MAS 547

Project 3

Jiahao Chang, Chen Xie, Xuanyi Lu, Pengzi Xu

PART 1

Background

JD is a Chinese online shopping platform, and it was officially launched in 2010. Based on the first half of 2015, Analysys.cn, a credible business information service platform that reflects the economic development of Chinese new media, released a special research report on China's internet online shopping app, which showed that JD's overall online shopping market accounted for 31.24% as the second place. According to a mobile data analysis report, JD has exceeded 10 million app downloads. Because the increase in ordering that makes the volume of the JD increasing, costs are increase as well, for labour cost, delivery, and packaging and so on. The old model will not be efficient, and we need to do some optimization to it.

The problem and the motivation

The main task of the product sorting system is to separate the products to be sorted by types of information based on the customer order data, the product with a same ID will send to the same storage location for the next process. The efficiency of product sorting has a great influence on the production and distribution of merchants, especially the labour cost. Therefore, we made this simple model to find out the problems and made optimization and improvement through the adjustment of parameters to make merchants achieve the optimal efficiency.

The goal of the study

The goal of this research is to explore the refinement and analysis of the transaction product sorting process, analyse the generated data, find which part could be optimized in the model and problems of the original sorting process, and finally optimize the process to discover the most effective sorting and delivery progression after optimization.

PART 2

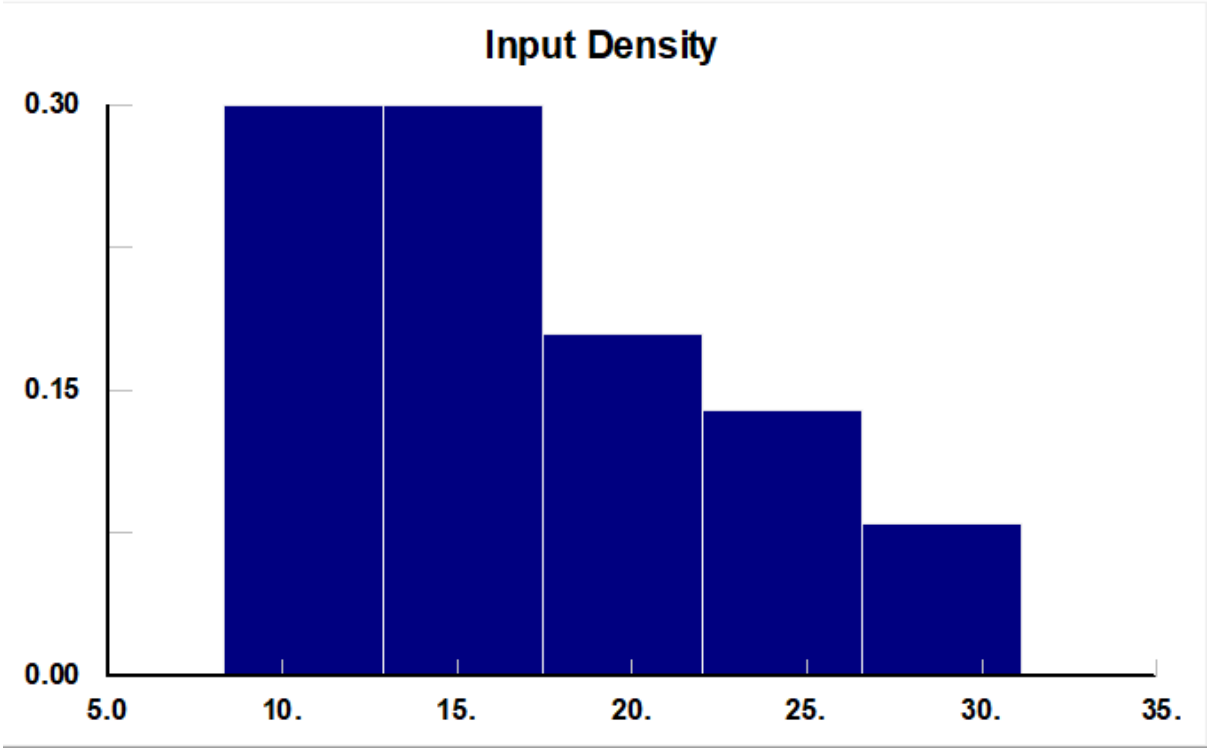

Figure 1

We will use Stat: Fit visualized data of JD. We will set 50 inputs of Stat: Fit. Those data are the processing time of packing of an order.

