SI + T - S$ periods, which we define as the net replenishment lead time (NRLT) for the stage.

We define the safety stock level, $SS(t)$, as the expected amount of inventory on hand at time t. This is the extra, "buffer" inventory we plan for on top of our expected demand. As we show in the appendix, the safety stock level at time t is a function of demand over the *preceding* NRLT. This is different than the base-stock level, which changes in anticipation of *upcoming* demand. The difference is because the replenishment order triggered by the base stock

does not result in on-hand inventory until $SI + T$ periods later. In other words, although we plan for the safety stock in advance, it does not materialize until the end of the replenishment lead time. Academic models typically use the base-stock level as their planning parameter and treat safety stock as a by-product, if at all. However, because commercial production-planning systems use safety stock as their input parameter and define it as *on-hand* inventory, it is important to get this timing correct.

This result is initially counterintuitive for most practitioners. They find it odd that safety stock changes as a function of the demand that comes before it, not after it. What they fail to appreciate is that at the time the order is placed to produce the safety stock for a period, the demand that will come before it is still unknown and will impact the supply that is carried into that period. This disconnect causes a common problem that Kraft Foods calls the landslide effect (Neale et al. 2009). Many companies experience a precipitous drop in inventory and service when transitioning from a high season to a low season because they plan their safety stocks based on forward-looking rules of thumb instead of demand from the backward-looking NRLT. These forward-looking targets pick up the low season prematurely, causing safety stock levels to drop and resulting in extremely poor service during the final periods of the high season. Kraft, for example, used to see service levels drop from 99 percent to 60 percent during the transition period for some highly seasonal products.

In addition to safety stock, there will also be pipeline (work-in-process) stock at a stage. As we demonstrate in the appendix, the pipeline stock level is primarily a function of the demand over the upcoming lead time. This future demand drives what we need to have in process today. However, the pipeline stock is also impacted by inbound and outbound service times. In particular, these times influence safety stock levels, and the pipeline stock must include planned adjustments to safety stock downstream because of nonstationary demand. However, these safety stock adjustments tend to counteract each other (especially with cyclic demand), so that over a sufficiently long horizon their impact on the average pipeline stock is negligible. We calculate and report the pipeline stock level by period but do not include it in our

optimization model because it is all but completely defined by the average demand rate and lead time.

Multistage Inventory Optimization

To model a multistage system, we use our single-stage model for each stage, where the inbound service time is a function of the outbound service times for the supplying stages, as Graves and Willems (2000) discuss. However, we need to characterize how the nonstationary demand is propagated to upstream stages in the multistage system. The demand seen by an upstream stage is the result of replenishment orders from its downstream stages. We derive the mean and variance of these orders for nonstationary demand in the appendix. We can make the following observations from these results. First, the expected demand (i.e., forecast) is shifted earlier in time by the NRLT as it is passed up the supply chain. In other words, the expected upstream demand at time t is a function of the expected downstream demand at time $t + $ NRLT. This just means that we plan ahead for any expected shift in demand. Second, the expected demand at upstream stages also consists of any adjustments to safety stock from their downstream stages. This is a classic example of the bullwhip effect—nonstationary demand is not only passed on; it is also amplified or dampened by changes to inventory levels in anticipation of changes in demand. Third, uncertainty is propagated differently than forecasts in a supply chain. The variance of demand is not shifted in time as it is passed upstream. Rather, it impacts upstream orders in the same period that it is experienced downstream. A downstream stage will plan for uncertainty through its safety stock, but it does not know how large its forecast error will be until it experiences the actual demand. Once this uncertainty is realized, it is immediately passed upstream as each stage recovers the difference between the actual demand and its forecast.

