Smarter Supply Chains through AI

DUNCAN KLETT

PREVIEW *In this article, Duncan Klett portrays a supply chain as a multifaceted control system in which a reduction in latency (delays) offers the most leverage for improved supply-chain performance. He sees a broad role for AI/ML in furthering this goal, through automating processes, validating data, segmenting items, and generating forecasts.*

INTRODUCTION

Every demand forecast has its flaws. Newer methods, such as AI and ML, can provide better forecasts, but even these will never be perfect. There is always some difference between predicted sales and actual customer demand. Still, an imperfect demand forecast is better than no forecast at all.

If we accept that a demand forecast is just a road map of future demand, then our goal should not be only to minimize forecast error but to improve supply-chain execution—what you do with the forecast matters! The more important metrics relate to supply-chain execution, such as projected and actual inventory, on-time delivery in-full, and counts of stock-outs. Put another way, if the "plan" is not executed through operations, then the forecast is of little value. Alternatively, supply-chain design, such as structure of the bill of materials or inventory policies, can exaggerate or mitigate the impact of inaccurate forecasts.

Supply-chain execution can be improved by recognizing that supply-chain management is really a control system, similar in concept to an automobile's cruise control. A principle of control-system theory is that delays in a control process can lead to instability.

Execution can also be improved by applying different strategies and methodologies to different segments of products, markets, and customers. Various forecasting methods, inventory strategies, and replenishment strategies can be tuned to different products and markets. In addition, AI/ML techniques are beginning to be used to manage supply chains. Applying the right strategies to the right product segments and generally reducing latency can enable organizations to respond to supply-chain variability more effectively to achieve business success.

SUPPLY-CHAIN CHALLENGES

Supply-chain management is really common sense: get the right stuff to the right place at the right time. The goals are almost always variations of

- Assuring customers on-time delivery
- Minimizing costs in purchase, inventory, scrap/waste, and overhead.

There are really only two types of output from supply-chain planning:

- Replenishing supply (consisting of source selection, transportation routing, replenishment quantity, and releasing orders); and
- Allocating resources (including materials and capacity) to meet demands.

In addition, you can manage various parameters and strategies, such as

- Inventory
- Demand "shaping"
- Product offerings, setting expectations, product design.

However, reality creeps in to add innumerable complexities to these simple concepts. For a single supply chain, there can be thousands or millions of parts to manage, perhaps in hundreds of different locations (or even thousands, if managing out to retail locations). Products become more and more complex in both function and manufacturing complexity. In many cases, the actual manufacturing processes required to produce a product are at the outer limits of current technology, resulting in highly variable yields and lead times.

A sense of the difficulty of managing a supply chain becomes immediately evident through a simple calculation. Consider an assembly that requires 100 different component parts. If each individual component is available 99% of the time, then the likelihood of having all 100 components available at the same time is 0.99 ^ 100 = 0.37 or 37%. Not a very good customer-service level!

In addition to product complexity, customer expectations have become more demanding and more difficult to meet. Customers often expect close to immediate delivery of "personalized" products, sometimes 12 time zones away from the actual factory! It is becoming increasingly difficult to meet these expectations, while also minimizing costs.

Product complexity and customer expectations are not even the end of the story. Supply chains operate globally and so are subject to disruption from natural and human-created events. Storms can disrupt transportation, tariffs can change cost structures and times for materials to cross borders, and global supply and demand for commodities change frequently.

In the end, supply-chain management is like any other business process in that it relies on a combination of people, process, and technology. Any change in the capabilities of one of these elements will often present risks and opportunities in overall system performance through corresponding changes in either or both of the other two.

CONTROL-SYSTEM THEORY FOR SUPPLY-CHAIN MANAGEMENT

Figure 1 is a simplified representation of a classic control system in which feedback from errors in achieving a target leads to adjustments in the production process.

