PNG: Effective Inventory Control for Items with Highly Variable Demand

Tovey C. Bachman, Pamela J. Williams, Kristen M. Cheman, Jeffrey Curtis, Robert Carroll

Interfaces
Vol. 46, No. 1, January–February 2016, pp. 18–32
ISSN 0092-2102 (print) | ISSN 1526-551X (online)

informs
http://dx.doi.org/10.1287/inte.2015.0829
© 2016 INFORMS

THE FRANZ EDELMAN AWARD
Achievement in Operations Research

PNG: Effective Inventory Control for Items with Highly Variable Demand

Tovey C. Bachman, Pamela J. Williams, Kristen M. Cheman
LMI, Tysons, Virginia 22102
{tbachman@lmi.org, pwilliams@lmi.org, kcheman@lmi.org}

Jeffrey Curtis
Defense Logistics Agency, Fort Belvoir, Virginia 22060, jeffrey.curtis@dla.mil

Robert Carroll
Office of Secretary of Defense, Washington, DC 20301, robert.w.carroll34.civ@mail.mil

LMI developed the PNG inventory control solution to manage inventory items with infrequent demand (i.e., isolated spikes in demand) as well as items with frequent, highly variable demand. Such items account for the majority of hardware stocked at the U.S. Defense Logistics Agency (DLA). The forecasting of demand for these items—no matter how sophisticated the forecasting method—had resulted in years of problems for DLA: excess inventory for some items, backorders for others, and excessive buyer workload. The implementation of PNG, a software package that consists of two inventory solutions, Peak Policy and Next Gen, allowed DLA to shift from trying to forecast each item individually to using a portfolio or risk-management approach to inventory control. Since DLA implemented PNG in January 2013, the agency has achieved its inventory-related goals for better customer service and reduced buyer workload, has experienced no inventory increase, and has saved nearly $400 million per year.

Keywords: inventory control; material management; risk management; optimization; simulation; supply planning.
History: This paper was refereed.

LMI, a not-for-profit government consulting firm, developed a unique inventory management tool to help the U.S. military's Defense Logistics Agency (DLA) better manage its inventory portfolio. The PNG software package combines two inventory solutions: Peak Policy (Bachman and DeZwarte 2007) and the Next Generation Inventory Model (Next Gen) (Bachman et al. 2011). With PNG's implementation, LMI shifted DLA away from using traditional forecasting methods to employing a more effective portfolio or risk-management approach to inventory control.

DLA uses PNG to consider the trade-offs among customer service, inventory value, and buyer workload goals. DLA can then make a single inventory decision that aligns with its organizational objectives—without separate investments in forecasted demands, safety stock, or order quantities. The result has been better customer service for the military servicemen that maintain and repair equipment for the U.S. warfighter—DLA's ultimate customer—and an impressive reduction in buyer workload, but with no increase in inventory. DLA is saving nearly $400 million per year—all from buying more of what sells and less of what does not sell.

Here, we briefly describe the terminology we use in this paper.

- *Demands* are requests for items from DLA's inventory.
- *Fill rate* is the percentage of orders filled completely when an order is received.
- *Inventory position* is DLA's stock on hand plus the stock that is due in from replenishment orders; a negative on-hand inventory position denotes backorders.

- *Reorder point*, denoted by s in (s, S) inventory control doctrine, is the lower control level. When an item's inventory position drops to (or below) the reorder point, a replenishment order is triggered.
- *Requisitioning objective*, denoted by S in (s, S) inventory control doctrine, is the upper control level. When an order is triggered, the order size is the difference between the inventory position and the requisitioning objective.
- *Echelon* is a stage of the supply chain. The items that PNG manages are stocked only at the wholesale echelon. DLA manages other items as retail stock; we do not include them in our discussion.
- *Wait time* is the number of days between the date an order is received and the date the last unit of material filling the order is issued.

Background

DLA manages wholesale inventories of consumable items for the U.S. Department of Defense and other federal agencies. (The military services have their own retail stocks of a lesser range of consumable items and retain management of repairable items.) By any standard, DLA is a large business. It has $38 billion in annual sales (all to branches of the U.S. government and allied forces) and $21 billion in inventory, handles 131,000 customer orders and 10,000 contracting actions per day, and has more than five million stock numbers. DLA manages nine supply chains that cover everything from aviation, maritime, and land systems to industrial hardware, construction equipment, fuels, medical supplies, subsistence, and clothing. Its five hardware supply chains are among the most difficult to manage. (DLA uses the term hardware generically for parts, equipment, and related items.)

The operating environment for the military is dynamic and unpredictable. Customer demands for hardware items are affected by wars, politics, military policy, and congressional budget turbulence. These causal factors are inherently unforecastable and drive a level of uncertainty that makes planning extraordinarily challenging.

Drifting failure patterns (patterns that are not well explained by theoretical probability distributions) for the military components within these supply chains compound the demand uncertainty. Adding to this challenge, DLA sometimes has difficulty getting vendors to bid for supply or manufacturing contracts. Contracts for military hardware items are issued at irregular intervals, and vendors often do not view DLA as one of their "best" customers. When vendors do work DLA items into their production schedules, orders are fulfilled at the manufacturer's convenience, not DLA's. As a result, lead times can range from a few months to a few years.

A Problem with Multiple Dimensions

DLA has long recognized that its hardware supply chains contain items with infrequent and highly variable demand; therefore, it segmented its hardware item populations in an attempt to better manage these items based on demand characteristics. Statistical, demand-based forecasting methods perform poorly for many of these items, no matter how sophisticated the forecasting algorithm. Therefore, we refer to these as items with unforecastable demand; forecasting simply leads to undesirable business outcomes (e.g., backorders, excess inventory, excess procurement workload, wasted working capital).

Commercial businesses might choose not to stock such slow-moving items, but DLA must stock such items because they are essential to support military missions. In the military, as with many industries, if a system is down for lack of a part, the mission is in jeopardy. Having the right part available can truly mean the difference between life and death.

To accommodate these problematic but essential items, DLA segmented its hardware items based on demand frequency. As Figure 1 illustrates, as demand becomes more frequent (and consistent), we see fewer items responsible for demand activity. The majority of DLA items are demanded infrequently (with demand in less than half the quarters). In addition, investment risk—the danger that DLA invests in inventory that ends up not being used—increases as the frequency of demand decreases (moving from right to left in Figure 1). The risk to customer service also goes up, because it is difficult to predict with any accuracy the specific items customers will actually need.

Realizing that forecasting did not work well for the slower-moving items, DLA had been managing items within the infrequent demand segment using

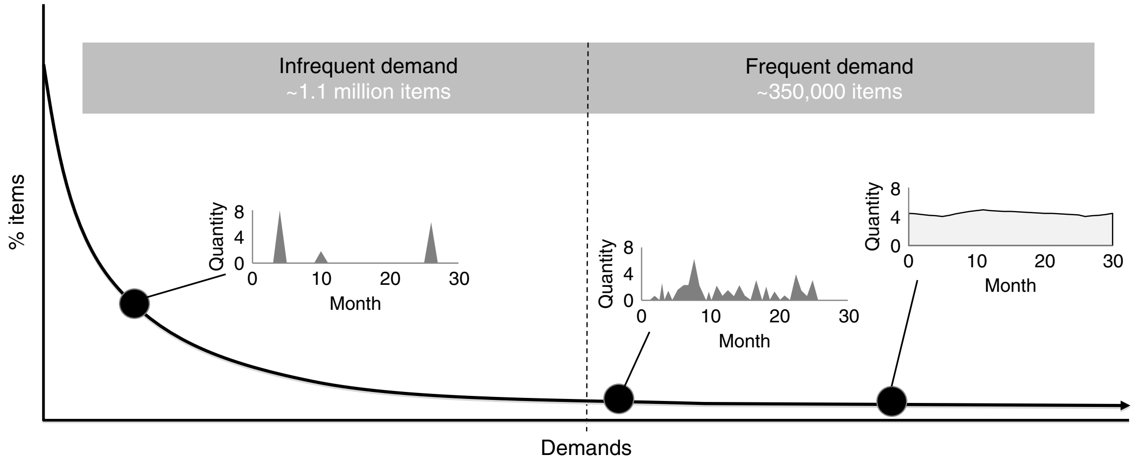

Figure 1: DLA segmented its hardware business according to frequency of demand to accommodate peculiarities in demand patterns.

a relatively simple set of business rules. Items with more frequent demand were managed using traditional forecasting and demand planning tools—until the agency realized that not all high-frequency items were forecastable.