With the stages linked by demand moving upstream and supply moving downstream, we are ready to consider the problem of minimizing total supply chain inventory holding costs subject to end-customer service constraints. In the GS modeling framework, the service times for the internal stages of the supply chain determine where we place safety stocks and are the decisions variables for the optimization model. The case of nonstationary demand raises a new question of whether to allow nonstationary service times

in the model. Graves and Willems (2008) show by example that it can be optimal to change service times when transitioning from one time phase to another. However, we believe that the most practical solution is to require service times to be constant. We defend this constraint with the following arguments. First, we note that most companies want a consistent safety stock strategy. They do not want to push safety stock to a distribution center (DC) during one period and pull it back to the factory the next, only to push it out to the DC again later. Constant service times result in constant safety stock locations. Second, Graves and Willems (2008) show that this restriction is unlikely to result in a significant cost penalty. Through analysis and a computational experiment, they demonstrate that the optimal service times are robust to changes in the demand parameters. Finally, constant service times simplify the model and reduce the computational burden. We note that whereas service times are required to be constant, inventory levels will change with the nonstationary demand, as we described in the previous subsection.

With the assumption of constant service times, we can generalize the objective function from Graves and Willems (2000) to the case of nonstationary demand. We formulate the optimization problem in the appendix and note that we can rewrite the objective function in terms of the average safety stock levels over the horizon at each stage. Thus, the nonstationary problem is equivalent to the stationary problem but with average instead of constant safety stock levels. Graves and Willems (2008) present reasonable conditions under which the average safety stock for nonstationary demand is a concave function of the inbound and outbound service times. Therefore, we can solve the nonstationary demand optimization problem with existing algorithms for stationary demand. For example, we can use the dynamic program in Graves and Willems (2000) for supply chains modeled as spanning trees or the algorithms developed by Humair and Willems (2006) for general acyclic networks. Although the averaging required for nonstationary demand increases the computational complexity of the algorithms, we have found that a typical supply chain of approximately 100 stages with a dozen time phases can be calculated in less than a minute and optimized in less than 10 minutes on a standard Microsoft-based server.

Extensions to the Basic Model

We can extend our nonstationary supply chain inventory model to account for a number of additional real-world complexities. In this subsection, we briefly discuss two of the most commonly used extensions: stochastic lead times and arbitrary review periods.

In practice, the replenishment lead time for a stage is rarely truly constant. When the uncertainty in the lead time is significant, it is important to be able to account for it in our safety stock model. Stochastic lead times can greatly increase the amount of safety stock required to meet the service target. We can extend the single-stage inventory model described previously to include stochastic lead times by conditioning on the possible realizations of the random lead time variable. In particular, we assume a discrete and bounded distribution for lead time and use the conditional expectation and conditional variance formulas to calculate the mean and standard deviation of demand over the appropriate NRLT window as before. Most of our observations from the deterministic lead time case have a stochastic lead time analog. For example, the expected demand is still shifted earlier in time as it is passed up the supply chain. However, instead of being shifted by a deterministic NRLT, it is now a probability-weighted average of the expected demand at the possible values of the random NRLT.

Whereas extending the single-stage model is relatively straightforward, stochastic lead times complicate the multiechelon optimization. Because the safety stock is now a function of the expected demand and the expected demand at upstream stages depends (through the NRLT) on the service times at downstream stages, stochastic lead times create a circularity that the dynamic program (DP) of Graves and Willems (2000) cannot readily address. The DP needs to know the mean demand at all stages; however, the mean demand depends on the service times, which are an output of the DP. We circumvent this issue by propagating demand for the optimization algorithm assuming all service times are zero. We then use the resulting "optimal" service times to re-propagate demand for all inventory-level calculations.

It is also common in practice to find review periods that are greater than one time period and potentially differ across the supply chain. When this is the case, we need to model the impact of the review period on the inventory dynamics at a single stage and on the propagation of demand throughout the supply chain. In particular, a non-unit review period extends the NRLT at a stage, increasing its safety stock requirements. It also alters the ordering behavior of the stage, causing orders to become lumpier and cycle stock to accumulate. Bossert and Willems (2007) describe how the GS framework can be extended to account for arbitrary review periods. We note that our approach for nonstationary demand works nicely with their model. By altering ordering behavior, arbitrary review periods create nonstationary demand in a supply chain even when the end-customer demand is truly stationary. Therefore, Bossert and Willems' algorithms are already set up to deal with nonstationary demand. When the end-customer demand is also nonstationary, we simply need to calculate all results over the entire time horizon instead of just for the outbound demand cycle created by the review periods.

