

Managing a Fashion Supply Chain

by **Satyashri Mohanty**



Implementing pull based supply chain solutions of Theory of Constraints involves significant paradigm shift from the conventional push based supply chain. When most companies, even the big names in many industries, are having push-based systems, the common experiential knowledge of most managers and even consultants are from such systems. As a result there is continual threat of a pull-based system being dismantled as most managers and consultants rely on their past experiences to “improve” systems.

This article was written for stake holders at corporate office of a large fashion company (referred as “The Fashion Company”), who despite getting fantastic results from pilot roll out of TOC solutions, in one of the divisions, were being “tempted” by consulting companies to go back to the push based system of forecasting.

Understanding results of pilot implementation

If we compare the actual off take of MTAM¹ SKUs (end product entity) for pilot wholesalers with the booking² quantities – the results provides us with an interesting insight. The actual off take varies widely from the initial booking quantities that they had given. Some qualities and shades sold twice compared to the booking quantity, while for others it is as low as 50% of the booking quantity. This kind of wide variation is seen in almost all wholesalers who have switched on to the MTA system, regardless of the difference in managerial competency, years of experience in business or IT systems available with them.

Remember, we are dealing here with perennial SKUs. If the forecasting error is so high for qualities and shades that are sold around the year, can we even imagine the errors that would happen with new products?

The Revelation

The people who have the deepest knowledge of their own local market (the wholesalers) are actually unable to forecast with any reasonable accuracy, even for products, which sell regularly. (In the seasons before the rollout of the MTA system, they have been suffering from high inventory and stock outs due to the gross errors in forecasting in perennials).

Possible Hypothesis

We are faced with either of the below two hypotheses to explain the phenomenon (of gross errors of forecasting by the wholesalers)

- All wholesalers do not use good “scientific” forecasting tools and hence suffer from using “gut” based forecasting decisions.
- It is futile to expect any significant improvement in accuracy. Forecasting demand of a fashion related SKU within a wholesaler's region for a booking season is, theoretically speaking, highly error prone.

The consulting community and software vendors have made us believe that the reason for poor forecasting is lack of use of “scientific tools”.

While the Theory of Constraints assumes hypothesis two is correct.

We have a problem!

How do we know which of the two is a valid hypothesis? If something is being promoted as a science, how do we invalidate that?

We will not go into philosophical debate of what is “science” and “non-science”. But we should be careful; the use of mathematical models does not necessarily mean we are being scientific! Math is not always Science!

Unless we get into a theoretical explanation of why hypothesis two is valid, the perceived sophistication of the tools will overwhelm us into accepting them under pressure of being “scientific”. (Rather, the fear of being labelled “unscientific” if we oppose those tools!).

Is there a theoretical fallacy behind these tools?

One of the latest branches of science called the Chaos Theory states that it is almost impossible to predict (with any reasonable accuracy) an output of a complex system because of non-linear feedback loops in reality. The relationship of variables in this system is not predetermined and a small change can have a disproportionate impact and vice-versa. (Remember the butterfly effect? – “a butterfly flapping its wings in China can lead to a tornado in Tokyo”!)

The demand of fashion related SKUs at selling points has the characteristics of a complex system and hence any mathematical model, which uses predefined linear or even non-linear relationships between variables, cannot have any reasonable accuracy. Even if we put 20 variables in our model, the error can throw us off gear. How do we know that market demand of fashion items is a complex system with non-linear feedback loops?

We can relate to experiences of sudden spurts in demand.



A delay in a competitor's supplies can cause a sudden surge in demand for our material in a location or a fashion icon wearing a specific style can cause a surge of demand in another. The news of increased high sales can, in turn, fuel the demand further.

Because of the non-linear feedback loops, the “noise” of the system is very high. So any forecast model tends to behave erratic, time-to-time, giving rise to huge errors, rendering it impractical for decision making. To nullify the effect of local variation, many people advocate an “aggregate forecast” or the top down forecast. Top down forecasting (which tries to take advantage of aggregation) or bottoms up forecasting (which takes advantage of knowledge of many local variables) or even a combination of the two does not yield better results when one has to forecast for large number of SKUs for a long horizon. Even if the aggregate forecast looks better, the advantages are lost when the inventory is eventually pushed to the point of maximum errors – the retail point.

