Contents lists available at GrowingScience

Engineering Solid Mechanics

homepage: www.GrowingScience.com/esm

Proposed method of forecasting cumulative effects of variation in manufacturing

Quinn Risch[a*]

[a]*Raytheon, United States*

ARTICLE INFO

Article history:
Received 24 January 2022
Accepted 12 March 2022
Available online
12 March 2022

Keywords:
Build Forecast
Manufacturing
Optimization
BOM
Supplied Parts

ABSTRACT

Manufacturing, in general, creates a finished good from a set of simpler supplied parts. Supplied parts are installed into higher assemblies, higher assemblies move into even higher assemblies, and eventually this terminates at the finished good. Delays or variation during the manufacturing process ripple all the way to the finished good, possibly from different branches of the build and possibly magnifying any individual effect. There is extensive literature regarding Lean Manufacturing and it provides strategies and business philosophy to deal with variation, however it offers little in the way of quantitative analysis on the effects of that variation upon the whole. Digital Twins and discrete event simulations can and have been used to model the impact of variation in its totality. Various papers on Digital Twins have explored how to model manufacturing, but very little on generalized behavior. (i.e. How schedule slips at the subassemblies impacts the delivery dates / quantities at the finished good level). This paper explores the analytical quantitative effects of input/sales variation through the manufacturing cycle and the resultant effect on the finished good manufacturing schedule/cycle. We demonstrate that even small random variations/interruptions propagate up the build chain, get reduced in magnitude and end up producing predictable reductions in the average build rate of the final product. Additionally, it is shown that the more supplied parts that comprise a finished good the greater the expected reduction in average build rate.

1. Introduction

Reducing waste, improving quality and increasing production are the main strategies to maximize customer value and producer profit (Taylor, 1911). The Toyota Production System (TPS) model of manufacturing first described in (Womack, Roos, Jones, & Carpenter, 1990) changed manufacturing from a buffered production model, where there was excess capacity dealing with variation, to a Lean model (James P Womack, 1996), a term first coined by John Krafcik (Krafcik, 1988). The concept of Lean manufacturing aims to reduce waste in a value stream by shortening time between order placement and delivery (Bhasin & Burcher, 2006).

Lean tells the manufacturer to identify customer value, map the value stream, create work flow, establish a pull system and use continual improvement (Mostafa, Dumrak, & Soltan, 2013; Jasinowski & Hamrin, 1995). The philosophy and principles of Lean produce results, but the philosophy and principles of Lean do not predict quantitative/numerical results (Pearce & Pons, 2019). From (Womack, Roos, Jones, & Carpenter, 1990), (Krafcik, 1988), (Mostafa, Dumrak, & Soltan, 2013), (Bicheno, 2016), and (Shingo, 1987) the study and practice of Lean manufacturing has been adopting strategies to maximize performance, but these principles never inform the reader what the maximum performance is.

Performance prediction/forecasting has typically been the domain of factory simulations and digital twins (Terkaj & Urgo, 2015). When large scale discrete event simulation tools are used (e.g. ProModel) there is a great deal of initial work to model

* Corresponding author.
E-mail addresses: quinn@rtx.com (Q. Risch)

doi: 10.5267/j.esm.2022.3.001

everything necessary/desired. For any sort of predictive power the simulations needs to encompass many thousands of units per trial and the simulation need to repeated several times in order to characterize the empirical distribution of the entire system (Law, 2008), which in turn characterizes factory performance (Hoad, Robinson, & Davies, 2010).

The current work provides a framework for analysis of how variation of completions at each lower level affects the finished good's completions/sales. This effect is similar to the Bullwhip effect (Lee, Padmanabhan, & Whang, 1997; Forrester, 1961; Siegele, 2002) where variation in customer demand has an amplified effect on the supply chain.

By taking the Bill Of Materials (BOM) and forming a drawing tree (**Fig. 1**), we define the direction of the variation propagation. Any delay or variation in one part/branch affects the next higher assembly (NHA). The entire build, for the purposes of analysis, will be dictated by the drawing tree and all work for each subassembly will be defined to start once all the composing parts are available.