Application Examples

Optiant, Inc. has integrated the nonstationary supply chain inventory model described in the previous section into its PowerChain suite of supply chain software applications. More than a dozen Fortune 500 companies, including Black & Decker, Hewlett-Packard, Intel, Kraft, and Proctor & Gamble, have applied it. In this section, we describe two examples of the successful application of this model. The first is a tactical implementation to support the sales, inventory, and operations planning process at Microsoft. The second is a strategic project to evaluate the benefits of using an inventory pool at Case New Holland.

Tactical Inventory Planning at Microsoft

When they hear the name Microsoft, most people think of software. With its popular products, such as the Windows operating system and Office applications suite, Microsoft, which had approximately $52 billion in software-related revenue in fiscal year 2008, is the world's largest software company. However, Microsoft is also increasingly a hardware company. It introduced the Xbox video game console in 2001 and the Zune digital music player in 2006. It considers consumer electronics one of its three key growth markets. Microsoft's Entertainment and Devices (E&D)

Division, which includes Xbox, Zune, personal computer (PC) hardware such as mice, keyboards, and webcams, and retail sales of consumer software, had over $8 billion in revenue in fiscal year 2008.

Microsoft's hardware business has experienced exponential growth in recent years. When the company realized that its software supply chain could not support the hardware business, it set about putting new business processes in place. Having implemented SAP's Advanced Planner and Optimizer (APO) for production planning in 2001, it decided that the next step would be to develop a robust monthly sales, inventory, and operations planning (SI&OP) process around APO. In 2005, Microsoft initiated projects to improve inventory management across its supply chain. The new SI&OP process would include four parts (Figure 4). The first is a monthly forecast process that uses statistical techniques and collaboration between business experts to forecast future demand. The second is a stock-keeping unit (SKU) stratification process in which SKUs are prioritized into A, B, C, and D classifications based on revenue, life cycle status, margins, and marketing factors. Microsoft then sets service level targets for each SKU based on these classifications. The third part of the SI&OP process is a statistical safety stock model to calculate time-phased finished-goods safety stock targets to feed into APO. The fourth part is the APO planning system to optimize weekly production plans to best reach the inventory targets.

It was critical that Microsoft's safety stock planning solution be able to address nonstationary demand. Microsoft views innovation as a key driver of growth and is frequently introducing new products. The average life cycle is only two years for PC hardware
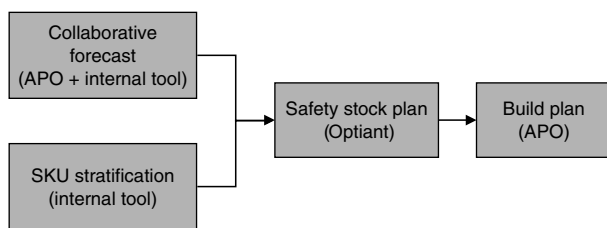

Figure 4: The flowchart shows the four parts of Microsoft's new SI&OP process: monthly forecasting, SKU stratification, safety stock calculation, and production planning.

products and five years for Xbox consoles. In addition, the E&D division's products are highly seasonal. Like many consumer products, they are in highest demand during the holiday shopping season. Typically, E&D generates over 40 percent of its annual sales during the months of October through December (Microsoft Corporation 2007). An adequate supply of products, such as the Xbox console, in the three months leading up to Christmas can mean the difference between a successful and disappointing fiscal year.