There are many sources of variability (change) in a supply chain:

- Lot sizes (order quantity minimums and multiples). Mismatches between typical demand quantities and replenishment lot sizes can be particularly problematic.
- Demand fluctuations and misalignment of actual and forecasted demand.
- Supply fluctuations such as actual lead times and yields.

However, perhaps the largest and easiest-to-manage factor is delay in responding to changes in supply or demand.

Key Points

- Supply-chain management is a complex control system in which latency (delay) in the control can lead to instability. While there are many other sources of instability in a supply chain, delay in responding to changes in supply or demand is perhaps the largest and easiest to manage.
- Traditional supply-chain planning systems consist of several unconnected processes such as Sales and Operations Planning (S&OP), Master Scheduling, Demand Planning, Order Promising, Supply Planning, Production Planning, Production Scheduling, Material Requirements Planning (MRP), and more. Each of these operates in isolation, adding further delay to supply-chain response.
- Removing control-system latency improves a system's responsiveness and reduces its instability. A digital model of the supply chain can significantly improve supply-chain responsiveness and reduce or eliminate the bullwhip effect. A companion digital model can provide higher on-time delivery (increased customer satisfaction) and lower inventory requirements.
- The areas of potential improvement from AI/ML lie in automating processes, validating and correcting data, generating forecasts, segmenting items and customers for appropriate management strategies, managing changes, recommending actions, and mitigating risks.

Figure 1. Classic Control System With Feedback

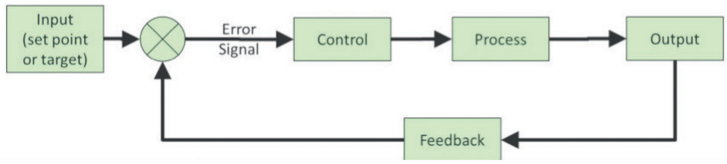
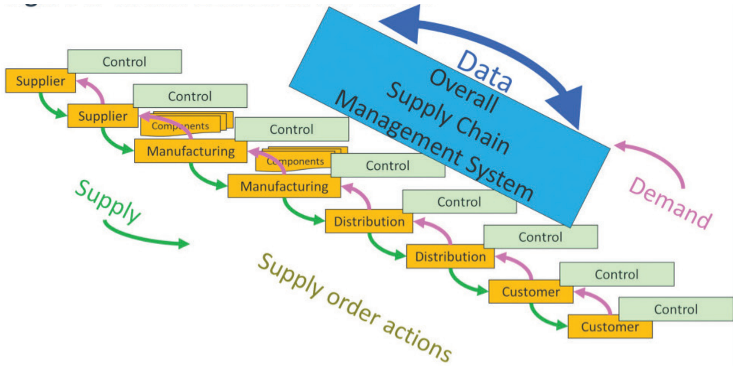


Figure 2. Multi-Level Supply Chain



If we think of supply-chain management processes as a complex control system, the impact of delay (latency of response) becomes very clear. Control systems are commonly understood and used extensively in engineering (***https://en.wikipedia.org/wiki/Controlsystem***). Their objective is to adjust various controls so that the actual process perfectly tracks to the desired target. Furthermore, the process should instantaneously track to changes in either the target setting or in the operation of the process. For example, an automobile's cruise control represents a simple control system that consists of a target (desired speed), a control (throttle setting and, more recently, braking), a process (the vehicle), and an output (actual vehicle speed). The difference between the target speed and the actual speed is an error signal which is used as a feedback loop to adjust the throttle setting.

Clearly, modern cruise controls show the effectiveness of such control systems.

Delay and the Instability Problem

Instability is a common problem with multiple levels in a supply chain, a result often attributed to the bullwhip effect ***https://en.wikipedia.org/wiki/Bullwhipeffect***).

Control-system theory reveals that delays in the control-to-feedback path lead to instability, and that instability can only be reduced by reducing the magnitude of the feedback signal. Such reduction is often called "damping." Unfortunately, reducing the feedback signal makes the system less responsive. The theory also shows that the system overreacts (bullwhip) unless the damping factor is less than 1/delay.