The high-use segment includes a mix of items with low demand variability (where standard forecast methods achieve good business outcomes) and items with high variability (where existing forecast tools just do not work well). Therefore, DLA's segmentation of its hardware population evolved from two to three segments (Figure 2).

- Items that are forecastable because their demand is frequent and has low variability;
- Items that are unforecastable because of infrequent demand;
- Items that are unforecastable because of frequent, but highly variable, demand patterns. (Frequent demand is defined as having demand in more than half the quarters. When the ratio of standard deviation of quarterly demand to mean quarterly demand exceeds 1.0, we consider the item's demand to be highly variable.)

Infrequent and unforecastable demand: Within the segment of infrequent, irregularly spaced demand, items experience demand spikes. Some demands occur months apart; others occur two, three, or even four years apart. Demand can be so sparse that keeping manufacturing capacity online is not economical. If the spikes are more regularly spaced, DLA would have a much easier problem; it could predict when demand might occur. Unfortunately, the problem is more complex than that, and the best forecast for most of these items is zero. But using a forecast of zero would result in zero stock—an unacceptable prospect for critical military items.

Frequent but still unforecastable demand: Within the segment of frequent, highly variable demand, items experienced demand in most months. The issue was in the variable size of the demand. During some months, the demand may have been for five units; in the next month, it may have been for 500 units. Demand frequency is high, but the timing and size of the demand is irregular. For this segment, commercial forecasting and planning software, no matter how sophisticated, has errors that can be as large as 200 percent, and the data do not fit standard probability distributions.

LMI's goal was to help DLA find an approach that would significantly improve business outcomes for these two challenging segments (infrequent, isolated spikes in demand and frequent, highly variable demand).

Performance: Evaluating the Competition

DLA decided to keep its easiest segment of the hardware population (shown on the right in Figure 2) under its current forecast-based inventory control.

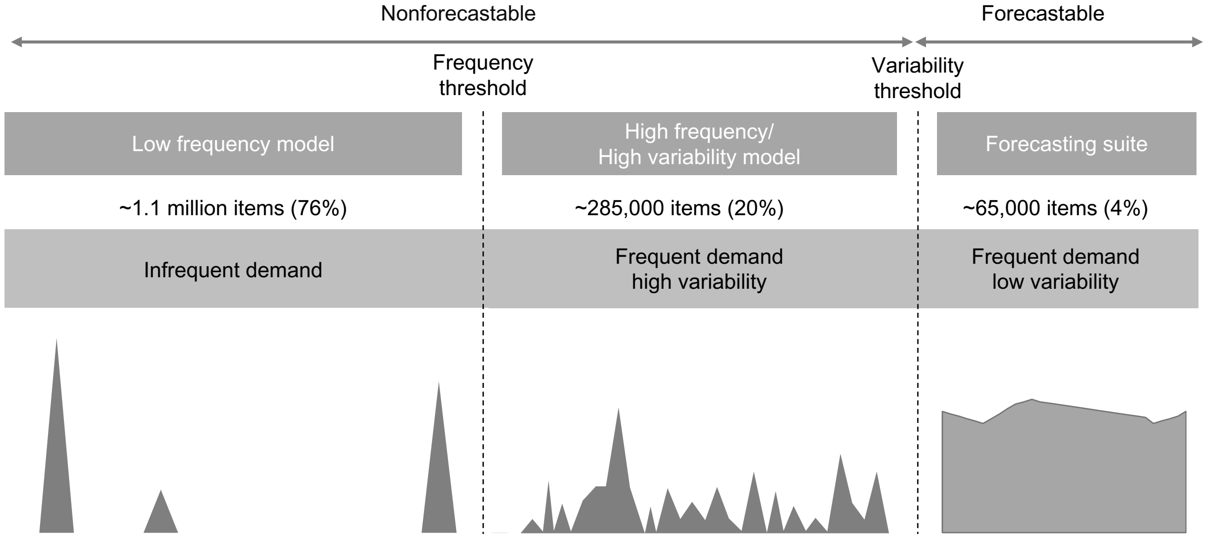

Figure 2: The segmentation of DLA's hardware business into three parts was an important step in achieving the ultimate solution.

For the other two segments, the agency considered alternatives that still fit into the planning structure of using a forecast or demand plan, safety stock (stock to cover demands that exceed a forecast of demand over a lead time), and an order quantity expected to cover demand over a fixed duration (also known as a coverage duration).

Using inventory simulations, DLA conducted a multiple-round competition of inventory forecasting solutions. Each round varied the forecasting piece of the planning logic. For some rounds, the goal was to improve customer service without increased inventory; other rounds focused on inventory reduction. DLA considered as many as 25 forecasting methods.

At the time DLA expanded the competition to include PNG, its goal was to reduce inventory value and procurement workload. PNG outperformed all other alternatives for all metrics (Table 1).

Because LMI's goal was to find the best solution for our client, we analyzed other alternatives (beyond those included in DLA's competition) for items with infrequent demand. These included Croston's method (Croston 1972) and bootstrapping-based stock levels. Croston's method forecasts demand sizes and time between demands. The forecasts are combined to produce a forecasted lead-time demand. Bootstrapping develops an empirical distribution of lead-time demand; Fricker and Goodhart (2000) describe bootstrapping and show an application. Using a simulation, we compared PNG with both Croston's method and bootstrapping. PNG required 30–50 percent less on-hand inventory to yield the same level of customer service.

Goals: Reduce inventory and replenishment workload

Scenario	Inventory $ (%)	Wait time (%)	Cash outlay (%)	% of orders	Fill rate (%)
Best of all other inventory control methods	−1.2	70.7	−4.4	−5.9	−10.7
PNG	*−7.1*	*−3.6*	*−7.5*	*−40.9*	*0.8*

Table 1: PNG outperformed more than 25 other inventory control methods tested by DLA.
***Note.* Table entries are percent changes from baseline.**

For the frequently demanded, but highly variable demand items, we tested several alternative safety stock methods. These methods assumed a wide variety of probability distributions (for the quantity demanded over a fixed lead time), including normal, gamma, negative binomial, lognormal, Laplace, Poisson, uniform, and Weibull. Variance estimates included those based on mean absolute deviation and mean-squared error. None of the safety stock methods we evaluated improved metrics for the highly variable segment. Statistical goodness-of-fit tests performed on the theoretical distributions rejected all methods (less than 0.5 percent probability of a false rejection), which is not surprising because many demand distributions are multimodal.

None of the alternative methods examined by either LMI or DLA for the problematic segments of DLA inventory yielded a significant improvement in bottom-line business metrics. The exception was PNG, which abandons the forecast and safety stock model for a radically different, risk-management-based approach.

In 2013, LMI implemented PNG for the five DLA hardware supply chains. It was a large-scale implementation, with more than 500,000 stock-keeping units (SKUs) that accounted for $1.5 billion in annual sales. (An SKU is an item or part number at a location; this initial implementation only covered items with a single location.)

Trade-Offs Among Business Objectives

DLA also needed to set priorities and configure the model to suit the agency's business goals. PNG presents supply chain managers with trade-off curves, which illustrate the three-way trade-off of different business outcomes. For example, the horizontal axis in Figure 3 presents on-hand inventory values; the vertical axis is average wait time, and the shading represents different levels of procurements (buys) awarded annually. Alternatively, the vertical axis can show fill rate.

By picking a single operating point—just one decision—an inventory supply chain manager can make a simultaneous three-way trade-off among operational goals (for example, customer service, inventory, and annual buys).

DLA supply chain managers started with the baseline (the black square at the upper right in Figure 3) or current process. They considered several options, such as pushing to reduce inventory value, with more modest improvements in customer service and reductions in buyer workload. Other options emphasized customer service and procurement workload.