The TOC Approach to the problem

It is a waste of time to even try and model a non-linear reality of the market. Neither can it be easily modelled nor it is required to be done.

So, instead of trying to improve the forecasting model, it is important to drastically improve the reaction time so that we suffer minimal from forecasting errors.

In case of wholesalers under MTA, we have a system which reacts 2 to 3 times a month to changes in market demand compared to the earlier scenario of booking for the entire season where reaction is only once in six months. The central stocking in the CWH and efforts of reducing production lead-time were with the purpose of improving the reaction time to market.

Due to faster reaction time of the supply chain, the wholesalers have suffered minimally from the forecasting errors. With low replenishment lead times to the wholesalers, the company is better equipped to deal with changes in market demand of each individual wholesaler.

The rule of dynamic buffer management³ makes any sophisticated forecasting model redundant. If an SKU does not sell fast, the rules of dynamic buffer management will kill it within 3 or 4 replenishment cycles, while if another SKU is moving fast, within 1 or two replenishment cycles, the norms go up. This ensures that inventory is always at the optimal level. The replenishment time is about 10 days for wholesalers under MTA, down from 6 months as under the booking cycle. You can estimate the speed of reaction time under the dynamic buffer management.

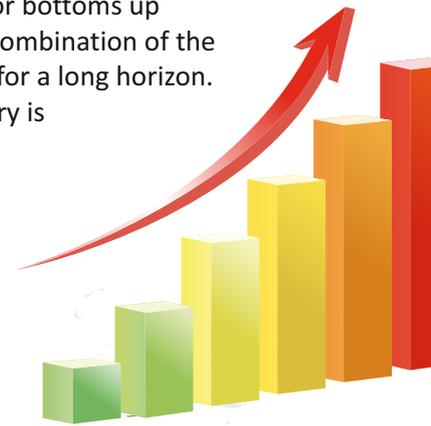
What about the non-MTA products?

It looks like we have a practical approach to solve the conundrum of high inventory and stock outs in case of MTA products. But a significant volume of business is from non-MTA where we expect forecasting errors to be even more pathetic.

Do we have to rely on a crystal ball or do we have to just learn to live with the bigger problem? Since that does not sound interesting, let us see if there is a way out.

For a moment let us switch topics and approach the problem from a different perspective – the retail selling point.

We have seen the errors of forecasting of wholesalers/retail points when they operate in a “2-season booking” method. For every season, there is always sediment of high inventory of some slow moving items left behind. This means, over time, we expect the average inventory holding of a retailer to increase – it



explains the inventory turns of 1 or 2 at the selling points. If 1 is the average inventory turns, we actually expect some SKUs to be lying in shelf to be older than a year for sure. Low inventory turns at selling point actually means that a significant portion of the shelf space (and the attention of the seller!) is actually occupied with slow items. Even with discounting schemes to clear shelf space, shelves available for new introductions are always limited. At the same time, the wholesalers with low ROI also have limited ability to take new products.

Combine the above effects with the fact that every year, there is large range offered by the company for new products; we have a rather interesting irony. Even when fashion companies introduce a large variety every year, the actual shop display does not undergo a radical churn frequently.

Now let us see if we can redefine the game with MTA in place.

The Fashion Company has a capability of having a large portfolio of SKUs under MTA, much more than the range that a wholesaler or a retailer can keep. With the wholesaler making more ROI (Return on Investment) under MTA, he will definitely try and shift his business to more MTAs. We are seeing this happen in many pilot wholesalers. He will experiment with new MTA SKUs because the quantities that he has to hold for the new ones will be dramatically less (maximum one month of sale) than what he would have done in 6-month booking system. If the product sells well, he will be replenished fast from the central warehouse. If products do not sell well, the rules of dynamic buffer management will try to reduce inventory rapidly and thus reducing the risks of the distributor and improving his portfolio. (Remember in MTA scenario, the dealer at any point of time, needs inventory only to cover for transportation time from central warehouse to his location).

For him and his retailers, he has actually added to his portfolio without waiting for collections. He can now churn his portfolio amongst the MTA SKUs both with him and the retail point. This means we will have new products at retail point without having to actually develop them.