Supply-chain delays are introduced in three primary ways:

- Time for "system" to react to a change in supply or demand for a part.
- Time for "control" to react to a change in supply or demand for a part.
- Time to propagate a change at one level to the next (e.g., through the product structure, between supply-chain systems, between different entities in a complex supply chain).

Multilevel Supply Chains

Figure 2 shows a simplified view of several levels in a typical supply chain.

Each level in the supply chain adds to the time it takes for the entire supply chain to react to a change. As discussed above, each extra delay contributes to instability.

Disjointed Supply Planning

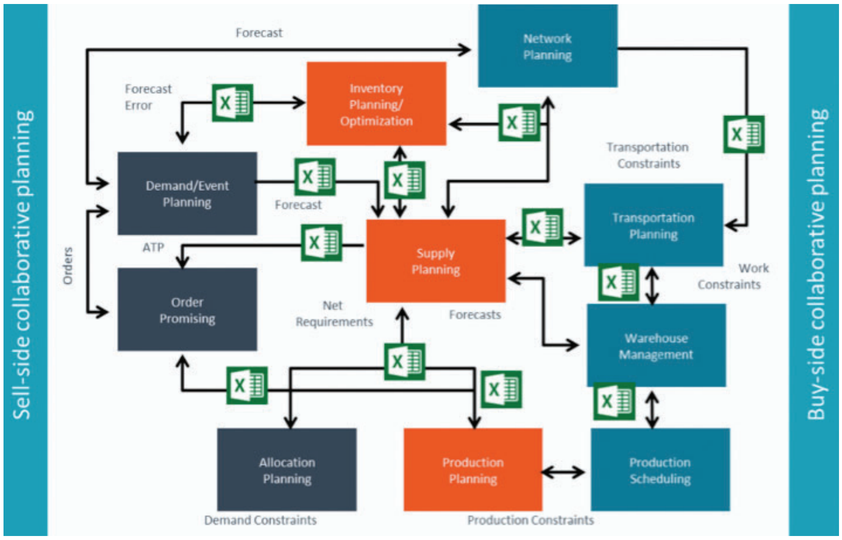
As illustrated in **Figure 3**, traditional supply-chain planning systems consist of several unconnected processes such as Sales and Operations Planning (S&OP), Master Scheduling, Demand Planning, Order Promising, Supply Planning, Production Planning, Production Scheduling, Material Requirements Planning (MRP), and more. Each of these operates in isolation. Data from one process is then passed to another process, often manually, using Excel or another tool such as a data warehouse. Each interface adds further delay to supply-chain response.

A Digital Model to Overcome Delays

An extension to control-system theory, designed specifically to overcome delay (latency) in system response, is named a Smith Predictor Model (***https://www.controleng.com/single-article/process-modeling-feedback-controllers/8d97811eab2a09c44f1be0c03e5d1b55.html***). This approach, as illustrated in **Figure 4**, uses a model of the physical process to predict its output prior to the actual process response. The process-model prediction is then combined with the actual system response to become the feedback input to control system logic. In terms of a supply chain, a digital model of the supply chain can be used to predict supply-chain response at all levels and through all elements at once, thereby removing the latency.

Removing control-system latency improves a system's responsiveness and reduces its instability. Therefore, a digital model of the supply chain can significantly improve supply-chain responsiveness and reduce or eliminate the bullwhip effect. A companion digital model can provide higher on-time delivery (increased customer satisfaction) and lower inventory requirements.

Figure 3. Disjointed Supply-Chain Management



APPLYING AI/ML

So how should we apply AI/ML to achieve improvements? For this discussion, I define AI as any system where a machine replaces or augments human activity. Even ML (machine learning) should be broadly defined to include any machine that gets better (learns) as it processes more data.

Several areas for potential improvement from AI/ML lie in Automating processes:

- Validating and correcting data
- Generating forecasts (including forecasts other than of demand)
- Segmenting items and customers for appropriate management strategies.

