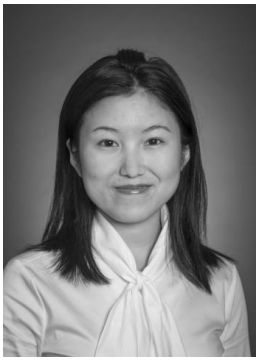
Capacity Planning for Biologics -- from Demand to Supply

By: Sifo Luo
Thesis Advisor: Postdoctoral Associate Ozgu Turgut
Topic Area: Pharmaceutical Production Planning, Biologics Manufacturing, Optimization

Summary:

This thesis provides a tool to capture both demand and supply uncertainty in pharmaceutical long range planning. In this thesis a stochastic optimization approach is followed to minimize the deviation from target capacity limit under different manufacturing and demand scenarios. The mixed integer linear optimization model incorporates the impact of demand and manufacturing variation on production allocation among manufacturing facilities through Monte Carlo generated scenarios. The thesis model is designed such that it can be used as a decision tool to perform robust capacity planning at the strategic level.



Prior to MIT, Sifo Luo graduated with a Master of Science degree in Information and Service Management from Aalto University, Finland. She worked in sales and operations in Microsoft and product management in Rovio Entertainment. After graduation, Sifo will join Converse-Nike Inc. as a strategy operations manager.

KEY INSIGHTS

1. Robust optimization helps XYZ Co. include manufacturing performance variation and uncertainty in biologics production planning. The model serves as a decision making tool by quantifying and visualizing capacity risk in a more accurate way.

2. Under demand upside, low manufacturing performance will increase production time and capacity required significantly.

3. Among all manufacturing parameters, success rate fluctuations bring a slightly more significant risk to current production planning. However, no single parameter should be given sole priority with regard to facility performance analysis.

Introduction

Biologics are large molecule drugs manufactured in living cells, often used in cancer treatment and other major diseases. Shortage of these products may endanger patients' lives and impact a company's market share. Stringent regulatory policies and complicated drug development technologies add uncertainty to supply chain planning for biologics. Once the drugs are approved, the challenge is to reliably supply therapies to patients in every market in which the drug has been approved. Constrained supply and time-consuming production setup in compliance require companies to plan far in advance. Therefore, the research sponsor, XYZ Co., a large pharmaceutical company, invests early in capacity and, on average, a new drug takes 8 to 12 years from patent filing to first sale. With long-range forecasting, risk and accuracy become important consideration.

Currently, XYZ Co.'s capacity planning is mainly constrained by the long-term demand and its forecast uncertainty. The impact of manufacturing level uncertainties, such as factory productivity and production success rate, is not entirely taken into account because only the expected self-reported values of production facilities are used in supply planning.

Being able to foresee potential capacity constraint is a key insight from the manufacturing planning point of view. XYZ Co. wants to find out the relative impact of variance of manufacturing technical parameters on active ingredients (API) production capacity. This thesis helps bridge the gap between long-range demand planning and manufacturing performance in pharmaceutical production planning.

Quantifying Demand and Supply Uncertainty

XYZ Co. provided the demand data on a single drug and the three manufacturing factories where it is produced. To account for the variation in demand,

three scenarios were generated: base case, upside case, and downside case. In addition, three technical parameters related with API manufacturing were chosen by XYZ Co. for this study – success rate, runs per week, and kilograms per run. Success rate is the expected ratio of runs (batches) successfully made over total batches. Runs per week and kilograms per run represent the output capability of each facility. Runs per week measures how many batches the site can run, given how long it takes to do a run based on all equipment that is being used. Kilograms per run is the average amount of material expected from a batch. To measure the influence of floating manufacturing parameters, two scenarios -- upside and downside -- were created. For each parameter, scenario values are assumed to be uniformly distributed between observed maximum and minimum value over the years. Capacity of manufacturing facilities is measured in weeks. The full capacity of each site is 52 weeks. Each site is identical with respect to the target capacity (80% of full capacity) and minimum target capacity (50% of full capacity). Each manufacturing site runs many production lines (i.e. produce multiple drug substances). The utilized capacity of a site for other drugs is captured as a baseload.