Microsoft selected Optiant's PowerChain suite for the safety stock planning component of its new SI&OP process. The software was implemented in 2006 by a project team consisting of representatives from Microsoft's global supply chain group and information technology teams, subject matter experts from the different product lines and geographic regions, and consultants from Optiant. The project was conducted in four phases. In the design phase, the team documented the business requirements and defined specifications for the software and its interfaces. In the build phase, the team added new features to the software and built the interfaces to Microsoft's data warehouses and APO. In the stabilize phase, the team tested the system and resolved all issues. In the acceptance phase, the team deployed the system and users validated results before putting the software into production. The team completed the total project in less than nine months.

Microsoft configured the PowerChain solution to produce weekly safety stock targets for a 52-week horizon. It used the nonstationary demand model described in the previous section with both stochastic lead times and arbitrary review periods enabled. The new statistically driven targets differed from Microsoft's old rule-of-thumb approach. Prior to this project, Microsoft had used a days-of-supply (DOS) safety stock rule for every finished SKU. Based on experience, intuition, or guesswork, supply planners would specify a number of days (usually 28) that the safety stock should cover. APO then converted this target to units by spreading it over the future forecasts until it was fully consumed. This approach resulted in safety stock targets that would adapt to changes in demand; however, they would do so in anticipation of future demand. Because safety stock
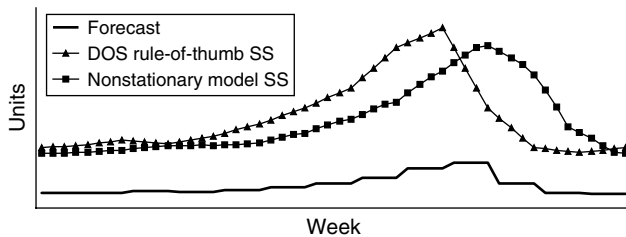
Figure 5: The graph shows a comparison of safety stock (SS) targets from Microsoft's old and new approaches. Targets change before they should under the old forward-looking DOS rule.

should be a function of the demand over the preceding NRLT, Microsoft's old targets were changing prematurely. Of course, there was also the issue of whether the planner had picked the right DOS target in the first place. Figure 5 shows an example of the difference between Microsoft's old approach and the correct nonstationary mathematics.

To help planners to overcome the psychological hurdle of letting an algorithm do what had been their jobs, Microsoft decided to implement "bumpers" in the software on the safety stock targets. Planners could specify an upper and (or) lower limit by SKU on the safety stock in terms of either absolute units or DOS. If the algorithm produced targets outside of these limits, the software would automatically round up (or down) to the appropriate limit and notify the planner that this had occurred. Some planners initially set the bumpers tightly around their old target values. However, as they saw the algorithms producing reasonable results and gained confidence in the approach over time, they gradually widened their limits.

Microsoft implemented the new safety stock solution worldwide for almost all its products sold through the consumer retail channel (approximately 10,000 SKUs). It is currently being used by approximately 30 supply planners to automatically calculate and feed time-phased safety stock targets to APO to guide production decisions. This relatively small number of planners is able to handle such a large number of SKUs because of the exception-based design of the solution's interface. Planners do not have the time to review the target for each SKU for every week. Instead, the software automatically calculates all targets and alerts planners only to those

targets that have changed significantly since the last run or have violated other user-specified rules.

Since implementing the nonstationary safety stock planning solution, the PC hardware business, which had metrics in place before the project, has seen an 18–20 percent increase in inventory turns and a 6–7 percent increase in fill rates. At the corporate level, Microsoft has seen net inventories steadily decline from a high of over $2.5 billion just before the solution went into production to the current levels of under $1 billion, all while revenues increased at an annual rate of almost 20 percent. For Microsoft's fiscal year 2008, the E&D Division reported its first profitable year. Qualitatively, Microsoft now has a consistent, automated, and data-driven process that reduces the emotionally charged debate in its monthly SI&OP process.

Strategic Inventory Planning at Case New Holland
Case New Holland (CNH) is a leading manufacturer of agricultural and construction equipment; its 2007 revenues exceeded $15 billion. The demand for its agricultural equipment is uncertain and highly seasonal. This demand is affected by macro conditions, such as projected crop prices and credit availability, and micro conditions, including when the harvest season begins in a given year because of rain and temperature. For example, Figure 6 shows the monthly demand for a utility tractor. Both the average demand and the demand uncertainty (as measured by the coefficient of variation, or CV) change significantly from month to month.