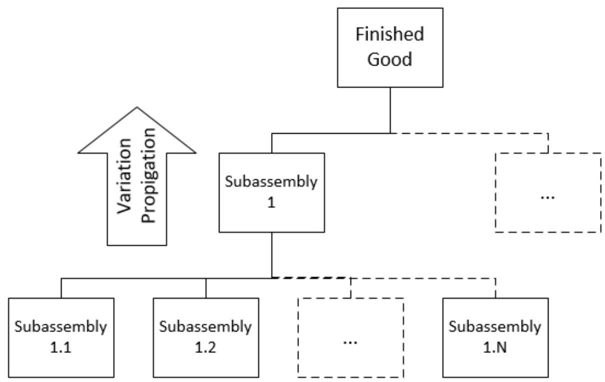

Fig. 1. Drawing Tree and Variation Propagation

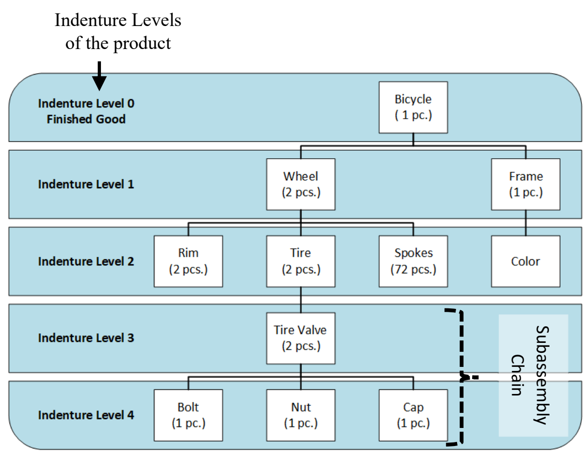

Fig. 2. Example BOM

Variation in each of these elements of the product affect the finished good slightly differently and most importantly all variation starts with the supplied parts. Starting with the supplied parts and their delivery schedules one can forecast the maximum achievable build rate for the finished good. With the predicted maximum build rate, a producer can improve the factory capabilities or adjust delivery schedules and or expectations.

The goal of the analysis is to define the ceiling of performance and not the actual factory performance. The analysis simplifies the problem by ignoring various resource limitations and performance variables of the value stream. The basic framework allows for a first order approximation of the non-linear effects of variation which helps predict maximum factory performance to whatever confidence level desired.

2. Materials and Methods

2.1. Assumptions

There are some basic premises that this paper will build upon that should be explicitly stated.

2.1.1. Assumption Number One

All Finished Goods are ultimately paced by the supplied parts. Regardless of how fast the build rate is at any higher level, one cannot build faster than the rate of the supplied parts. The equations use only the build rate and variation of the supplied parts and thus any forecast is the ceiling of performance (i.e. the maximum sales of the finished good).

2.1.2. Assumption Number Two

Any assembly that is composed of lower indenture parts will typically have some portion of the build that is started before all the subassemblies are installed (see Fig. **3**, blue diamond to grey diamond).

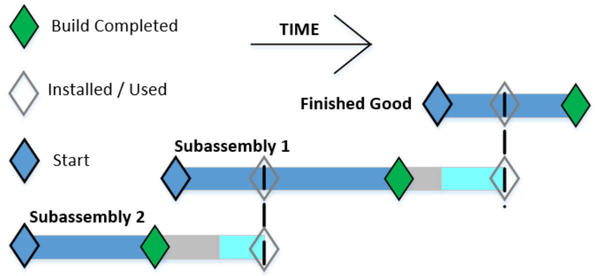

Fig. 3. Subassembly Build Time Overlap

This paper will assume the time from ‘Start to Install’ (blue to grey diamond) is much smaller than the time from ‘Install to Build Completion' (grey diamond to green diamond). That is we will assume all parts are kitted at the start of any work. Accordingly, in the general work up of the solution the possibility of this overlap will be ignored. This assumption allows the total build time (i.e. finish good performance) to be calculable from a smaller number of variables and therefore allows a simpler analysis. The general conclusion resulting from this assumption will be the same.

2.2. Calculating Cumulative Distributions (CDF)

2.2.1. Supplied Parts

Without a loss of generality, one can rely on the Central Limit Theorem[1] and extend the results to cumulative totals (rather than averages). Long-term behavior of the completions/sales can be treated as being Gaussian (because of the Central Limit Theorem) (Montgomery, 2014).

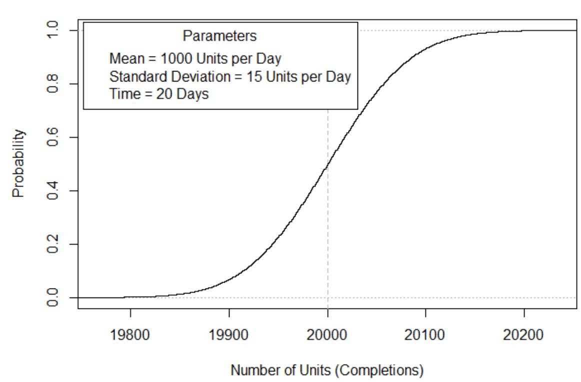

Fig. 4. Example CDF plotted

Fig. 4 illustrates the long-term probability of completions at some time for one particular supplied part. Each supplied could potentially start production at different times, and therefore we must ensure whatever mathematical model can take this into account.