First, we will look at the distribution of the data. The lower bound cannot be zero, this is because a zero-minute packing time of an order does not make sense. In figure 1, the minimum packing time is around 7, which means the minimum time it takes to process an order until it goes to the next step is about 7 minutes. The lower bound shown in the figure 1 is 5 minutes per order, which means it won't take less than 5 minutes to packing an order. Normal distribution has a skewness of 0 because that is symmetrical, in the JD case, we can predict that there is a skewness for this distribution. The fat tail is on the left and the longer tail is on the right, the distribution seems to have a right skewed.

descriptive statistics

data points	50
minimum	8.33
maximum	31.15
mean	16.6914
median	15.18
mode	12.65
standard deviation	6.09434
variance	37.1409
coefficient of variation	36.5118
skewness	0.616507
kurtosis	-0.603415

Figure 2

Descriptive statistics would provide more statistical information to our analysis. Skewness is a measure of the asymmetry of the probability distribution of a random variable about its mean. In other words, skewness tells you the amount and direction of skew. According to the general rule of thumb, If skewness is less than -1 or greater than 1, the distribution is highly skewed. If skewness is between -1 and -0.5 or between 0.5 and 1, the distribution is moderately skewed. If skewness is between -0.5 and 0.5, the distribution is approximately symmetric. In our JD case, the skewness is equal to 0.62, that means the distribution is moderately skewed. The mean of 16.69 is more than the median of 15.18, which proves the distribution skewed to the right. Kurtosis tells the height and sharpness of the central peak, relative to that of a standard bell curve. Our case has a kurtosis of -0.6034, that means the bell curve will have a more uniform, flat, look than the normal distribution.

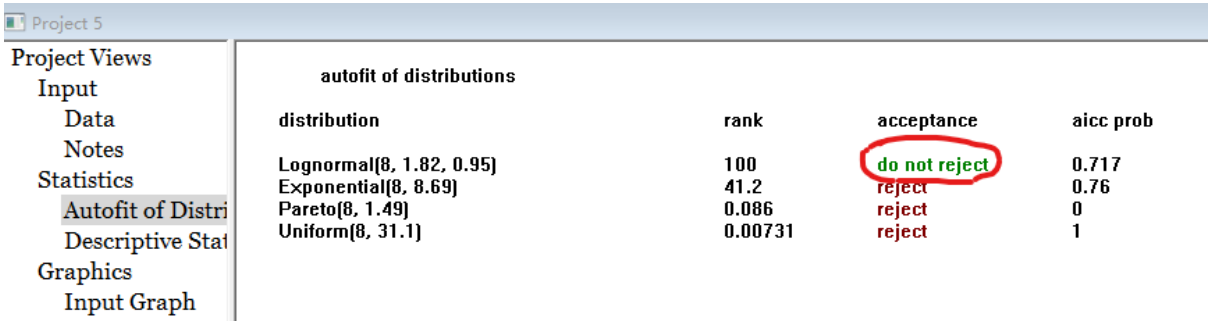

Figure 3

Normal distribution should not be fit in our case, this is because we have set the lower bound to the distribution. Figure 3 shows which distribution would be the choice we use in the Simio. Exponential distribution, Pareto distribution, and uniform distribution all reject our

non-hypothesis (the population follows the distribution.). Lognormal distribution will be our best choice to implement.