In North America, agricultural equipment is sold through a network of thousands of independent dealers. Because of economies of scale in manufacturing,
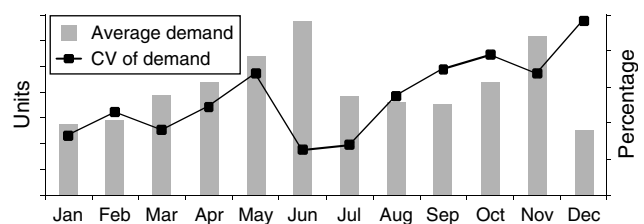

Figure 6: The graph shows an example of nonstationary demand for a CNH utility tractor. Both the mean and uncertainty (coefficient of variation (CV)) vary significantly by month.

individual product families tend to be produced in a single plant and then shipped to satisfy worldwide demand. For example, utility tractors are manufactured at a plant in Jesi, Italy. These tractors have engines that range from 50 to 110 horsepower. At a retail price ranging from $25,000 to $50,000 depending on the configuration, these tractors are highly maneuverable while still able to operate the common implements of a farming operation.

In the traditional distribution system, the utility tractors manufactured in Jesi are shipped directly to dealers. Because of the uncertainty of demand and long lead times caused by the long shipping distances, it was common for some dealers to have too much inventory and others too little. Furthermore, popular models would sell quickly; however, less popular models might sit for some time, leaving dealers with an incorrect product mix. In 2003, CNH sought to redesign its distribution network to become more responsive to dealer demands, while reducing the total amount of its working capital tied up in inventory. To accomplish this in North America, the company considered establishing a dealer pool in Dublin, Georgia. This dealer pool, which would stock popular configurations of Jesi tractors, would allow dealers to hold fewer tractors in inventory while receiving higher service levels and faster service times.

CNH modeled the Jesi postponement problem with the nonstationary inventory model that we described earlier in the paper as implemented in Optiant's PowerChain suite. The model, which included each popular configuration of Jesi tractor serving the different customer regions in North America, had 81 SKU locations. To accurately capture the nonstationary nature of demand, CNH divided the time horizon into eight phases. The model enabled CNH managers to quantify the costs and benefits of the dealer pool. They could compare the inventory costs incurred at the Dublin pool to the reductions in inventory it would enable in the dealer network. The ability to model seasonal demand was important to ensure user acceptance and produce actionable safety stock results. As a result of the modeling effort, CNH decided to implement the Dublin pool. After implementing the pool, total supply chain inventories decreased by more than 20 percent.
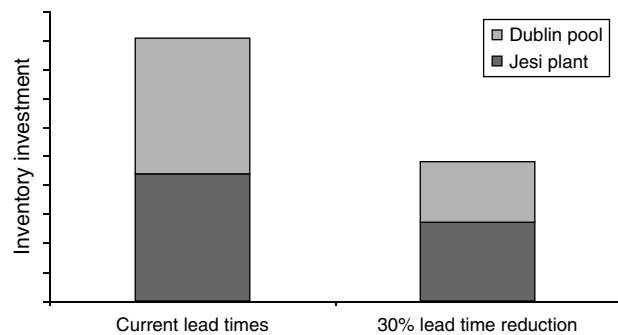

Figure 7: The graph shows an example of a sensitivity analysis that CNH performed. Reducing lead time and lead time variability by 30 percent reduced total inventories by almost 50 percent.

Once decision-support models are in place, they make performing sensitivity analyses easy. Users often underestimate this benefit of implementing such models. A significant portion of CNH's continuous-improvement activities focused on reducing lead time and lead time variability. The multiechelon inventory model allowed users to rapidly quantify the impact of lead time reductions in different parts of their supply chain and prioritize improvement initiatives (Figure 7).