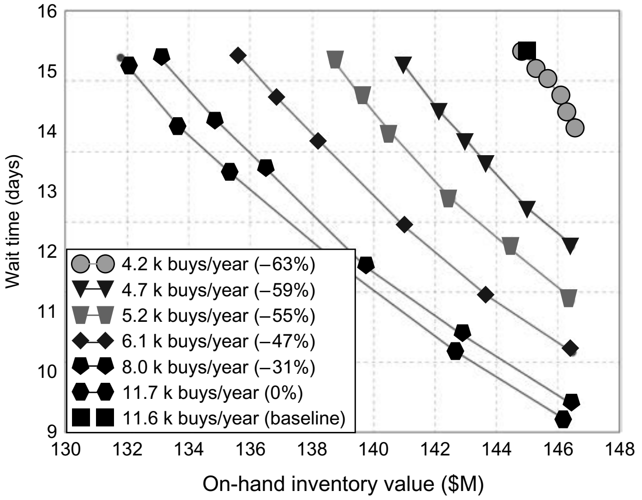

Figure 3: With PNG, DLA can select a point on a curve that best aligns with its overall business goals.

PNG offered DLA up to 30 possible trade-offs in long-term metrics, and an analysis of the near-term effects (e.g., obligation authority—or the annual expenditures on replenishing material—or number of immediate buys) for each option.

DLA did not pick points at the enterprise level. Instead, each supply chain manager selected a point that best met his (her) unique goals. Overall, DLA chose to keep average on-hand inventory about the same, while reducing both the number of buys generated annually and customer wait time. Each supply chain manager revisits the trade-off curves annually.

Once a trade-off curve point is selected, a simulation model projects a range of annual expenditures across simulation trials. The one-year expenditure associated with a trade-off point can be one factor considered in the point selection.

Selecting a point on a particular trade-off curve translates to fixed PNG model parameters for each supply chain and specific weights for the three business metrics. The PNG model parameters are then used during subsequent quarterly processing to generate an (s, S) pair for each item (Scarf 1960). This quarterly review updates levels to reflect changes to item demand transactions and item characteristics, such as lead times, prices, and asset position; but it does not change the PNG settings established during the annual process.

Peak Policy

Peak Policy (Peak) is analogous to a surge protector for an electrical circuit. The normal voltage is a constant 12 volts (V), but sometimes there are voltage spikes (e.g., the largest voltage spike might be 50,000 V). We may not be able to afford a surge protector sophisticated enough to protect against a 50,000 V spike, but we can afford one that will protect against 10,000 V. This may cover 90 percent of the voltage spikes but still keep the protection level affordable. The normal no-demand situation is analogous to the standard 12 V, and the infrequent demands are the voltage spikes.

Instead of forecasting, Peak determines the (s, S) pair for each item in a population using a simulation-based hedging strategy that balances the risks of being either out of stock or overinvesting. Because Peak considers the overall population as a portfolio (much like a stock portfolio), it does not attempt to predict the outcomes for individual items.

Peak also uses all of an item's historical demands for the previous five years to develop a risk profile for future demand, even if the only demand occurred five years ago. This means that Peak evaluates the possibility that future demand will be as large as any quarterly demand within that five-year window. Experimenting with shorter windows (i.e., one, two, and three years) showed little improvement in customer service. At four years, however, customer service dramatically improved, and it improved even further at five years. Moving to six- and seven-year windows improved service further but required significantly greater inventory value. Five years offers the best balance between the contrasting goals of excellent service and low inventory investment. Peak does not assume another demand spike will occur, nor does it assume full coverage for a similar demand spike. Peak simply uses past demand to illustrate a range of demands that could occur under unforeseen circumstances, such as a war.

Figure 4 illustrates the steps of the Peak process, from raw data to trade-off curves.

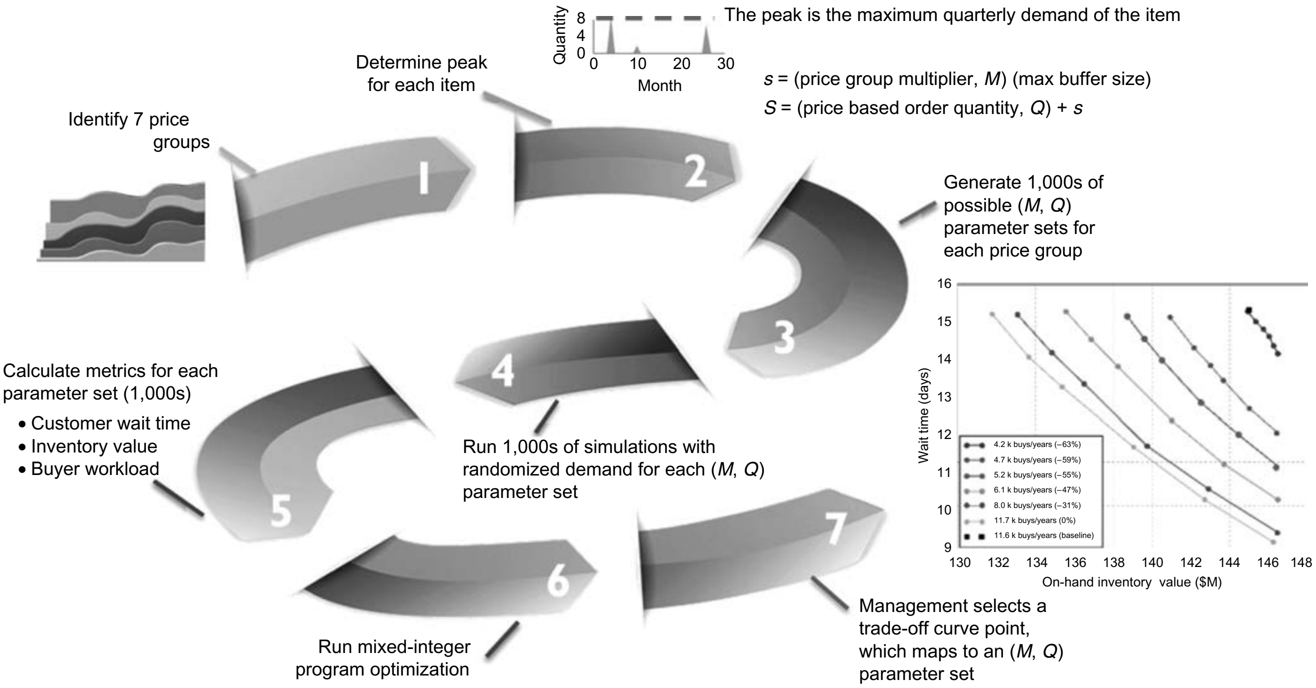
Quantity
8
4
0
0
10
20
30
Month
The peak is the maximum quarterly demand of the item
Determine peak for each item
s = (price group multiplier, M) (max buffer size)
S = (price based order quantity, Q) + s
Identify 7 price groups
1
2
3
Generate 1,000s of possible (M, Q) parameter sets for each price group
4
Calculate metrics for each parameter set (1,000s)
• Customer wait time
• Inventory value
• Buyer workload
Run 1,000s of simulations with randomized demand for each (M, Q) parameter set
5
6
7
Run mixed-integer program optimization
Management selects a trade-off curve point, which maps to an (M, Q) parameter set
Wait time (days)
16
15
14
13
12
11
10
9
130
132
134
136
138
140
142
144
146
148
On-hand inventory value ($M)
4.2 k buys/years (−63%)
4.7 k buys/years (−59%)
5.2 k buys/years (−55%)
6.1 k buys/years (−47%)
8.0 k buys/years (−31%)
11.7 k buys/years (0%)
11.6 k buys/years (baseline)

Figure 4: Peak follows seven steps to go from raw data to trade-off curves for items with infrequent demand.

First (step 1 in Figure 4), the items are sorted by unit price, and we identify the 10th, 25th, 50th, 75th, 90th, and 95th percentile prices. These price breaks form the boundaries for seven price groups.

Next, for each individual item, the model finds the largest spike in demand (i.e., the peak) over a 20-quarter period (Figure 4, step 2). The item's reorder point is set to a yet-to-be-determined multiplier (M) times the peak demand. To keep buyer workload down, Peak uses price-based order quantities and replenishes less expensive items in larger batches than expensive items. The requisitioning objective for an item is its reorder point plus a yet-to-be-determined order quantity (Q).

The best (M, Q) sets are determined through simulation and optimization. The Peak algorithm creates thousands of sets of possible multipliers and order quantities for each price group (Figure 4, step 3); these are sets of model parameters.