Model Formulation:

Sets:

M Set of manufacturing factories/sites
T Timeframe in years {2018...2025}
API Active pharmaceutical ingredient
DL Set of demand levels {1 = base demand, 2 = high demand, 3 = low demand}
S Stochastic scenarios within each demand level

Parameters:

SR Manufacturing success rate per site (stochastic per *M, T, API, DL, S*)
RW Number of production runs per week (stochastic per *M, T, API, DL, S*)
KGS Kilograms of API per production run (stochastic per *M, T, API, DL, S*)
D Drug substance requirement, in kilograms (stochastic per *T, API, DL, S*)
BaseUsage Utilized capacity of each site, in weeks

Decision Variables:

ThputM Non-negative variable to capture manufacturing amount, in kilograms (per *M, T, API, DL, S*)
SlackThput Non-negative variable to capture manufacturing volume in case extra capacity is needed, in kilograms (per *T, API, DL, S*)
ExtraThput Non-negative variable to capture manufacturing volume in case total capacity does not reach the minimum capacity level, in kilograms (per *T, API, DL, S*)
W Non-negative variable to capture site capacity utilization measured in weeks (per *M, T, API, DL, S*)
P Binary variable showing whether or not a site is used (per *M, T, API, DL, S*) (1 = the site is used for production; 0 = the site is not used for production)
Final_Weeks Non-negative variable showing the maximum of Weeks among all scenarios of a demand level set (per *M, T, API, DL*)
XW+ Non-negative variable captures the excess of 'Weeks+BaseUsage' from target value (i.e. 80% of 52 weeks) (per *M, T, API, DL, S*)
XW- Non-negative variable captures the slack of 'Weeks+BaseUsage' from target value (i.e. 80% of 52 weeks) (per *M, T, API, DL, S*)

Objective function:

$$\text{Min } \sum_{M,T,API,DL,S}(XW^{+}_{m,t,api,dl,s} + XW^{-}_{m,t,api,dl,s} + U1 * \text{P}_{m,t,api,dl,s}) + U2 * \sum_{T,API,DL,S}(\text{ExtraThput}_{t,api,dl,s} + \text{SlackThput}_{t,api,dl,s}) \quad (1)$$

where *U1* is a small penalty number that limits the total allocated number of sites for production; *U2* is a big penalty number for using extra capacity when the existing capacity is not maxed out or underutilizing a facility that creates non-negative slack capacity.

Subject to:

$$\text{W} = \frac{\text{ThputM(m,t,s,api,dl)}}{\text{SR(m,t,s,api,dl)} * \text{RW(m,t,s,api,dl)} * \text{KGS(m,t,s,api,dl)}} \quad (2)$$

$$\sum_{M} ThputM_{m,t,api,dl} \pm ExtraThput_{t,api,dl,s} \mp SlackThput_{t,api,dl,s} = D_{m,t,api,dl,s} \quad (3)$$

$$Minimum\ Target\ Capacity * P_{m,t,api,dl,s} \leq W_{m,t,api,dl,s} + BaseUsage \quad (4)$$

$$W_{m,t,api,dl,s} + \text{BaseUsage} \leq Site\ Full\ Capacity * P_{m,t,api,dl,s} \quad (5)$$

(where *P* is functional when *BaseUsage* = 0; i.e. if *W* = 0 & *BaseUsage* = 0, P = 0)

$$\text{Final\_Weeks}_{m,t,api,dl} \geq W_{m,t,api,dl,s} \quad (6)$$

$$W_{m,t,api,dl,s} - \text{Target Capacity} \leq XW^{+}_{m,t,api,dl} \quad (7)$$

$$\text{Target Capacity} - W_{m,t,api,dl,s} \leq XW^{-}_{m,t,api,dl} \quad (8)$$

(1) The objective function has three parts. Part one is capacity allocation through minimizing the deviation from the target capacity limit; part two is site selection by minimizing the site used; part three is demand fulfillment through minimizing the unsatisfied or excess demand respectively.
(2) Capacity conversion constraint describes the way XYZ Co. measures capacity of their manufacturing facilities. Capacity is measured in weeks through dividing the yearly production volume by the conversion factor -- runs per week multiplies kilograms per run multiplies success rate.

(3) Demand constraint limits the annual production volume to be as close to the annual demand as possible. If total *ThputM* -- production in kilograms -- exceeds demand, *ExtraThput* is positive; if it is under demand, *SlackThput* is positive. *SlackThput* and *ExtraThput* are auxiliary decision variables that are used to prevent infeasibility of the model in case demand cannot match exactly with the production volume.
(4) Upper capacity limit constraint: Site binary variable *P* is determined by capacity *W* and taken capacity *BaseUsage*. Only when *W* and *BaseUsage* are 0, *P* is 0.
(5) Lower capacity bound: to make sure P is 1 if the sum of $W_{m,t,api,dl,s}$ and *BaseUsage* is positive.
(6) Among all the scenario solutions, *Final_Weeks* takes the riskiest capacity usage among scenarios of each demand level set. This is to guarantee the robustness of the approximation for a capacity utilization risk i.e. overutilization.
(7) This defines the positive deviation from target capacity: W+BaseUsage > Target
(8) This defines the negative deviation from target capacity: Target > W+BaseUsage

Production Allocation and Site Selection

The model tries to limit the production to as few facilities as possible while fulfilling the annual demand requirement. This means that site that has the smallest deviation from the target capacity level will be allocated first. As a result, the model prioritizes production at sites that reach the capacity more easily. At the same time, the model also avoids the under-utilization of a facility.