[1] The Central Limit Theorem only relies upon the assumption of a finite variance, and this constraint upon the sales distribution is reasonable, if not required.

2.2.2. Next Higher Assemblies (NHA)

Supplied parts ultimately become part of the larger assembly, the next higher assembly (NHA). Ultimately, the entire build is composed of the supplied parts, and it makes sense that the Finished Good's performance ceiling will be paced by the supplied parts.

The movement/transfer of parts from a lower indenture level to the NHA level cannot happen instantaneously and therefore, one expects some delay before assembly at the NHA level can start. Our analysis must contain a delivery delay relationship of parts from one indenture level to the NHA.

Build and delivery time is always greater than zero thereby decreasing the amount of time that items are actually being built and likewise reducing the number of parts received at the next higher assembly (NHA).

2.2.3. Application to the Finished Good

Getting to the last NHA, the Finished Good, is achieved in a stepwise manner across the bill of material starting from the supplied parts and working up the drawing tree (to the Finished Good). At each indenture, the delivery/transportation time for that part contributes to a delay at the NHA. Accounting for delivery time in the NHA is straightforward and it effectively reduces the production time at the NHA (Fig. 5).

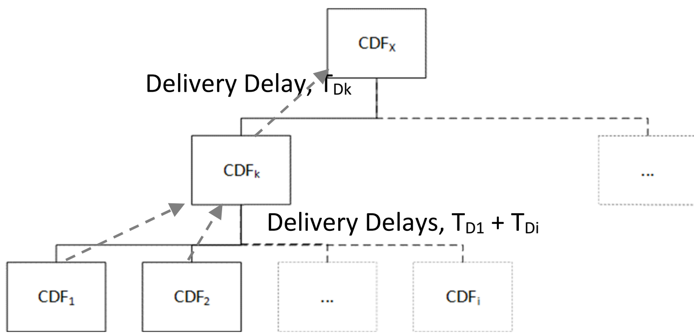

Fig. 5. Build and Delivery Delays Propagating Up the Drawing Tree

The result is the CDF for the finished good solely as a function of the supplied parts (Fig. **6**).

2.2.4. Finished Good Example

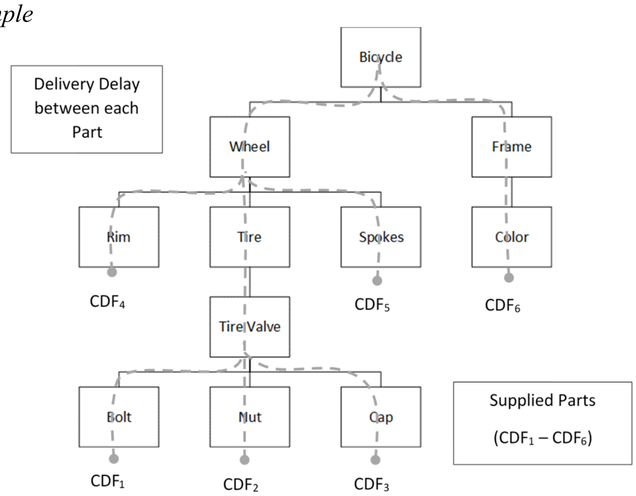

Fig. 6. CDF Evolution Up the Drawing Tree

The resulting calculation from even the simple case of Fig. **6** will be quite complex.

The resultant CDF shows the dependency of the finished good on all the supplied parts. All drawing trees with the same quantity of supplied parts will have a similar CDF equation, but with different build times and delays, rates and standard deviations.

2.2.5. *CDF to Likelihood Calculations*

Calculating the CDF provides a way to calculate the statistical likelihood, L, of achieving some number of finished goods, N, at some future date (i.e. the expected number of finished goods over some period of time).

$$L = [1 - CDF_{Final}(N)] * 100 \quad (1)$$

where L is the likelihood of achieving N completed units at T_{end} (from 0 to 100%). Using Eq. (1) a factory can estimate the maximum delivery schedule of a finished good in accordance to their tolerance for risk (i.e. confidence level). A very conservative risk posture of 3σ (99.73% likelihood) is shown in Fig. 8.