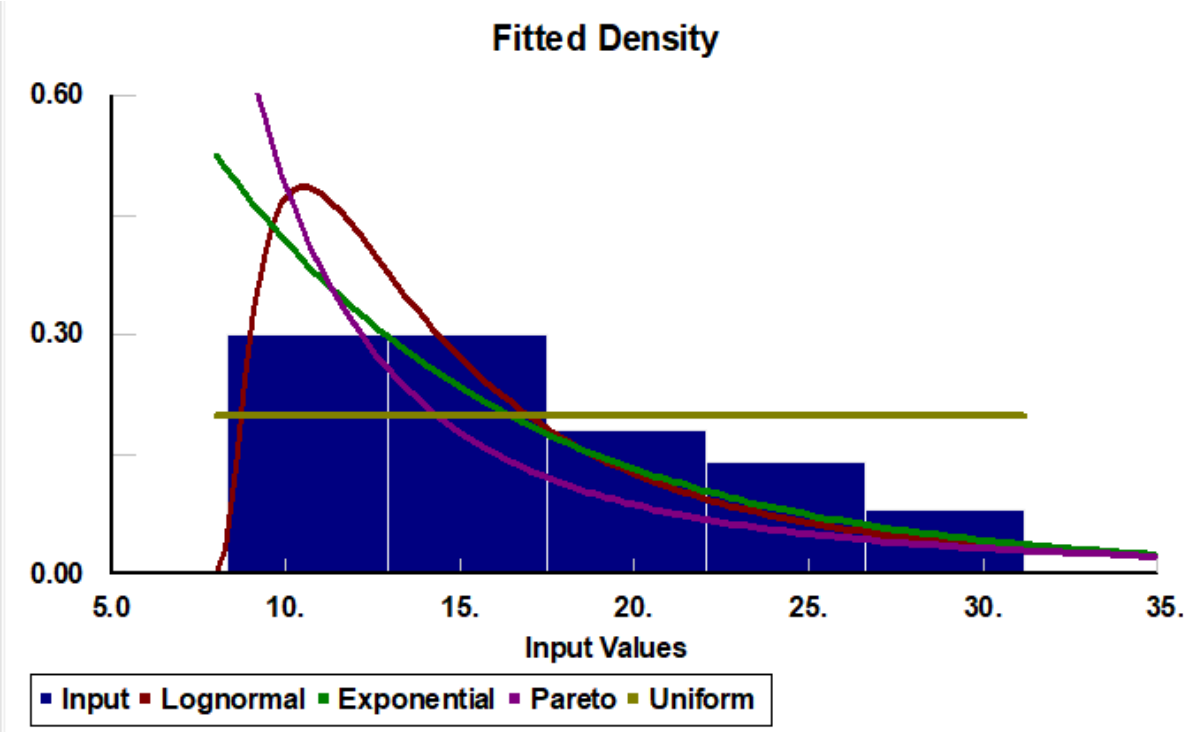

Figure 4

By taking a comprehensive view of the distributions, figure 4 would explain why the lognormal is the best choice. The red line is the lognormal distribution, it matches the shape of our right-skewed JD distribution.