More and more wholesalers will take advantage of the large offering of MTAs by The Fashion Company to add freshness to their portfolio, while enjoying very high ROI. We expect the volume of business from MTA to go up significantly while at the same time, the stores will have more freshness than before.

The Fashion Company will experience a new paradigm – we need not always develop new products to introduce new products at the retail counters.

So what about the fashion icon status of The Fashion Company?

With more business coming from MTA SKUs, wholesalers may select less from the new products offered. Now on the face of it – it looks like The Fashion Company might lose its “fashion” icon status. (Even though in the MTA system, we are actually introducing new products much faster without developing them every time).

But let us see if there is actually a huge opportunity instead of a problem.

If we look at the 6-month booking system where the agents are asked to select from a plethora of new SKUs on a specific day, we will find many SKUs not getting enough attention.

It is not only humanly impossible to provide effective attention to each of the SKUs in the large offering but when wholesalers have low inventory turns (about 2), they have restricted capital to invest in many new products. So fear of locked capital becomes the dominant factor in selections.

Combine this with the fact that wholesalers do make huge forecasting errors. So, though we might be developing products, which are inherently winners, the wholesalers and agents may not pick them up, and at same time they could be selecting inherent losers assuming otherwise.

Remember the key to a fashion supply chain is fast reaction and not accurate forecasts. So instead of asking the dealer for forecast requirements for a 6-month season, we can ask for shorter horizons to reduce his errors in forecasting.

But we have a problem with more seasons. In current environment when wholesalers start buying for the season at start of the season, it forces The Fashion Company to rely on forecast for first 1-2 months of the season – so every season starts with “blind” zone. More seasons means more blind zones. Each period of blind zone exposes The Fashion Company to the mess of forecasting errors – high inventory and stock outs. Does not seem like a great idea.

But let us try to understand why seasons lead to periods of blind zone. It comes from lack of reliability of supplies and hence wholesalers wait till closure of supplies to start the next season. So if we improve reliability of supplies, we can ask wholesalers to order without waiting for start of a season but they would want a smaller season.

And now we have a win-win!

We can increase seasons, which will reduce the period of forecast for dealers, and at the same time the dealer need not wait for end of supplies to start for next season – no blind zone for The Fashion Company. But wait; there is one more negative. More seasons might decrease the quantity per SKU and cause more problems in the plant.

When one offers a large collection in one big lot, we have an issue of SKUs not getting enough attention. With many distributors going through a large portfolio, the chances of getting small quantities per SKU are very high. But when we offer a small collection each time, the focus is on few SKUs. We move to many seasons but each season has a smaller portfolio. With focused attention on a smaller range, we should start getting more quantities per SKU. Yes we have to find a way to reach out to the wholesalers very fast with each collection story.

At this stage of implementation, the wholesalers should be already enjoying high ROI and hence their ability to invest in new products is enhanced dramatically.

Now we have a solution where all new products go through a pure MTO (made to a confirmed order) route without need to forecast. When a product starts selling more, we can put them in the MTA. The MTAs, which are not selling well, will be killed by the dynamic buffer management rules.

We will have a system, which does not require a sophisticated forecast model either for MTA or for MTO. With increase in MTA business, the quantity per SKU for MTA will increase. The plants will have more quantities per SKU and that will improve the output of spinning and weaving.

What does all this do to the financial performance of The Fashion Company? Significant!!

Yes the changes are fundamental paradigm shifts in approach to market and they are not easy. But then who wants an easy life?

¹Made to Availability is the TOC replenishment solution which depends on actual consumption to trigger replenishment in entire supply chain.

²The fashion company has a process where wholesalers have to place their confirmed requirements upfront for a booking season. There are two seasons in a year, each lasting for 6 month

³With drastic reduction in replenishment time, dynamic buffer management uses sales rate of past replenishment period and compares it with speed of replenishment to continuously adjust norms. If MTA items remain in red zone for entire replenishment period, the norm is adjusted upward and vice-versa.

Vector Consulting Group (www.vectorconsulting.in) is the leader of 'Theory of Constraints' consulting in India. Vector has been working closely with some of the well known FMCG, Engineering Goods, Custom Manufacturing and Auto Components companies to improve their overall profitability through supply chain effectiveness.