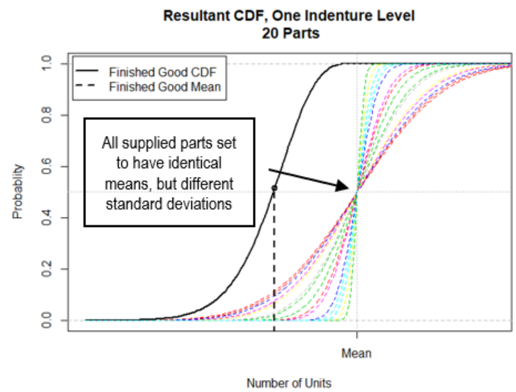

Fig. 7. CDF of Supplied Parts and Finished Good

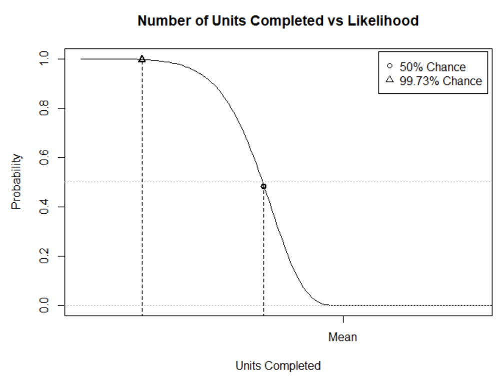

Fig. 8. Likelihood of the Finish Good Achieving Some Number of Completed Sales (from Fig. **7**)

Any Finished Good completion schedule can be forecast (simple or complex). Fig. 9 uses the example of Fig. 6 illustrate the differences between the Supplied Parts delivery rates and the maximum delivery rate of the Finish Good.

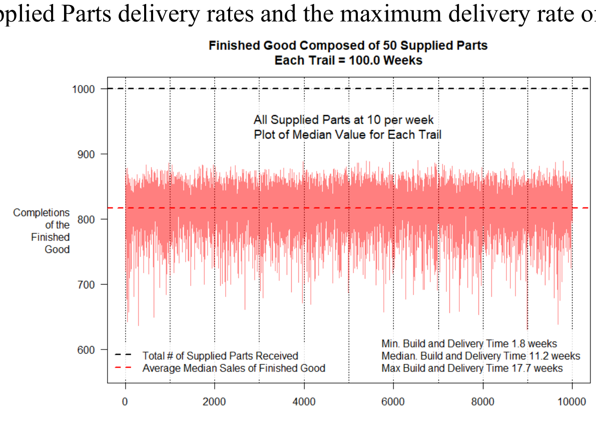

Fig. 9. Median Performance of Finish Good

A factory can forecast performance and change behavior in accordance to quantified predictions of behavior and their tolerance for risk.

3. Results

3.1. Effect of Indenture Level

Typical manufacturing wisdom states that as the number of parts increase then the expected variation in finished good increases. This can be shown to be false. As expected, given several numerical simulations, the average performance of the next higher assembly decreases with each supplied part added. Fig. **10** & Fig. **11** show that the range of sales of the Next Higher Assembly (red) decreases with the addition of more supplied parts (shown in black) despite the fact that all the supplied parts all have the same average performance (same mean build rate and same variation/standard deviation).