Conclusions

In this paper, we presented a practical model for managing inventory in supply chains with nonstationary, stochastic demand. We generalized the single-stage inventory equations to account for the necessary changes in inventory levels over time. We also characterized how nonstationary demand is propagated in a supply chain and discussed how existing algorithms for stationary demand can be modified to optimize safety stocks across a supply chain. We described the application of the model by Microsoft to generate weekly safety stock targets to feed its production-planning system and by CNH to evaluate the benefits of an inventory-pooling strategy.

Many of the observations that result from our non-stationary model are initially counterintuitive. It is not immediately obvious that safety stock should be a function of backward-looking demand or that demand forecasts and demand uncertainty propagate differently through a supply chain. Most practitioners have only been exposed to stationary demand

inventory models, if any; their intuition for managing nonstationarity has not been developed. This is a clear case in which operations research-based decision-support systems can add value.

Although our model captures many of the real-world complexities typically associated with this problem, there are several areas worthy of further investigation. First, constraints on production capacity are common in highly seasonal businesses. It is often cost prohibitive to carry peak-demand capacity year-round. Extending the model to explicitly account for the impact of shared production capacity on time-phased safety stock targets would be of value. It also is common to find that suppliers quote longer lead times during a high season. Extending the model to include nonstationary replenishment lead times would address this situation. Further research to improve the computational performance of the optimization algorithm via approximations would be useful. Finally, simple rules or insights to guide practitioners when setting time-phased service level targets in the absence of accurate backorder costs would be valuable.

Appendix

In this appendix, we present a number of the mathematical details of our nonstationary supply chain inventory model. Interpretations and observations about the results are contained in the main body of the paper. Throughout this section, we assume that SI represents the inbound service time for a stage (i.e., the maximum of its suppliers' outbound service times), T represents its lead time, and S represents its outbound service time. We let $d(a, b)$ denote the demand over the time interval $(a, b]$ and assume that demand in period t has mean $\mu(t)$ and standard deviation $\sigma(t)$.

Single-Stage Inventory Equations
This subsection contains the derivations of the equations for base stock, safety stock, and pipeline stock at a single stage. We begin by deriving the base-stock level, $B(t)$, and safety stock level, $SS(t)$, at time t. Suppose we stand in period t. The order placed in period $t - SI - T$ will have just arrived. This order brought the inventory position at period $t - SI - T$ up to the base-stock level $B(t - SI - T)$. Because all orders prior to period $t - SI - T$ will also have arrived by

period t, the inventory position $B(t - SI - T)$ represents the total supply to inventory by period t, net of all demand up to and including period $t - SI - T$. In period t, demand from period $t - S$ is filled. The total demand filled from inventory since period $t - SI - T$ is $d(t - SI - T, t - S)$. Combining these observations, we can express the inventory on hand in period t, $I(t)$, as the difference between the total supply received and total demand filled by period t:

$$I(t) = B(t - SI - T) - d(t - SI - T, t - S). \quad (1)$$

To avoid a stockout, we require $I(t) \geq 0$. From Equation (1) we see that this requirement equates to

$$B(t - SI - T) \geq d(t - SI - T, t - S).$$

In practice, we create demand bounds by specifying a service level, $\alpha(t)$, which represents the probability of meeting all demand received in period t. If we assume independence over time and that the demand over the NRLT can be approximated by a Normal distribution, this service level can be achieved by setting

$$B(t - SI - T) = \sum_{i=t-SI-T+1}^{t-S} \mu(t) + z(t-S) \sqrt{\sum_{i=t-SI-T+1}^{t-S} \sigma^2(t)},$$

where $z(t)$ represents the number of standard deviations implied by $\alpha(t)$. Translating time units and substituting $\text{NRLT} = SI + T - S$, we get the following equation for the base-stock level:

$$B(t) = \sum_{i=1}^{\text{NRLT}} \mu(t+i) + z(t+\text{NRLT}) \sqrt{\sum_{i=1}^{\text{NRLT}} \sigma^2(t+i)}. \quad (2)$$