The model then simulates randomized demand spikes (using an empirical distribution based on actual demand history) over a period of years, typically running 20 trials (20 possible futures). The same demand patterns are used to stress test thousands of model parameter sets (Figure 4, step 4). Our approach differs from other simulations that assume either a constant demand rate or a theoretical textbook probability distribution of demand.

From the simulations, Peak projects population-level outcomes for customer service, inventory value, and buyer workload (Figure 4, step 5) for each parameter set. These population-level outcomes are the coefficients of the objective function and constraints of a mixed-integer program (MIP) problem (Figure 4, step 6). A branch-and-cut solver finds the best (M, Q) sets to attain various trade-offs between business outcomes. Appendix A includes the MIP formulation.

Each trade-off curve generated corresponds to a relatively narrow range for annual buys, almost a level curve. In step 7 of Figure 4, a manager selects an operating point on a trade-off curve that aligns with the organization's long-term goals (three to five years out). The point selected maps to a parameter set (a specific M and a specific Q for each price group).

Finally, to set an item's (s, S), the algorithm determines the price group in which the item falls and applies that price group's M and Q, as we explain earlier, to set (s, S) for each item. Point selection (determining Ms and Qs) is typically done annually, whereas the (s, S) pairs are updated quarterly to reflect changes in the demand history.

The Next Generation Inventory Model

Now let us turn to Next Gen, our solution to the problem of frequent, but highly variable demand. The Next Gen model frames decisions about the size and timing of buys according to empirical distributions of demand quantities and the time between demands. Although Next Gen was designed for items with frequent, albeit variable demand, the model also works well on more consistent demand, where forecast-based approaches have historically been applied.

Next Gen starts with an original stream of customer orders—five years of orders if they are available; at least two years of orders are required. Next Gen first builds histograms (Figure 5, step 1) for order interarrival times and order sizes, and weights more recent observations more heavily, which allows the model to adapt to changing demand over time. (Balancing responsiveness to a slowly developing trend without overreacting to spikes in demand was a major challenge.)

Using renewal theory (Ross 1992), and building on the work of Sahin (1979, 1982), Next Gen estimates the probabilities of the inventory position for a wide range of potential (s, S) pair combinations (Figure 5, step 2.a) directly from the histograms. The inventory-position probability distribution is computed from n-fold convolutions of the demand-size histogram.

From convolutions of the interarrival time and demand-size histograms, we also arrive at probabilities for different amounts of customer demand (Figure 5, step 2.b) for the fixed lead time in the input data, which may result in a multimodal, badly behaved distribution (Figure 6).

Next Gen generates multiple sets of penalty factors for backorders, inventory value, and annual buys (Figure 5, step 3.a). It builds a cost function for each set (Figure 5, step 3.b), which adds terms for requisition backorders and inventory position (beyond the Sahin objective function). This balances service for items with small demand sizes with items with large demand sizes and prevents overinvesting in a few items at the expense of all the others. However, this new cost

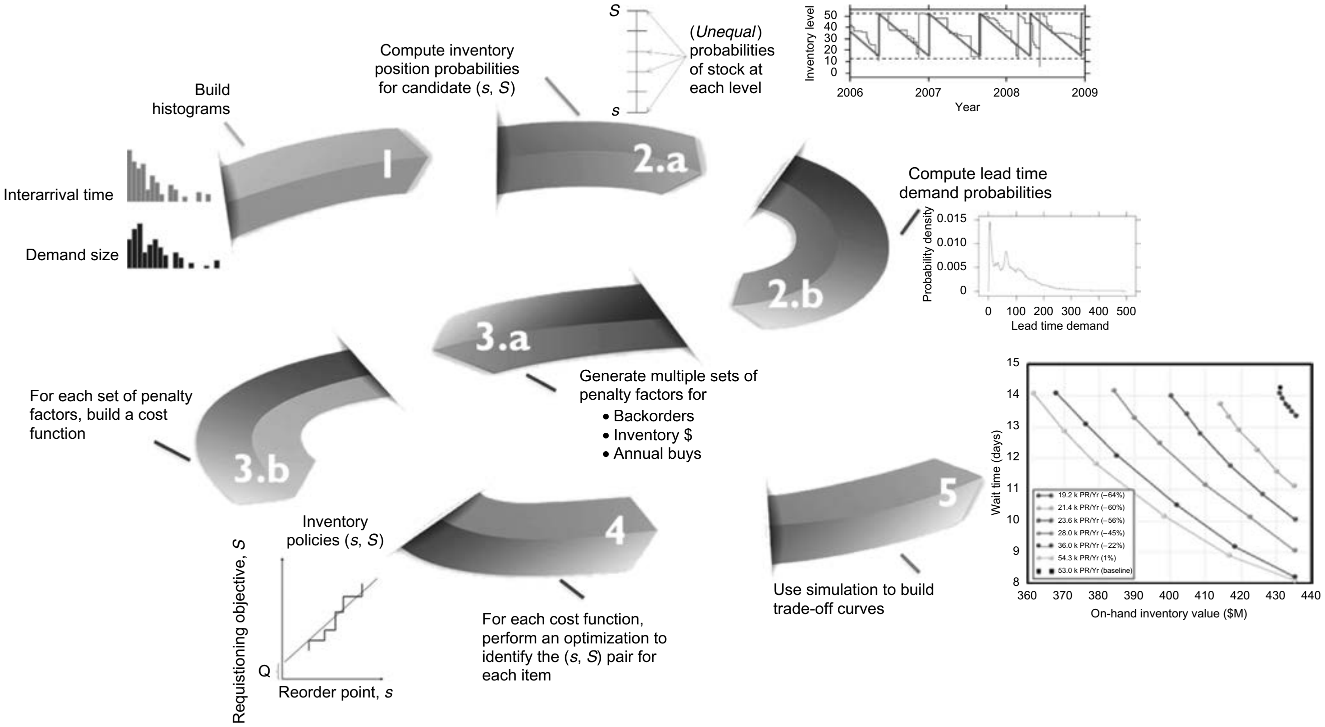
Build histograms
Interarrival time
Demand size
1
Compute inventory position probabilities for candidate (s, S)
S
s
(Unequal) probabilities of stock at each level
Inventory level
50
40
30
20
10
0
2006
2007
2008
2009
Year
2.a
2.b
Compute lead time demand probabilities
Probability density
0.015
0.010
0.005
0
0
100
200
300
400
500
Lead time demand
3.a
Generate multiple sets of penalty factors for
• Backorders
• Inventory $
• Annual buys
3.b
For each set of penalty factors, build a cost function
4
Inventory policies (s, S)
Requisitioning objective, S
Q
Reorder point, s
For each cost function, perform an optimization to identify the (s, S) pair for each item
5
Use simulation to build trade-off curves
Wait time (days)
15
14
13
12
11
10
9
8
360
370
380
390
400
410
420
430
440
On-hand inventory value ($M)
19.2 k PR/Yr (–64%)
21.4 k PR/Yr (–60%)
23.6 k PR/Yr (–56%)
28.0 k PR/Yr (–45%)
36.0 k PR/Yr (–22%)
54.3 k PR/Yr (1%)
53.0 k PR/Yr (baseline)

Figure 5: Next Gen follows five steps to go from raw data to trade-off curves for items with highly variable demand.

function does not satisfy the conditions assumed by the Z-F search (Zheng and Federgruen 1991, 1992); therefore, we approximate using an alternate cost function that does. We provide additional details in Appendix B.

At this point, we have a formula (the cost function) with values for penalty factors; however, we still need an (s, S) pair for each item that minimizes the total cost. To evaluate the cost function for a particular (s, S) pair (for every item), Next Gen calculates probabilities for backorders, on-hand inventory, and annual buys to get the expected values for the cost terms.