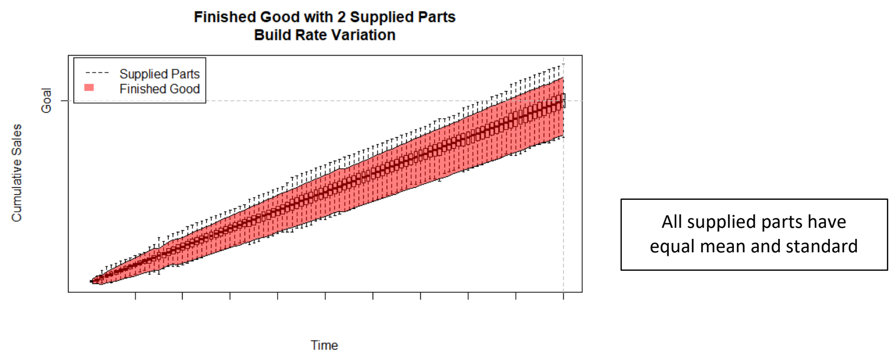

Fig. 10. Finished Good Variation, 2 Supplied Parts

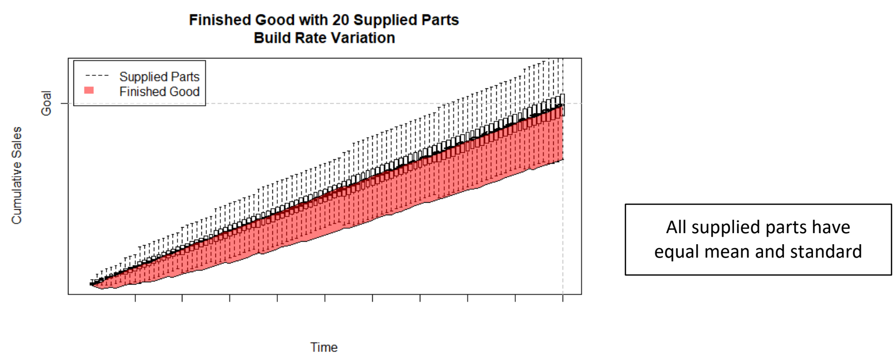

Fig. 11. Finished Good Variation, 20 Supplied Parts

Although the schedule performance of the Finished Good decreases with an increase in supplied parts, as would be predicted; the variation behaves contrary to expectation: variation at the Finished Good level decreases with an increase in supplied parts. Given the in depth mathematical analysis of this paper, we have determined that as the number of supplied parts increases, the Finished Good is completed more predictably, but predictably slower.

3.2. Effect of Subassembly Chain/Number of Indentures

One would predict any subassembly chain (i.e. building in series, Fig. 13) has a theoretical maximum build rate that is equal to the supplied part. However, a subassembly chain (the set up parts built in series) will compound delays (taking the maximum at each level and adding them together), therefore ΔT will decrease at each level thus reducing time of performance and total Finished Goods produced. Driving delivery times at each level towards zero decreases this effect, however, the effect will never go away entirely.

Even as delivery time goes to zero, there is the practical matter of actual performance. If the manufacturing at the higher stages of build is planned to have the same capacity and variability as the delivery schedule of supplied parts, then the Finish Good is doomed to be slower than the supplied parts' delivery schedule merely due to random chance (see Fig. 12).

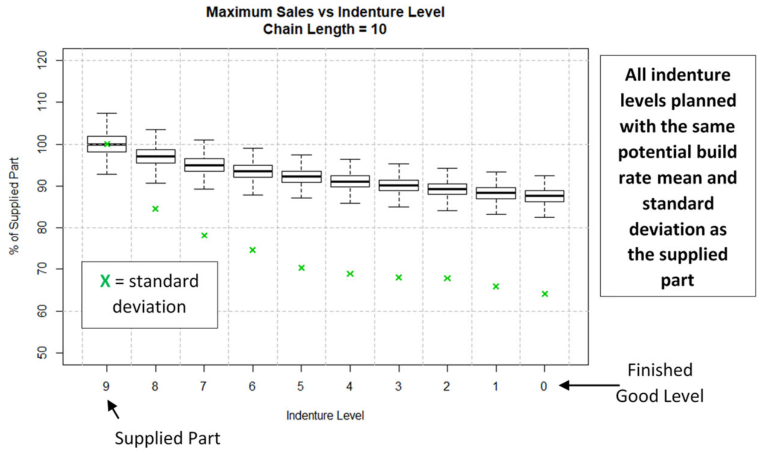

Fig. 12. Supplied Part Variation Eating Away at Finished Good's Average Performance

To avoid this slow down each next higher assembly will have to have a zero delivery delay and consume all the supplied parts as quickly as they arrive for the theoretical maximum to be achieved. This means that the build rates further up the chain will have to be able to consume even the largest variation from the lowest level immediately (i.e. the average build rate at the NHA will need to be the supplied part mean plus the +3σ). The choice for a business is to either introduce inefficiencies at the higher factories by requiring excess capacity at the higher indenture levels, or to accept and plan for production rates at higher indenture levels that are less than the supplied parts (and slower than the lower than the rates at each subsequently lower indenture level).