Taking the expected value of Equation (1) and applying Equation (2), we get the following equation for the safety stock level:

$$SS(t) = z(t-S) \sqrt{\sum_{i=1}^{\text{NRLT}} \sigma^2(t-S+1-i)}. \quad (3)$$

To derive the pipeline (work-in-process) stock, $W(t)$, we again assume that we stand at period t. All replenishment orders initiated on or before period $t - SI - T$ will have completed processing by period t. All replenishment orders initiated after period $t - SI$ will still be at the suppliers and will not yet have

entered our process. The orders in between constitute our pipeline stock. If $O(t)$ denotes the replenishment order initiated in period t, we can express the pipeline stock at time t as

$$W(t) = \sum_{i=t-SI-T+1}^{t-SI} O(t). \qquad (4)$$

In the next subsection, we derive the expectation and variance of the replenishment orders for our adaptive base-stock model. Taking the expectation of Equation (4) and borrowing Equation (8) from the next subsection, we do some algebra to arrive at the following equation for the expected pipeline stock level:

$$E[W(t)] = \sum_{i=1}^{T} \mu(t-S+i) + SS(t-SI) - SS(t-SS-T).$$

Multistage Demand Propagation

In this subsection, we derive the expectation and variance of the replenishment orders that result from our single-stage model. These orders become the demand at upstream stages in our multiechelon system. By definition of an adaptive base-stock policy, a stage's order at time t, $O(t)$, is determined by the following rule, where $IP(t)$ represents the inventory position in period t after demand fulfillment but before order placement, and $D(t)$ represents the demand:

$$O(t) = B(t) - IP(t)$$
$$= B(t) - (B(t-1) - D(t)). \qquad (5)$$

We note that demand is a random variable; however, the base-stock levels are known functions of time and demand parameters. We can use Equation (5) to determine the expectation and variance of $O(t)$:

$$E[O(t)] = E[B(t) - B(t-1) + D(t)]$$
$$= B(t) - B(t-1) + \mu(t), \qquad (6)$$
$$Var[O(t)] = Var[B(t) - B(t-1) + D(t)]$$
$$= Var[D(t)]$$
$$= \sigma^2(t). \qquad (7)$$

Substituting the base-stock from Equation (2) into Equation (6), we do some algebra to arrive at

$$E[O(t)] = \mu(t + NRLT) + \Delta SS(t), \qquad (8)$$

where $\Delta SS(t)$ is the safety stock adjustment for period t and is given by

$$\Delta SS(t) = z(t + NRLT) \sqrt{\sum_{i=1}^{NRLT} \sigma^2(t+i)}$$
$$- z(t + NRLT - 1) \sqrt{\sum_{i=1}^{NRLT} \sigma^2(t+i-1)}.$$

Multistage Optimization Problem

This subsection uses the single-stage and demand propagation results to formulate the multistage optimization problem. We first use Equation (7) to propagate demand variability throughout the supply chain. Specifically, we let ϕ_{ij} represent the number of units from upstream stage i required per unit at downstream stage j and A represent the set of arcs in the supply chain network. Assuming independence between stages, we propagate variances as follows:

$$\sigma_i^2(t) = \sum_{(i,j) \in A} \phi_{ij}^2 \sigma_j^2(t).$$

We propagate service levels using a more complicated weighted average of the variances at each stage; we omit the details in the interest of space.

We are now ready to pose the optimization problem using Equation (3) to calculate safety stocks. However, because the inbound and outbound service times at a stage are decision variables in our multistage framework, we modify our single-stage notation. We write the left side of Equation (3) as $SS_j(t, S, SI)$ to acknowledge its dependence on these variables, where the subscript j denotes the stage. We then formulate the multistage optimization problem as

$$\text{Min} \ \sum_{t=1}^{H} \sum_{j=1}^{N} h_j SS_j(t, S, SI)$$

$$\text{s.t.} \quad SI_j + T_j - S_j \geq 0 \quad \text{for all } j = 1, \ldots, N;$$