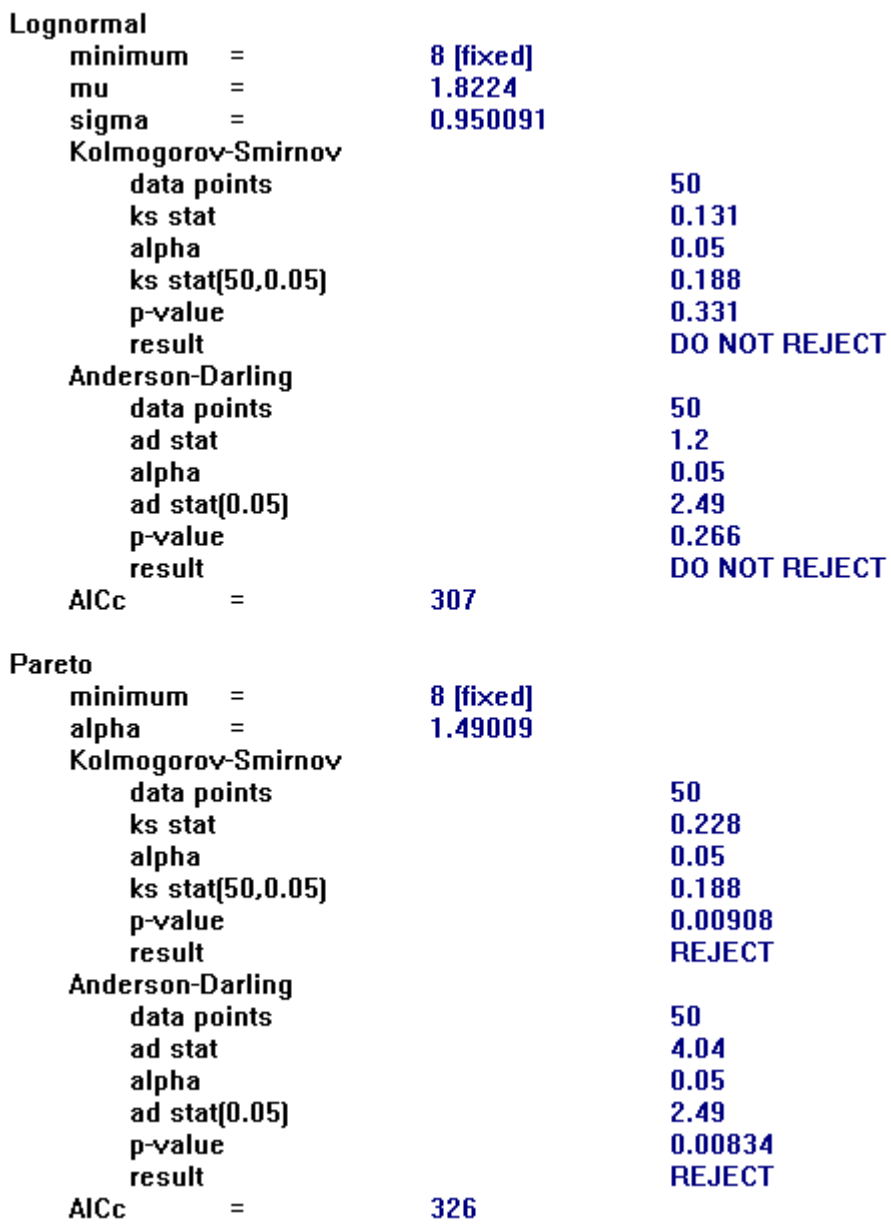

Figure 5

Figure 5 shows the goodness of fit test. P-value gets larger will decrease the significance. Lognormal distribution has the highest p-value among other distributions. That is the reason why lognormal distribution is the best.

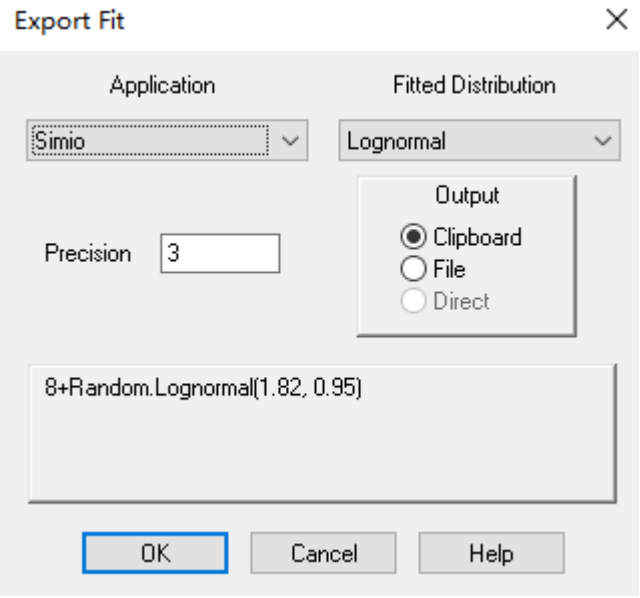

Figure 6

Figure 6 is the format we will use in the Simio model.