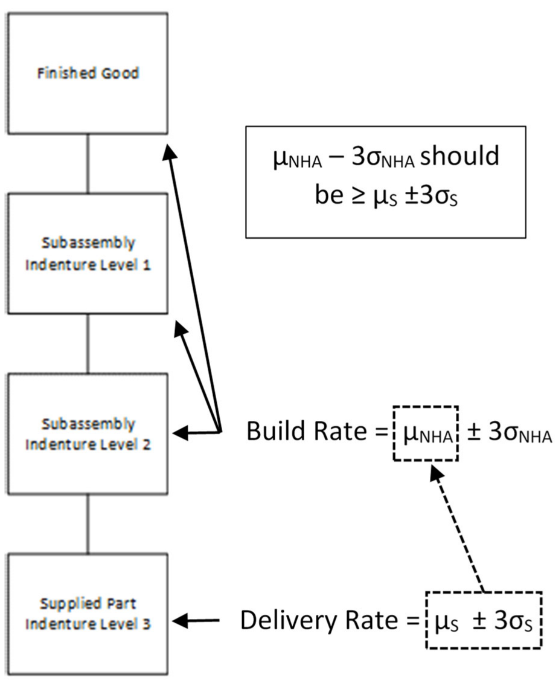

Fig. 13. Single File Subassembly Chain (a build in series)

4. Discussion

Common wisdom would suggest that the structure of the drawing tree would play no role in how to plan manufacturing, We predict that is not the case. Broad and flat drawing trees behave different from long/deep subassembly chains (Fig. 14 & Fig. 15). Delivery delays become less and less significant as time goes on (as the delivery delay becomes a smaller percentage of total time passed), and ultimately the number of parts composing a finished good is not as important as the number of supplied parts.

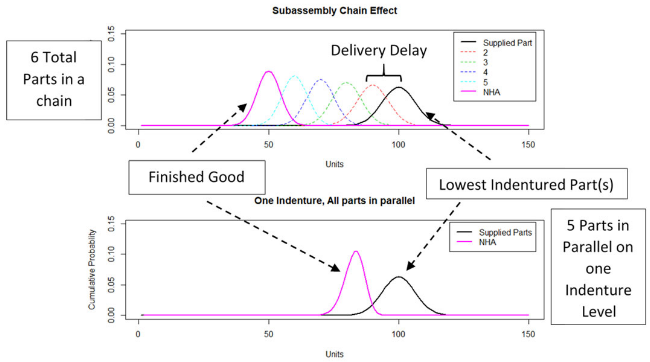

Fig. 14. CDF, Subassembly Chain (series) versus Flat Indenture (parallel)

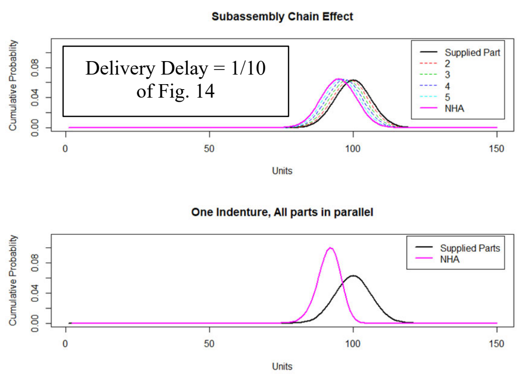

Fig. 15. Small Delivery Delay versus Flat Indenture

By not accounting for even tiny amounts of variation in the supply chain, any manufacturing plan will underperform expectations and it will likely adversely affect business. When even optimistically small variation is applied, one can better estimate/approach real factory performance. One can, in general, predict that the greater number of supplied parts will lead to slower average build rate of the finished good, however the variation of the build schedule for the finished good will be smaller than any of the supplied parts.

5. Conclusion

Any Finished Good can have its theoretical maximum completion schedule calculated. When a factory is armed with this knowledge, there can be a reasonable assessment of resources and constraints. The equations outlined in the appendix can be applied in a number of ways, either in a software package, like Excel, or R (as was done by the author, https://cran.r-project.org/) or the equations could be implemented in a purpose built enterprise application. Any effort to account for variation in the build will help more honestly forecast the actual achievable build rate of the finished good.

Acknowledgment

The author would like to thank all the colleagues that provided input and listened to his ramblings on the subject. Without their patience, this paper would have never come to fruition.

References

Bhasin, S., & Burcher, P. (2006). Lean Viewed as a Philiosophy. *Journal of Manufacturing Technology Management, 17*(1), 56-72.
Bicheno, J. &. (2016). *The Lean Toolbox* (5th ed.). Buckingham: Picsie.
Forrester, J. W. (1961). *Industrial Dynamics.* Cambridge, Massachusetts: MIT Press.
Hoad, K., Robinson, S., & Davies, R. (2010). Automated selection of the number of replications for a discrete-event simulation. *Journal of the Operational Research Society, 61*, 1632-1644.
James P Womack, D. T. (1996). *Lean Thinking* (1st ed.). Mishawaka, IN, USA: Taylor & Francis.
Jasinowski, J., & Hamrin, R. (1995). *Making It in America: Proven Paths to Success from 50 Top Companies.* New York City: Simin & Schuster.