Under high demand scenarios production sites are utilized more and the likelihood of reaching the upper capacity limit is higher. Alternatively, when demand is low, production sites seek to reach the lower capacity boundary and the potential feasibility of adding more production lines is higher. All three parameters – success rate, kilograms per run, and runs per week – impact the capacity utilization. Low values of manufacturing parameters lengthen the production time required for the facility to meet the demand levels. Alternatively, high values of manufacturing parameters reduce the production time requirements of the facility. When all three parameters are at low levels (Figure 1), production for drug X is at the highest risk of interruption or delay. As demand increases, the unfulfilled demand increases creating a backlog. When demand is at the upside level (Figure 1), all production facilities are fully utilized. Figure 2 shows the unfulfilled demand requirement under each demand scenario.

Figure 2 Unfulfilled demand under low manufacturing performance

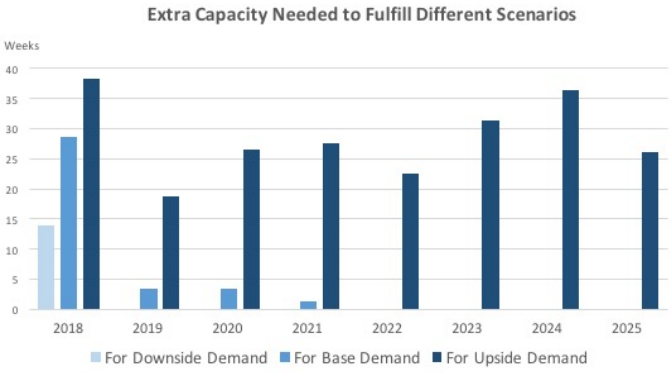


Figure 3 Extra capacity needed to fulfill different demand scenarios

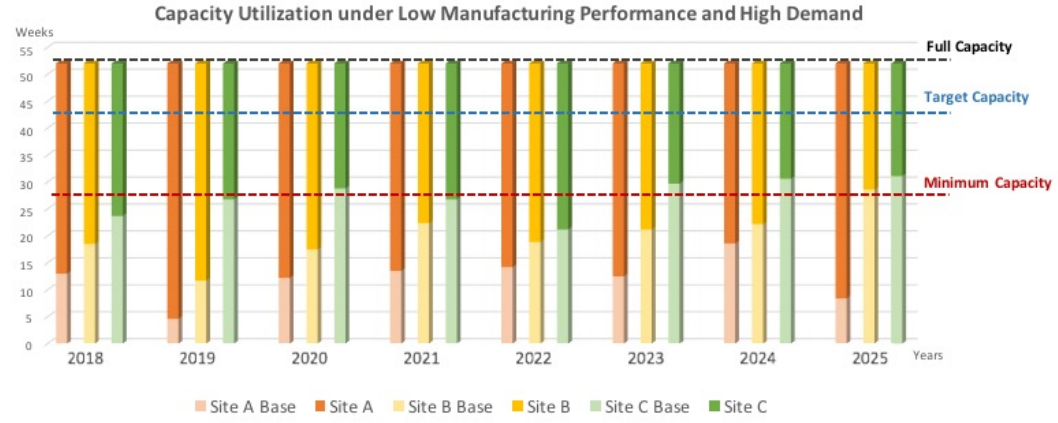


Figure 1 Under high demand, facilities run out of capacity when all three manufacturing parameters are low

The extra capacity needed to satisfy the unfulfilled demand is shown in Figure 3.

According to the sensitivity analysis, none of the manufacturing parameters are significantly different in regards to their capacity deviation from the base case scenario. This implies that no single parameter should be given sole priority with regards to facility performance analysis.

Conclusion

The model helps the company visualize capacity risk through experimenting with different manufacturing and demand scenarios. By capturing both the variability of in-house manufacturing parameters and demand, this research first provides a decision making tool for production facility selection and allocation in biologics raw material manufacturing. Additionally, this thesis uses this model to close the gap between manufacturing parameter analysis and strategic capacity planning in order to assist the long range planning department in communicating manufacturing related risk within the company.