For each item, the Z-F algorithm performs a stair-step search within the (s, S) plane, adjusting the s and S at each step and then reevaluating the cost function. The result is one (s, S) pair for each item, which minimizes this particular cost function (Figure 5, step 4). In addition, combinations of (s, S) sets minimize the other cost functions (each with different penalty factors) that were generated. To build trade-off curves (Figure 5, step 5), Next Gen assesses each (s, S) set using LMI's financial and inventory simulation model FINISIM™, which projects wait time, fill rate, inventory value, and procurement actions. We intentionally use a simulation with spliced segments (with random starting points and lengths) of the original demand history, thus ensuring that the demand stream does not assume that the lead-time demand and inventory-position probabilities in the Next Gen model are correct. This independent assessment (against many possible

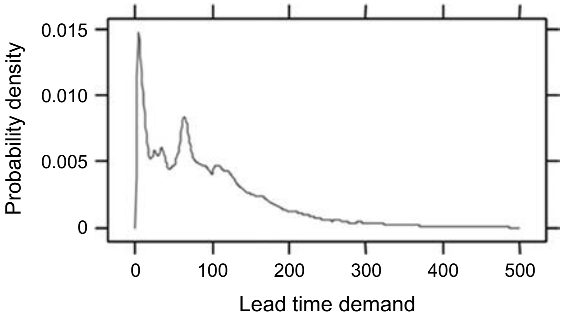

Figure 6: Next Gen can accommodate badly behaved, multimodal distributions.

futures) gives us confidence that the Next Gen trade-off curves are robust and present practical options. To our knowledge, other solutions do not assess their stock levels for robustness against demand patterns that do not meet their mathematical assumptions.

On a single CPU, and in the most computationally intensive case (tuned for large investment and few procurement actions), the average runtime is 0.62 seconds per (s, S) calculation. Tuning to a particular on-hand inventory level requires approximately 10 iterations (6.2 CPU seconds). For a 300,000-item population, this gives a total run time of 1.86 million CPU seconds (520 CPU hours) to tune to any given investment target.

Much like Peak, the Next Gen results are presented as a family of trade-off curves, parameterized by the number of annual replenishment actions generated. We do not need to make separate decisions on safety stock or order quantity, or to accept a demand plan with unknown cost implications. The operating point selected maps directly to an (s, S) pair of levels for each item.

Within Next Gen, we build histograms from the actual demand arrival process, make them adaptive (but not overly reactive), and use renewal theory to compute business metrics directly (without introducing errors from forecasting and distribution fitting). We extended the literature to target minimization of both requisition and unit backorders and to spread customer support across items, all while maintaining an efficient solution for the entire population. We also developed efficient techniques (incorporating improvements in convolution speed, rescaling the unit of issue, parallelizing the algorithm, and running on a 64-node cluster) to reduce run time by two orders of magnitude.

Next Gen differs from the few competing solutions (of which we are aware) that use empirical demand probabilities in several important ways:

- It uses empirical probabilities to build trade-off curves, thus enabling a single, integrated decision. Only a few methods use empirical distributions and a fill-rate goal, but such methods do not integrate cost, workload, and customer service.
- It predicts outcomes only at the overall population level.
- The explicit treatment of the demand arrival process, rather than starting with an empirical distribution for lead-time demand, enables (s, S) to evolve dynamically over time, without overreacting.

Development and Implementation Timeline

The development of PNG was a long but worthwhile journey. From one perspective, development followed a familiar path—from the problem definition to the model formulation phases, we tested our models against real-world data, compared them with alternatives, and then implemented a solution. From another perspective, the journey was unusual. The failure and ultimate abandonment of standard inventory control led to a years-long search for alternatives. Ultimately, finding a solution required a radical shift in perspective.

The development timeline for PNG spans 14 years (Figure 7), which is a testament to the difficulty of the problems, and the resourcefulness and determination it took to solve them.

In 1999, production at the U.S. Department of Defense maintenance depots was being delayed by shortages of repair parts, and the Office of the Secretary of Defense (OSD) asked LMI to determine the nature and cause of the parts shortages. We discovered that approximately 40 percent of the stock numbers that were delaying repairs were items with infrequent demand.

We sought to characterize the demand pattern for roughly 300,000 such items. As we moved a replenishment lead-time window through each item's demand history, we recorded how many windows captured any demand (irrespective of quantity). We found that more than 90 percent of the lead-time demand observations were zero (i.e., no observed demands).

If the usual situation for items with infrequent demand was no demand, any positive demand was an aberration. We realized that, rather than forecasting demand, we needed to protect against the risk that a demand might occur. Thus, we redefined the problem in terms of risk management.

With this new risk-management mindset, we developed inventory control algorithms that balanced protection against unpredictable demand spikes with affordable investment. Those algorithms became Peak Policy.

By 2004, our simulation studies showed that the new Peak algorithms delivered more cost-effective support for more than 20 different item populations (containing from a few hundred to more than 20,000 items), including aircraft, maritime systems, land systems, engines,

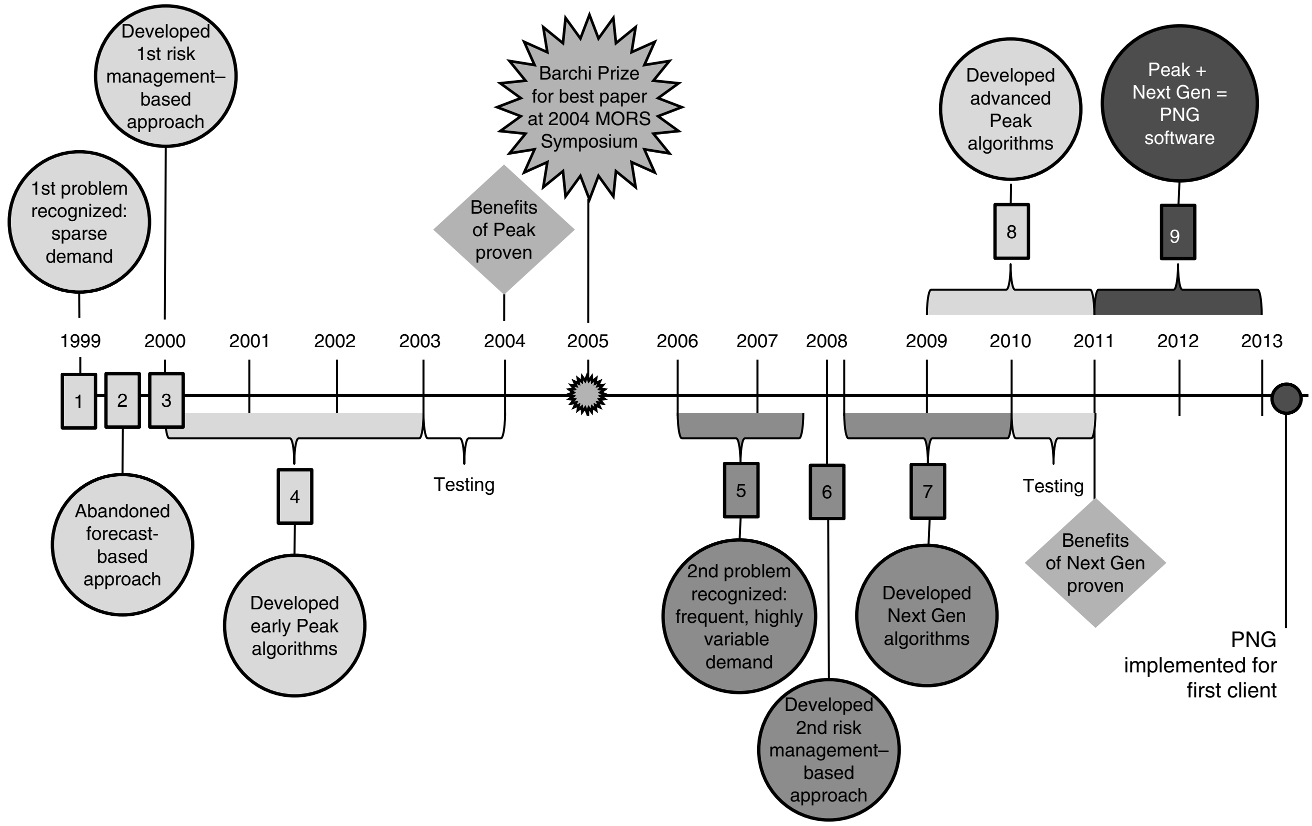
Developed 1st risk management–based approach
1st problem recognized: sparse demand
Barchi Prize for best paper at 2004 MORS Symposium
Benefits of Peak proven
Developed advanced Peak algorithms
Peak + Next Gen = PNG software
8
9
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
1
2
3
4
Testing
5
6
7
Testing
Abandoned forecast-based approach
Developed early Peak algorithms
2nd problem recognized: frequent, highly variable demand
Developed 2nd risk management–based approach
Developed Next Gen algorithms
Benefits of Next Gen proven
PNG implemented for first client

Figure 7: The development of PNG took many years. Even after the benefits of Peak were proven, it took nearly another decade to implement the solution.

and communications and electronics categories. We also showed that the benefits were reproducible across item populations and robust over time.