Remember, though, that the most important input to any system that uses data, including AI, is clearly the data. Therefore, processes must be in place and used to achieve as accurate and timely data as practical.

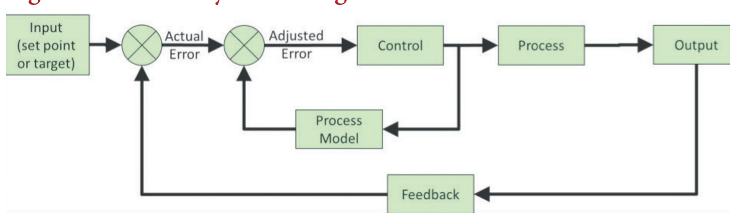
Automating Processes

Automation occurs when a specific action (which can include "do nothing") is triggered based upon observed data and some combination of calculations or rules. This requires that a programmer, user, or learning algorithm recognizes a pattern and then triggers an action. For example, when the projected inventory at a distribution center (DC) for a set of parts falls below a specified level and if the supplying site has sufficient inventory to satisfy this demand, a stock transfer order (STO) is created automatically to prevent the below-target situation. The STO would have an associated due date, transfer quantity, and other information.

Material Requirements Planning (MRP) is actually a form of automation that has been in general use since the mid-1960s. MRP recommends replenishment orders based on projected demand, current supply, and new planned supply.

An extension of process automation would use machine learning (ML) to recommend or perform actions based on its observations of human activity and the ensuing results. A machine could be trained to learn which actions are most successful in improving the results from various situations. Think of computers already beating human masters in chess and Go. Similar learning approaches might be applicable to managing supply chains.

Figure 4. Control System Using Process-Model



In addition, AI/ML approaches might even be better used in setting planning parameters (rules and data involved in source selection, lot sizes, transport routes, etc.) than in managing the individual transactions. Given more effective planning parameters, standard processes could then manage most of the transactions.

Demand Sensing

Statistical models for forecasting demand rely on historical data to predict future demand. The past, however, is not always a good predictor of the future. In those cases, the dream is to find other data sources—leading indicators—that are better predictors of future demand. Some such data, such as national holidays and recognizing which days fall on the weekend, are known well in advance. Other data such as weather and weather forecasts, product promotions, and internet search trends can also be included in the calculations (or learned response) to generate better demand forecasts.

In one case, the company's challenge was to manage daily delivery of product to numerous retail stores, such that inventory was minimized while always having product available to sell. Learning historical trends with holiday information, the simulated result reduced stock-outs from five days to one, and did this using 10% less inventory.

However, our experience has shown that the "smarter" (i.e., more complicated) an algorithm, the more resistance there is from users to actually employ it. Simply put, if they don't understand it, they don't trust it, and they won't use it! To overcome this resistance, we have utilized processes that allow users to set thresholds as to when a recommendation is accepted. Visualization techniques such as Shapely analysis (***https://arxiv.org/abs/1705.07874***) can also help users understand the significance of the different data (features) as they affect a particular prediction.

In **Figure 5**, "until holiday" and "weekday" add to the initial forecast quantity for a specific period while "month," "week," and "since holidays" reduce it. Thus, the revised forecast is 17.26 compared to the base forecast of 20.89.

Remember, though, that a better forecast does not ensure better supply-chain results. The supply chain must still execute in order to get the right stuff to the right place to satisfy the actual demand.

Beyond Demand Forecasting

Most discussions of forecasting for supply chains deal with demand. A demand forecast is important, but so are other factors such as production times, shipping times, and yields. Models and techniques for forecasting these can be quite different from those for forecasting demand. For example, demand data is often aggregated by adding the observed quantities in a period to generate a single historical number for that period. Clearly, shipping times for receipts on the same day should not be added together when predicting future lead times. Further, for predicting yields there is a physical maximum of 100%.