$$SI_j \geq S_i \quad \text{for all } (i, j) \in A;$$

$$S_j \leq s_j \quad \text{for all end-customer stages } j;$$

$$S_j, SI_j \geq 0, \text{ integer} \quad \text{for all } j = 1, \ldots, N;$$

where h_j denotes the inventory holding cost per unit per time period at stage j, s_j the maximum acceptable service time for customer stage j, H the number of

periods in the horizon, and N the number of stages in the supply chain. The objective is to minimize the safety stock holding cost over the planning horizon. The constraints ensure that the NRLTs are nonnegative, each stage's inbound service time is as large as the maximum service time quoted to it, and the service times to customers meet their requirements.

Last, as in Graves and Willems (2008), we rewrite the objective function in terms of the average safety stock levels over the horizon to facilitate its solution with existing algorithms for the stationary-demand problem:

$$\sum_{t=1}^{H}\sum_{j=1}^{N} h_j SS_j(t, S, SI) = H \times \sum_{j=1}^{N} h_j \left(\frac{\sum_{t=1}^{H} SS_j(t, S, SI)}{H} \right).$$

References

Billington, C., G. Callioni, B. Crane, J. D. Ruark, J. U. Rapp, T. White, S. P. Willems. 2004. Accelerating the profitability of Hewlett-Packard's supply chains. *Interfaces* **34**(1) 59–72.

Bollapragada, S., T. E. Morton. 1999. A simple heuristic for computing nonstationary (s, S) policies. *Oper. Res.* **47**(4) 576–584.

Bossert, J. M., S. P. Willems. 2007. A periodic-review modeling approach for guaranteed service supply chains. *Interfaces* **37**(5) 420–435.

Chen, F., J. S. Song. 2001. Optimal policies for multiechelon inventory problems with Markov-modulated demand. *Oper. Res.* **49**(2) 226–234.

Ettl, M., G. E. Feigin, G. Y. Lin, D. D. Yao. 2000. A supply network model with base-stock control and service requirements. *Oper. Res.* **48**(2) 216–232.

Graves, S. C. 1999. A single-item inventory model for a nonstationary demand process. *Manufacturing Service Oper. Management* **1**(1) 50–61.

Graves, S. C., S. P. Willems. 2000. Optimizing strategic safety stock placement in supply chains. *Manufacturing Service Oper. Management* **2**(1) 68–83.

Graves, S. C., S. P. Willems. 2008. Strategic inventory placement in supply chains: Nonstationary demand. *Manufacturing Service Oper. Management* **10**(2) 278–287.

Humair, S., S. P. Willems. 2006. Optimizing strategic safety stock placement in supply chains with clusters of commonality. *Oper. Res.* **54**(4) 725–742.

Johnson, M. E., E. Anderson. 2000. Postponement strategies for channel derivatives. *Internat. J. Logist. Management* **11**(1) 19–35.

Karlin, S. 1960. Dynamic inventory policy with varying stochastic demands. *Management Sci.* **6**(3) 231–258.

Microsoft Corporation. 2007. Annual Report 2007. Retrieved June 1, 2008, http://www.microsoft.com/msft/reports/default.mspx.

Morton, T. E., D. W. Pentico. 1995. The finite horizon nonstationary stochastic inventory problem: Near-myopic bounds, heuristics, testing. *Management Sci.* **41**(2) 334–343.

Neale, J. J., S. P. Willems, J. Beyl. 2009. Misalignment in the timing of nonstationary demand and supply: The landslide effect. Working paper, School of Management, Boston University, Boston.

Scarf, H. 1960. The optimality of (S, s) policies in the dynamic inventory problem. K. J. Arrow, S. Karlin, P. Suppes, eds. *Mathematical Methods in the Social Sciences*. Stanford University Press, Stanford, CA, 196–202.

Treharne, J. T., C. R. Sox. 2002. Adaptive inventory control for nonstationary demand and partial information. *Management Sci.* **48**(5) 607–624.