Could Peak also improve weapon system readiness? Replaying history with Navy and DLA data in a simulation, we showed fewer aircraft would have been out of operation for lack of parts had Peak been used, and without increasing the dollar value of inventory. We presented the methodology and results at the 2004 Military Operations Research Society (MORS) Symposium and won the Barchi Prize for best paper (Bachman 2007).

We went live with a population of aviation supply chain items in 2004. Shortly after, in 2005, DLA started an enterprise resource planning implementation, which delayed further Peak implementation.

In 2006–2007, motivated by our earlier success with Peak Policy, we decided to create a new risk-management approach for items with more frequent, but volatile demand. Because we had more data than were available for our infrequent demand case, we looked to maximize the use of that data. Our hypothesis was that, if we looked at demands over a long enough period, causal factors behind the demand uncertainty (e.g., changeable maintenance programs and uncertain failure patterns) would be reflected in the data—not perfectly, but, perhaps, well enough to achieve better business outcomes.

From 2008 to 2010, we developed Next Gen, a new model based on the raw demand process. Next Gen used five-year demand histories and advanced mathematical algorithms to develop stock levels, but it did not use a demand plan, forecast, safety stock, or separate order quantity.

In 2011, Next Gen passed rigorous simulation testing and was ready for use. Separately, we found we could improve the Peak algorithms, rendering them capable of finding better solutions more quickly. This new approach proved both effective and efficient, and made large-scale application of Peak possible (tens of thousands to hundreds of thousands of items).

Between July 2012 and January 2013, LMI coordinated with DLA's information technology services, supply chains, and headquarters to implement PNG. Changing the mindset—convincing people to move away from a forecast-based business view and consider inventory management based on a balance of risks and probabilities—was difficult. Effective communication, training, and close coordination across the enterprise were significant factors in our successful implementation.

In January 2013, DLA began using PNG to control the inventory levels for 500,000 of its most difficult-to-manage SKUs. Just two years later, metrics show that PNG is delivering impressive benefits.

Benefits

DLA supply chains set the following goals during the trade-off curve point-selection process:

- Achieve 90 percent fill rate (in accordance with the DLA director's operations order).
- Target a significant reduction in procurement request (PR) workload.
- Improve both wait time and on-hand inventory value.

Just over two years after its implementation, PNG helped DLA make considerable progress toward these goals. (Although wait time was initially one of the three goals established by the supply chains, fill rate eventually became the focus metric for customer service and DLA placed less attention on wait time.)

Increased fill rate: Greater fill rate translates into better support and higher weapon system readiness. For items managed by PNG, fill rate—the percentage of orders filled completely on the same day they are received—increased four percentage points, without increasing on-hand inventory value. The item population under Peak (items with sporadic demand) improved by 10 percentage points, from 72 to 82 percent. The Next Gen items (items with frequent, but highly variable demand) improved by four percentage points, from 85 to 89 percent.

Reduced PR workload: Before implementation, DLA generated 369,000 PRs each year for the items the agency selected for PNG management. PRs are contractual documents that describe the material to be obtained. PNG reduced the number of PRs by 35 percent to 239,000. If we assume that one-third of the PRs are processed manually (a conservative estimate) and use a DLA-provided labor cost of $400 per request, DLA saves nearly $18 million per year—$90 million in savings over five years.

Reduced PR cancellations: Cancelled PRs are down 70 percent, from 145,000 to 45,000 per year. Using a DLA-provided labor cost of $75 to cancel a PR, this saves DLA $7.5 million per year—$37.5 million over five years.

Improved dollar value of on-hand inventory: On-hand inventory value decreased from $3.3 billion in January 2013 to $2.7 billion in October 2014. The reduction was achieved mostly through an aggressive disposal policy; however, because PNG ensures that DLA continues to buy more of what sells and less of what does not sell, PNG contributes to the inventory reduction. More importantly, with PNG, DLA will avoid a new buildup of excess inventory after disposal.

More efficient use of working capital: Before PNG, that is, in calendar year 2012 (CY12) and prior years, dollars invested in replenishment stock were running much higher than revenue generated from sales. Figure 8 shows a steady decline in the ratio of procurement dollars to sales dollars. (Over the same period, this ratio actually increased for the population of items

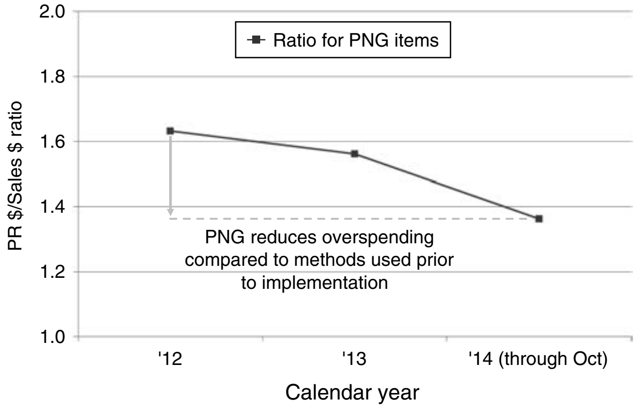

Figure 8: PNG reduces overinvestments resulting from volatile forecasts.

not assigned to PNG—their overspending continued to rise.) With PNG, DLA is using its working capital more efficiently and seeing a reduction in overspending. By reducing the gap between dollars spent on procurements and dollars earned on sales, DLA will save $400 million annually.

Improved warfighter support: The greatest benefit is occurring beyond DLA. Having the right items on shelves directly correlates to an increase in fill rate for mission-critical items, which means greater availability of weapon systems and warfighter readiness.

Organizational Impact

The DLA and LMI teams continue to meet to discuss implementation activities and monitor metrics and progress. The supply chains actively engage in the annual trade-off curve process, reviewing achievements from the previous year and setting goals for the next. DLA's supply chains are shifting their thinking and better managing their inventories by balancing risks, rather than forecasting demand.

Lessons Learned

From an operations research perspective, what have we learned?

- Defining the right problem is critical. The problem was not about improving forecasts; it was about buying the right quantity at the right time.
- When modifying any of the standard methods fails to solve the problem, seeking a fresh approach is appropriate.
- Persistence pays off. It took 14 years, from initial search to solution implementation. The last five years were spent winning support for the PNG approach.
- A close partnership between operations research practitioners and the client organization is a key to success. Funding, unwavering support, and review of results by members of the DLA research and development program were critical elements in PNG's success.

Additional Applications

Several other industries may benefit from our approach, including automotive, commercial aircraft, marine hardware, apparel, consumer goods, food and beverage, manufacturing, oil and gas, pharmaceuticals, and medical supplies. The high-technology and electronics industries, and those that provide service and aftermarket parts, would also benefit from PNG.

Following PNG's successful implementation at DLA, OSD has initiated a project to extend the software to reparable items managed by the military services.

Appendix A. Peak Policy

Let m_i and q_i equal the number of possible values for multipliers (M) and order quantities (Q) for price group i. For example, price group $i = 1$ might have six possible multiplier values (i.e., $m_i = 6$) and four possible order quantity values (i.e., $q_i = 4$), which results in 24 different candidate (M, Q) pairs for price group $i = 1$. There are N price groups and K candidate multiplier and order quantity pairs per price group: $K = m_i \times q_i$.

The problem of finding Peak Policy parameters (multipliers M and order quantities Q by price group) to meet a given set of trade-off goals is formulated as an MIP, where simulation outputs for wait time, on-hand inventory value, and PRs are key input parameters for the problem. The baseline stocking policy refers to the metrics achieved by the business prior to conducting the trade-off analysis using the MIP.