Krafcik, J. F. (1988). Triumph of the Lean Production System. *Sloan Management Review, 30*(1), 41-52.
Law, A. M. (2008). Output Data Analysis for a Single System. In A. M. Law, *Simulation Modeling and Analysis* (pp. 485-547). New Delhi: Tata McGraw-Hill Publishing Company Limited.
Lee, H. L., Padmanabhan, V., & Whang, S. (1997, April). Information Distortion in a Supply Chain: The Bullwhip Effect. *Management Science, 43*(4), 546 - 558.
Montgomery, D. C. (2014). In *Applied Statistics and Probability for Engineers (6th Edition)* (p. 241). Wiley.
Mostafa, S., Dumrak, J., & Soltan, H. (2013). A framework for lean manufacturing implementation. *Production & Manufacturing Research, 1*, 44-64.
Pearce, A., & Pons, D. (2019). Advancing lean management: The missing quantitative approach. *Operations Research Perspectives, 6*(100114).
Shingo, S. (1987). *The Sayings of Shigeo Shingo: Key Strategies for Plant Improvement.* (A. P. Dilon, Trans.) New York City, New York: Productivity Press.
Siegele, L. (2002). Chain reaction: Managing a supply chain is becoming a bit like rocket science. *The Economist, 362*(8258), 13-15.
Taylor, F. W. (1911). Chapter 2: The Principles of Scientific Management. In *The Principles of Scientific Management* (pp. 30 - 144). New York City: Harper & Brothers.
Terkaj, W., & Urgo, M. (2015). A Virtual Factory Data Model as support tool for simulation of manufacturing systems. *Procedia CIRP, 28*, 137 - 142.
Womack, J. P., Roos, D., Jones, D. T., & Carpenter, D. S. (1990). *The Machine that Changed the World.* New York City: Free Press, Simon & Schuster.

A.1. Appendix

A.1.1. Derivations/Equations

Any subassembly's sales over time is described by a distribution, SD. SD is independent and random.

$$SD[\mu,\sigma] = \{X_1, X_2, \dots X_n\} \tag{A.1}$$

where μ is the mean of the sales for the time period, σ is the standard deviation of sales for that time period, and X is the random sales number. The Sample Average, S_n is:

$$S_n = \frac{X_1 + X_2 \dots + X_n}{n} \tag{A.2}$$

where *n* is the number of samples. Applying the Central Limit Theorem yields,

$$\lim_{n\to\infty} SD[\mu,\sigma] \rightarrow N[\mu, \frac{\sigma^2}{n}] \tag{A.3}$$

where N is the Normal (Gaussian) distribution and σ^2 must be finite. To get to the average cumulative total, CT, from the Sample Average, S_n, one only needs to multiply by n.

$$n * S_n = n\left(\frac{X_1 + X_2 \dots + X_n}{n}\right) = (X_1 + X_2 \dots + X_n) = C_T \rightarrow n\mu \tag{A.4}$$

With the Central Limit Theorem any Sales Distribution, SD, is transformed into a Gaussian distribution with the original mean and variance multiplied by n.

$$\lim_{n\to\infty} n * S_n = X_1 + X_2 \dots + X_n \rightarrow N[n\mu, n\sigma^2] \tag{A.5}$$

For the application to a sales distribution the variable n is a unit of time measure and for convenience we can change the notation.

$$n \xRightarrow{becomes} T$$

$$\lim_{T\to\infty} T * S_T = C_T = X_1 + X_2 \dots + X_T \rightarrow N[T\mu, T\sigma^2] \tag{A.6}$$

A.1.2. Dependency of Upstream Sales Distribution

Each subassembly has a sales distribution (previously derived) that can be described as:

$$Sales\ Distribution\ of\ jth\ Subassemby\ =\ SD_j(\mu_j, \sigma_j, Max_j, Min_j) \tag{A.7}$$

where SD is the probability distribution function (PDF), μ_j is the mean, σ_j is the standard deviation, Max_j is the upper range of the distribution, Min_j is the lower range of the distribution. The cumulative maximum sales of the Finished Good is the minimum of the cumulative sales of all subassemblies that compose it, i.e. the distribution of the minimum of the subassemblies sales distributions.

$$FLOOR\left\lfloor \begin{matrix} SD_1(\mu_1, \sigma_1, Max_1, Min_1) \\ SD_2(\mu_2, \sigma_2, Max_2, Min_2) \\ \ldots \\ SD_j(\mu_j, \sigma_j, Max_j, Min_j) \end{matrix} \right\rfloor \rightarrow SD_{min}(\mu_m, \sigma_m, Max_m, Min_m) \tag{A.8}$$

where SD_{min} is the distribution of the minimum of the subassemblies' PDFs (1 through j).