PART 3

Flowchart

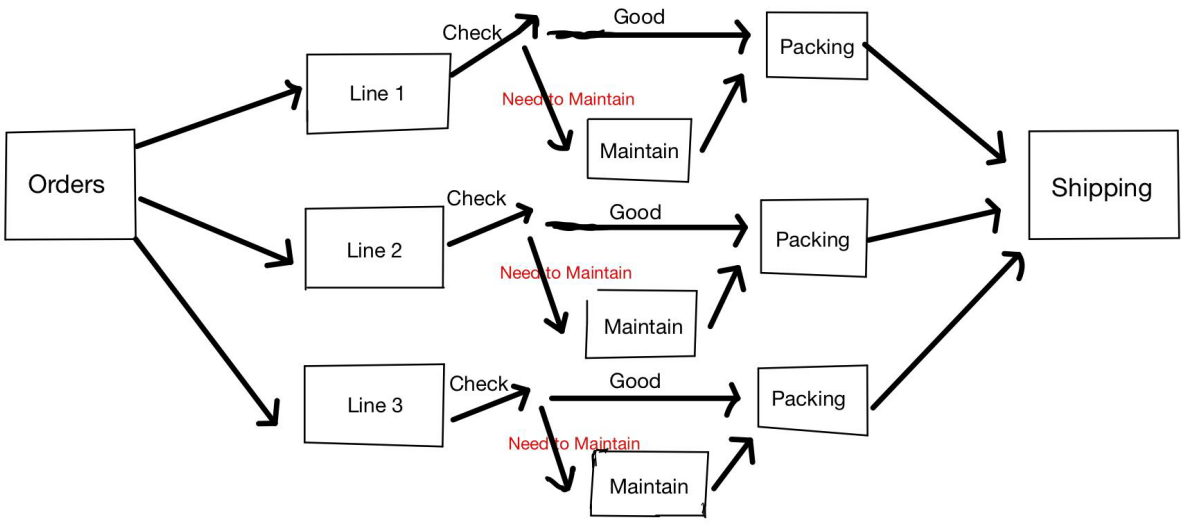

Figure 7

The original hypothetical process based on the problem that There are three types of products in the outlets, and we name they A, B, C in the model. The first thing we must distinguish which type of type pf the product (A or B or C) that customer wants, when customers place an order. The storage of products would be damaged by humid environment. If the condition of the product is good, it will be sent to packaging station after check. Then the product will be collected and shipped.

Potential improvement

Potential improvement 1: It is possible to increase maintain and packaging workstations' utility rate. Therefore, we are considering combined the corresponding maintaining and packaging for each product line to reduce at least 3 people and increase utilization.

Potential improvement 2: Due to the product backlog of Check place 1; thus, rising the initial capacity from 1 to 2. The increased capacity would reduce its utilization but can effectively decrease the number of products for 1, which can lead to a higher productivity or faster cash flow for the overall model.

PART 4

Simulation Model

Our SIMIO simulation model will have three resources and three entities that is because different of arrival rate among each product, so we must perform it, otherwise it won't be present. We have nine servers those will stand for check place, maintain, and pack workstations in the model, and one sink. Path will be used to make connections instead of connectors that is because the distance among stations and delivery time are not counted.