Parameters

$i =$ price group, $(i = 1, \ldots, N)$;
$j =$ candidate (M, Q) pair for price group i, $(j = 1, \ldots, K)$;
$OH_{ij} =$ average on-hand inventory value for parameter pairing j of price group i;
$OH_{\text{base}} =$ average on-hand inventory value for the baseline stocking policy;
$OH_{\text{scale}} =$ scaling factor for average on-hand inventory constraint (this is the percentage of OH_{base} that is permitted for the average on-hand inventory resulting from the MIP);
$PR_{ij} =$ procurement workload for parameter pairing j of price group i;
$PR_{\text{base}} =$ average procurement workload for baseline stocking policy;
$PR_{\text{scale}} =$ scaling factor for procurement workload constraint (this is the percentage of PR_{base} that is permitted for the procurement workload resulting from the MIP);
$WR_{ij} =$ average wait time for parameter pairing j of price group i;
$WR_{\text{base}} =$ average wait time for the baseline stocking policy;
$WR_{\text{scale}} =$ scaling factor for average wait time constraint (this is the percentage of WR_{base} that is permitted for the waiting time resulting from the MIP).

The parameters OH_{ij}, PR_{ij}, and WR_{ij} are the outputs of a simulation model using Peak Policy.

Decision Variables

$X_{ij} = 1$: When (M, Q) candidate pair j from price group i is used; otherwise $= 0$.

Objective function. The business outcome metrics are on-hand inventory, wait time, and procurement workload. Any of these three outcome metrics can be minimized and the others constrained. To illustrate, we present the on-hand inventory value objective function as $\sum_{i=1}^{N}\sum_{j=1}^{K} OH_{ij}\, X_{ij}$.

Constraints. The procurement workload projected by Peak Policy should not exceed some percentage of the procurement workload associated with the baseline stocking policy. If procurement workload is not an issue, PR_{scale} may be greater than 1. We have

$$\sum_{i=1}^{N}\sum_{j=1}^{K} PR_{ij}\, X_{ij} \le PR_{\text{scale}}\, PR_{\text{base}}.$$

The wait time projected by Peak Policy should not exceed some percentage of the wait time associated with the baseline stocking policy. We have

$$\sum_{i=1}^{N}\sum_{j=1}^{K} WR_{ij}\, X_{ij} \le WR_{\text{scale}}\, WR_{\text{base}}.$$

There is, at most, one Peak multiplier and Peak order quantity per price group i. We have

$$\sum_{j=1}^{K} X_{ij} = 1, \quad \forall\, i.$$

Binary Constraints. The decision variables are binary, such that

$$X_{ij} \in \{0, 1\}, \quad \forall\, i, j.$$

Appendix B. Next Generation Inventory Model

The Next Gen model computes (s, S) levels for items in a population for use with either a continuous-review or periodic-review ordering doctrine. Next Gen consists of two parts: an analytical optimization model and a stochastic simulation. The optimization model computes (s, S) levels for items, and the simulation model independently assesses (s, S) levels in terms of business metrics for customer service, inventory value, and replenishment workload. This appendix describes the optimization model.

Optimization Algorithm

Zheng and Federgruen (1991), using an abstract objective function with specified properties, showed how to efficiently solve for (s, S) policies when a positive cost to replenish exists and shortages are backordered. This was a major breakthrough, because it had been known since the 1950s (Arrow et al. 1958) that the objective function was badly behaved, with multiple local minimums. Search algorithms prior to Z-F (Veinott and Wagner 1965), although faster than an exhaustive search, were not computationally efficient enough to be practical. Recent books (Axsäter 2006, Zipkin 2000) cite Z-F as the best relevant algorithm. To our knowledge, we are the first to apply the Z-F algorithm to a practical problem; however, the Zheng and Federgruen (1991) paper did not express the objective function in terms of observable data. We needed to construct an objective function from observable data and, if possible, show it satisfied the Z-F conditions.

The population's objective function, whose solution is a set of (s, S) levels for every item, is the sum of the items' objective functions. The item problems are linked by a set of common penalty factors for backorders, carrying inventory, and replenishment actions. Therefore, it suffices to work with an item's objective function.

Let $c(s, S)$ be the expected cost per unit time, for a single item, in equilibrium. We have

$$\begin{aligned} c(s, S) &= \frac{1}{M(\Delta)}\left\{\frac{K}{\mu_a} + \sum_{j=0}^{\Delta-1} m(j)G(S-j)\right\} \\ &= \frac{1}{M(S-s)}\left\{\frac{K}{\mu_a} + \sum_{j=0}^{S-s-1} m(j)G(S-j)\right\}, \end{aligned}$$

where

$\Delta = S - s$

$M(j) =$ expected time until the next order is placed, giving starting inventory position of $s + j$, and

$G(S-j) =$ instantaneous back-ordering and holding cost, and a lead time after inventory position is at $S - j$.

Sufficient conditions on $G(y)$ for the Z-F algorithm to work are

1. $-G(y)$ is unimodal—$G(y)$ is monotone decreasing to the left of an interval (possibly a single point), where it is constant and attains its minimum value, and monotone increasing to the right of that interval;
2. $\lim_{|y|\to\infty} G(y) > \min_y G(y) + K$, where $K > 0$ is the fixed cost to place an order;
3. $c(s, S)$ and $G(y)$ satisfy the relationship

$$c(s-1, S) = \alpha_\Delta c(s, S) + (1 - \alpha_\Delta)G(s),$$

$$\text{where } \alpha_\Delta = \frac{M(\Delta)}{M(\Delta+1)}.$$

Construction of Basic Objective Function

Sahin (1979) developed an objective function for a continuous review (s, S) policy based on arbitrary demand size and interarrival time distributions. Because our customer had histories of transactional demands (stock number, date, quantity), we could build histogram distributions for demand interarrival times and demand sizes, and use those to construct the objective function.

Notation

$c(s, S) =$ expected cost per unit time for reorder point, s, and requisitioning objective, S, $\Delta = S - s$;

$A(t) =$ CDF for requisition interarrival times;

$A_n(t) =$ n-fold convolution of $A(t)$, $n > 0$; $A_0(t) = 1$;
$b(k) =$ density for requisition sizes, assumed discrete;
$B(k) =$ CDF of requisition sizes;
$b_n(k) =$ n-fold convolution of $b(k)$;
$B_n(k) =$ n-fold convolution of $B(k)$; $B_0(t) = 1$;
$g(y) =$ probability density for lead-time demand;
$G(y) =$ short-term back-ordering and holding cost with inventory position y;
$H(y) =$ short-term holding cost with inventory position y;
$P(y) =$ short-term back-order cost with inventory position y;
$h =$ unit holding cost;
$K =$ administrative cost to place a replenishment order;
$\mu_n =$ mean interarrival time;
$p =$ back-order penalty;
$R(x) =$ renewal function for requisition sizes for a total quantity demanded x; $R(x) = 0$;
$r(x) =$ renewal density of requisition sizes for a total quantity demanded x; $r(0) = 1$.

The state probabilities of the inventory position, given (s, S), are then

$$\Pr(IP = j) = \frac{r(S-j)}{1+R(S-s-1)}, \quad j < S \quad \text{and}$$

$$\Pr(IP = S) = \frac{1}{1+R(S-s-1)}.$$

The probability distribution for quantity demanded in a replenishment lead time is

$$\begin{aligned} g(j) &= \Pr(d(t-L,t)=j) \\ &= \sum_{k=1}^{\infty} \Pr(j \text{ units in } k \text{ arrivals}) \Pr(k \text{ arrivals}) \\ &= \sum_{k=1}^{\infty} b_k(j) \cdot \frac{1}{\mu_a} \cdot \int_{u=0}^{L} \underbrace{[1-A(u)]}_{\Pr(\text{no demand from } 0 \text{ to } u)} \\ &\qquad \cdot \underbrace{[A_{k-1}(L-u) - A_k(L-u)]}_{\Pr(k \text{ arrivals from } u \text{ to } L)} \, du. \end{aligned}$$

Note that $g(0) = 1/\mu_\alpha \int_0^\infty [1 - A(u+L)]du$.