As the number of subassemblies that comprise a Finished Good increases the resultant distribution of minimums (Equation A.8) has its mean, μ_m, approach the smallest minimum of the subassemblies.

$$\lim_{j \to \infty} \mu_m = Min_m = FLOOR\left\lfloor \begin{matrix} Min_1 \\ Min_2 \\ \ldots \\ Min_j \end{matrix} \right\rfloor \tag{A.9}$$

The general solution to find the distribution of the minimum for a set of subassemblies that comprise a Finished Good is:

$$\begin{aligned} (1 - CDF_{min}) &= (1 - CDF_1)(1 - CDF_2) \ldots (1 - CDF_j) \rightarrow \\ CDF_{min} &= 1 - (1 - CDF_1)(1 - CDF_2) \ldots (1 - CDF_j) \rightarrow \\ \frac{d}{dx} CDF_{min} &= SD_{min} = \frac{d}{dx}\left[1 - (1 - CDF_1)(1 - CDF_2) \ldots (1 - CDF_j)\right] \end{aligned} \tag{A.10}$$

where CDF_j is the Cumulative Distribution Function of the jth subassembly, SD_{min} is the Probability Distribution Function of the minimum of the subassembly sales distributions, and CDF_{min} is the Cumulative Distribution Function of the minimum of the subassembly sales distributions.

The resultant CDF can then be used to predict the probability of sales as a function of N, the number of units at a given time.

$$\mu, 3\sigma \xrightarrow[Theorem]{Central\ Limit} SD[\Delta T\mu, \Delta T\sigma^2] = \frac{1}{\sqrt{2\pi\Delta T\sigma^2}}\, e^{-\frac{(N - \Delta T\mu)^2}{2\Delta T\sigma^2}} \tag{A.11}$$

$$CDF = \frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{N - \Delta T\mu}{\sigma\sqrt{2\Delta T}}\right)\right] \tag{A.12}$$

The time to which the product at any one level is made, ΔT, must take into account delivery or transportation delays from one indenture to the next.

$$\Delta T \rightarrow \Delta T + \sum_{i=1}^{M} \Delta T_{Di} \rightarrow \Delta T + D(i, \ldots, M) \tag{A.13}$$

where T_{Di} is the delivery delay of the part to its NHA and then to its next HNA, etc. and D(i, …, M) is identically the same as the summation notation, but more compact (i to M is chain from the first part to the last/Finished Good – Fig. A-16).

$$CDF_i = \frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{N - [\Delta T + D(i,\dots,M)]\mu_i}{\sigma_i\sqrt{2[\Delta T + D(i,\dots,M)]}}\right)\right] \rightarrow$$

$$CDF_i(N, \Delta T + D(i,\dots,M), \mu_i, \sigma_i) = \frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{N - [\Delta T + D(i,\dots,N)]\mu_i}{\sigma_i\sqrt{2[\Delta T + D(i,\dots,N)]}}\right)\right] \rightarrow$$

$$CDF_k = [1 - [1 - CDF_1(\Delta T + D_1(i,\dots,N), \mu_1, \sigma_1)]\,[1 - CDF_2(\Delta T + D_2(i,\dots,N), \mu_2, \sigma_2)] \dots [1 - CDF_i(\Delta T + D_k(i,\dots,N), \mu_i, \sigma_i)]\,] \quad (A.14)$$

where D_1, D_2 through D_k are the various paths from the supplied part to the Finished Good (thus tabulating the delivery delay from the supplied parts arrival to delivery to the Finished Good).

CDF_k is the cumulative distribution function for the Sale Distribution of the Finished Good. This is the equation that this paper derives its conclusions from.

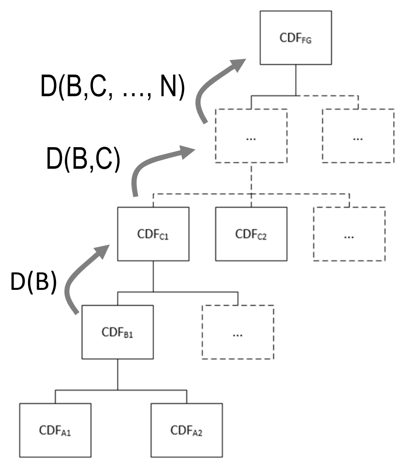

Fig. A-16. Delivery Delay Schema