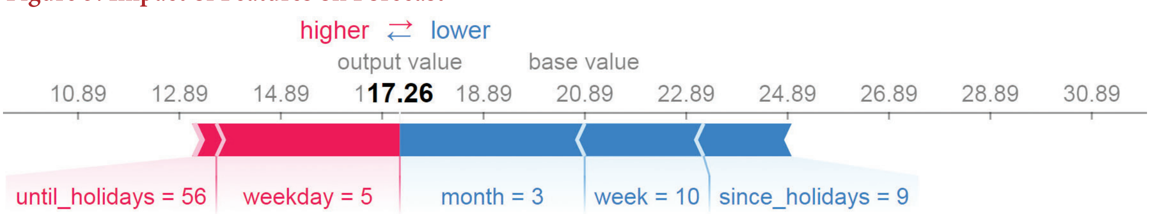
For predicting future values for some of these factors, we have found it is better not to aggregate the historical data at all. Each observation is then used for learning and subsequent predictions of future values.

Segmentation

Reiterating a comment above, supply-chain management does just two things: replenishment and allocation. Behind these two functions, though, are a myriad of supporting operations, such as

- Policies (including inventory targets, replenishment)

Figure 5. Impact of Features on Forecast



• Supply order management
• Demand prioritization and sequencing
• Demand forecasting
• Sales promotions.

There is no single approach to supply-chain management that will provide good results across all products and customers. The Pareto Principle (also known as the 80/20 rule) is well known. A lesser-known extension of that principle is the Glenday Sieve (***www.repetitiveflexiblesupply.com***). Glenday's research has found that while 20% of a company's products represents 80% of its business, just 6% of the products typically represents 50% of its business. At the other extreme, the last 30% of products represents just 1% of its business. Therefore, it makes sense that strategies for production and inventory for the top 6% would be different than those for the bottom 30%.

For example, for high-runner products, even a few days of inventory might represent a significant investment, so replenishment should occur frequently. But for these parts, forecasts based on history often have less error due to the significance of large data volumes to time-series forecasts. Furthermore, with frequent replenishment, only a relatively small percentage of demand is needed in buffers to achieve desired service levels. The risk of carrying extra inventory is small because it will be needed in the next period.

On the other hand, products with infrequent and irregular demand are difficult to forecast. However, carrying enough supply to cover even a year's expected demand for these products might not represent a large investment. Also, depending on product commonality and distribution requirements, it might be preferable to hold inventory at "upstream" locations where different finished goods might use common components.

From a holistic perspective, balance the value of incremental improvements in forecast accuracy against the cost of achieving that improvement, or of carrying additional inventory to overcome forecast errors. The concept is to match execution and planning strategies with the characteristics of specific products and markets.

AI cluster analysis (***https://en.wikipedia.org/wiki/Clusteranalysis***) can be used to help identify groups of products and customers that exhibit similar demand, supply, and other characteristics that might therefore be managed using similar strategies.

LESSONS LEARNED

Supply-chain management systems should strive to reduce the time required for supply-chain execution to respond to changes.

• Apply AI/ML processes in steps from detection, prediction, recommendations, and then to automation
• Allow people to understand what the system is recommending or doing and why
• Facilitate people to add their insight to make incremental system improvements and to adjust automated recommendations or automated actions.

Duncan Klett was a cofounder in 1984 of Kinaxis, whose RapidResponse product offers an interactive tool for supply-chain planning. As a certified professional engineer with more than 30 years of experience with analytics and software solutions, Duncan now acts as a liaison between manufacturers, the Kinaxis design team, and executives. Prior to starting Kinaxis, Duncan held various positions with Bell Northern Research and Mitel Corporation. He also has been involved with several professional and community organizations and regularly gives lectures at universities and industry events.

dklett@kinaxis.com

This article appeared in Winter 2020 issue of Foresight magazine, a publication available for members of the International Institute of Forecasters, and is made available to you with permission of the publisher.