The expected number of replenishment cycles per unit time is

$$\frac{1}{\mu_a(1+R(S-s))}.$$

Following Sahin, the expected cost of replenishment, carrying inventory, and of outstanding backorders, in equilibrium, is then

$$\begin{aligned} c(s,S) &= \left(\frac{K}{\mu_a} + \sum_{j=0}^{\Delta-1} r(j) \left[h \sum_{k=1}^{S} kg(S-j-k) \right.\right. \\ &\qquad \left.\left. - p \sum_{k=-1}^{-\infty} kg(S-j-k) \right]\right) \cdot (1+R(\Delta-1))^{-1} \\ &= \frac{1}{1+R(\Delta-1)} \left\{ \frac{K}{\mu_a} + \sum_{j=0}^{\Delta-1} r(j)[H(S-j)+P(S-j)] \right\} \\ &= \frac{1}{1+R(\Delta-1)} \left\{ \frac{K}{\mu_a} + \sum_{j=0}^{\Delta-1} r(j)G(S-j) \right\}. \end{aligned}$$

All the terms in this objective function are computable from observable data. We showed that the instantaneous cost function G satisfies the Z-F conditions, enabling us to solve for (s, S) using the Z-F algorithm with our objective function.

Modified Instantaneous Cost of Backorders and Holding Inventory

Simulation experiments against randomized historical demand revealed that the basic Next Gen objective function did not provide the desired support for DLA's mission because

1. unit measures (i.e., expected units back-ordered) are dominated by the service level for large requisitions (not appropriate for DLA); and
2. minimization of the cost function can be achieved by focusing on high levels of customer service for a few items at the expense of the others. (DLA's customers require a high level of support for a wide range of items.)

To remedy this situation, we modified $G(y)$ by adding two terms to the cost function, so that it becomes

$$G(y) = H(y) + H_{IP}(y) + P_u(y) + P_r(y),$$

where $H_{IP}(y)$ is the cost to carry on-hand plus on-order inventory, $P_u(y)$ is the cost of outstanding unit backorders, and $P_r(y)$ is the cost of outstanding back-ordered requisitions (irrespective of quantity).

Cost terms for both inventory position (IP) and on-hand inventory overlap, but the addition of the IP perturbation term enabled us to address problems (1) and (2) in the objective function above, while keeping G as a function of y alone (for better mathematical tractability).

The modified $G(y)$ did not satisfy the Z-F conditions; therefore, we developed a unimodal approximation $\hat{G}(y)$.

The cost function must be unimodal for the optimization procedure to avoid being caught in a nonoptimal local minimum. We made a simple unimodal approximation to the conditional cost function by flattening out the valleys corresponding to local minima before building the full cost function. Given a possibly nonunimodal function f defined on $\{0, 1, 2, \ldots, N\}$, achieving a global minimum on $[x_m, x_i]$, we use the unimodal approximation f_u defined on $\{0, 1, 2, \ldots, N\}$ by

$$f_u(0) = f(0)$$

$$f_u(N) = f(N)$$

$$f_u(x) = \begin{cases} f(x), & \text{if } f(x) \le f(x-1), \\ f(x-1), & \text{otherwise,} \end{cases} \quad \text{for } x \le x_m;$$

$$f_u(x) = \begin{cases} f(x), & \text{if } f(x) \ge f(x-1), \\ f(x-1), & \text{otherwise,} \end{cases} \quad \text{for } x > x_i.$$

Finally, we work with the new cost function $c(s, S)$:

$$c(s,S) = \frac{1}{1+R(\Delta-1)}\left\{\frac{K}{\mu_a} + \sum_{j=0}^{\Delta-1} r(j)\hat{G}(S-j)\right\}.$$

The cost function for a population of N items is simply $C(s,S) = \sum_{i=1}^{N} c(s_i, S_i)$.

Acknowledgments

The DLA Weapon System Sustainment Program Research and Development Program made the development of Peak Policy and Next Gen possible. Cliff Lanphier, Jeffrey Pinnick, and Tammy Sabo led the DLA planning implementation team. Representatives from DLA Information Operations were instrumental in creating and maintaining the data transmission interface. We also acknowledge the business process analysts at DLA Aviation, Land, Maritime, and Troop Support for the training provided to supply analysts and the contracting shop. Carol DeZwarte was instrumental in the early success of Peak Policy. Ericson Davis formulated Peak Automation as a mixed-integer linear programming problem and coded the algorithm. Joel Lepak, Karl Kruse, and John Westbrook were key contributors to the development of the Next Gen algorithm. We also thank our coaches, Ken Fordyce of Arkieva and Ananth Iyer of Purdue University, for their helpful suggestions on improving this article.

References

Arrow KJ, Karlin S, Scarf H (1958) *Studies in the Mathematical Theory of Inventory and Production* (Stanford University Press, Redwood City, CA).

Axsäter S (2006) *Inventory Control*, 2nd ed. (Springer, New York).

Bachman TC (2007) Reducing aircraft down for lack of parts with sporadic demand. *Military Oper. Res.* 12(2):39–53.

Bachman TC, DeZwarte CA (2007) Method of determining inventory levels. U.S. Patent 8,165,914, filed May 18, issued April 24, 2012.

Bachman TC, Kruse K, Lepak J, Westbrook J (2011) Method and computer system for setting inventory control levels from demand inter-arrival time, demand size statistics. U.S. Patent 8,600,843, filed September 19, issued December 3, 2013.

Croston JD (1972) Forecasting stock control for intermittent demands. *Oper. Res. Quart.* 23(3):289–303.

Fricker RD, Goodhart CA (2000) Applying a bootstrap approach for setting reorder points in military supply systems. *Naval Res. Logist.* 47(6):459–478.

Ross SM (1992) *Applied Probability Models with Optimization Applications* (Dover, Mineola, NY).

Sahin I (1979) On the stationary analysis of continuous review (s, S) inventory systems with constant lead times. *Oper. Res.* 27(4):717–729.

Sahin I (1982) On the objective function behavior in (s, S) inventory models. *Oper. Res.* 30(4):709–724.

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

Veinott AF Jr, Wagner HM (1965) Computing optimal (s, S) inventory policies. *Management Sci.* 11(5):525–552.

Zheng Y-S, Federgruen A (1991) Finding optimal (s, S) policies is about as simple as evaluating a single policy. *Oper. Res.* 39(4):654–665.

Zheng Y-S, Federgruen A (1992) Corrections to "Finding optimal (s, S) policies is about as simple as evaluating a single policy." *Oper. Res.* 40(1):192.

Zipkin PH (2000) *Foundations of Inventory Management* (McGraw-Hill, New York).

Tovey C. Bachman has worked in operations research and logistics for more than 25 years, leading breakthroughs in modeling, inventory analysis, and space logistics. His research focus with LMI has been on mathematical modeling and analysis of complex systems. He invented Peak Policy for managing inventories of items with infrequent demand and was the lead developer of the Next Generation Inventory Model for managing items with frequent but highly variable demand. He holds two patents for the Peak and Next Gen inventory management suite. Tovey earned his PhD in mathematics from the University of Michigan in Ann Arbor and his bachelor's in mathematics from Reed College in Portland, Oregon.

Pamela J. Williams is a senior consultant at LMI and is the project lead for the PNG implementation at DLA. Her research interests include inventory control, machine learning, and nonlinear optimization. She earned her MS and PhD in computational and applied mathematics from Rice University and a BS in mathematics from the University of Kentucky.

Kristen M. Cheman is a consultant at LMI. She earned her MS in applied mathematics from North Carolina State University, an ME in systems engineering from the University of Virginia, and a BS in mathematics from Allegheny College.

Jeffrey Curtis is the executive director, Logistics Support Directorate (J34) for DLA Logistics Operations. He has served in a range of positions with the federal government, including as the executive director for DLA's Materiel Policy, Process, and Assessment Directorate, where he managed logistics policy, plans, programs, and operations for all supply classes managed by DLA. Jeffrey graduated from The Ohio State University with a bachelor's degree in business administration and a double major in logistics and marketing. He graduated from the Federal Executive Institute's Leadership for a Democratic Society program and is an honor graduate of the Army's Operations Research and Systems Analysis Military Applications Course. He is a member of the Defense Acquisition Corps and is DAWIA Level III Certified in the contracting field.

Robert Carroll has worked in logistics in the Department of Defense for almost 30 years. He is a logistics specialist in the Office of the Assistant Secretary of Defense for Logistics and Materiel Readiness, where his focus is on supply chain integration. Most of his career was at the Defense Logistics Agency, working at DLA Headquarters in Fort Belvoir, Virginia; DLA Troop Support Supply Center in Philadelphia, Pennsylvania; and DLA Aviation Supply Center in Richmond, Virginia. Robert holds a bachelor of business administration degree in human resources from the University of New Mexico and a master's of public administration with a concentration on procurement in the public sector from Pennsylvania State University.