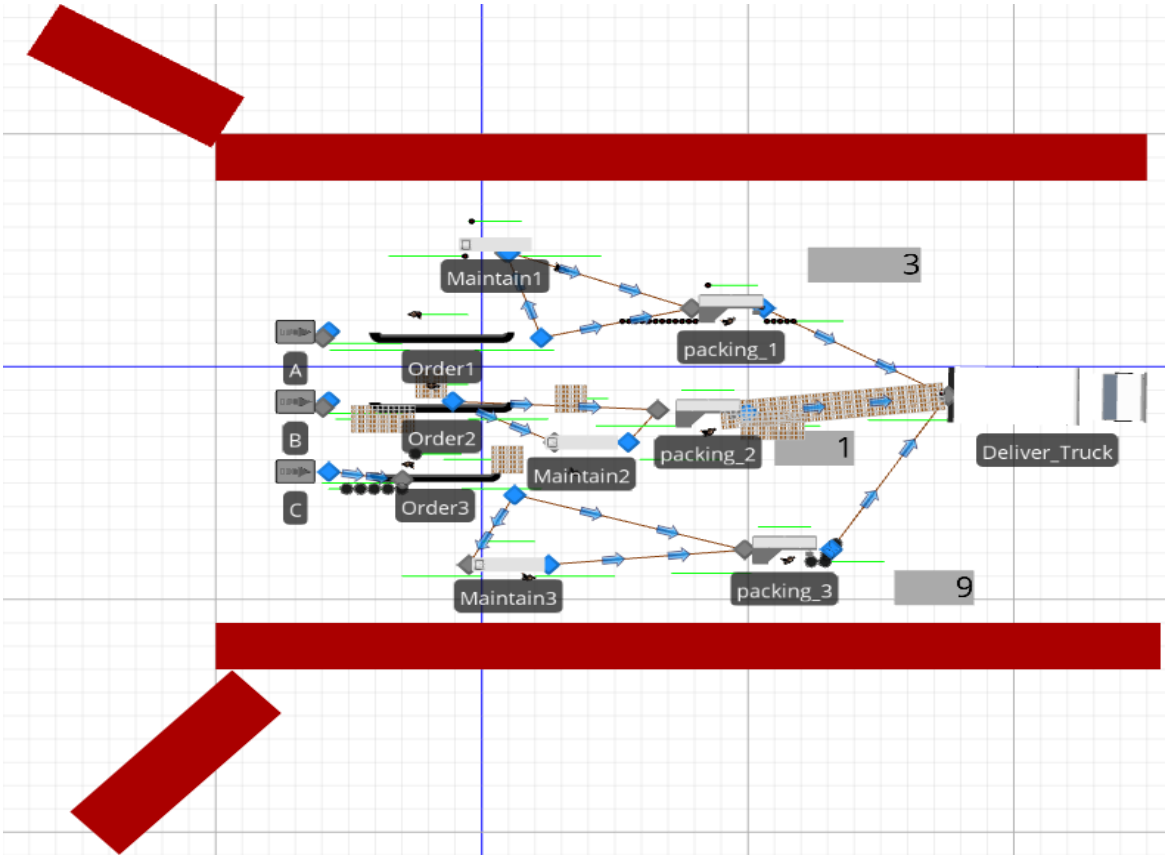

Figure 8

For products A, B, and C, because we have not found certain real-world data by finding distribution, but during the research, we found the potential distribution of A would be exponential with a mean of 16 per minutes; the potential distribution of B would be uniform with a range between 4 and 11 per minutes; the potential distribution of C would be triangular with range of 4 and 14, mean of 11. The processing time of packing station is 8 + Random.Lognormal(1.82 , 0.95) based on the goodness-of-fit finding. Since there are branches after the check place, the products to be inspected enter the packaging station and

the maintain station respectively according to a certain proportion, so the output node of the check point needs to be output according to the weight, and the set proportion to the products in the corresponding path.

PART 5

steady-state queueing theoretic results

Setting the running time as 1000 hours to get the result in figure 8.

For the three entity objects: based on figure 8, apparently, order line 1, total of 3958 pieces were processed, and 3360 pieces were processed. The average product retention time was 68.05 hours, and the maximum retention time was 155.69 hours. The average number of Order1 is 271, and the maximum is 638. For order line 2, a total of 8006 pieces were processed, 6706 pieces were processed, the average product retention time was 76.97 hours, the maximum retention time was 166.06 hours, the average quantity was 622.02 pieces, and the maximum was 1310 pieces. For order line 3, there were 6206 pieces in total, 5420 pieces were processed, the average product retention time was 65.57 hours, the maximum retention time was 130.36 hours, the average quantity was 405.72 pieces, and the maximum was 796 pieces.

Object Type	Object Name	Data Source	Category	Data Item	Statistic	Average Total
ModelEntity	OrderA	[Population]	Throughput	NumberCreated	Total	3,958.0000
				NumberDestroyed	Total	3,360.0000
			FlowTime	TimeInSystem	Average (Hours)	68.0501
					Maximum (Hou...	155.6905
					Minimum (Hou...	0.3521
					Observations	3,360.0000
			Content	NumberInSystem	Average	271.0011
					Maximum	638.0000
	OrderB	[Population]	Throughput	NumberCreated	Total	8,006.0000
				NumberDestroyed	Total	6,706.0000
			FlowTime	TimeInSystem	Average (Hours)	76.9663
					Maximum (Hou...	166.0559
					Minimum (Hou...	1.3551
					Observations	6,706.0000
			Content	NumberInSystem	Average	622.0229
					Maximum	1,310.0000
	OrderC	[Population]	Throughput	NumberCreated	Total	6,206.0000
				NumberDestroyed	Total	5,420.0000
			FlowTime	TimeInSystem	Average (Hours)	65.5731
					Maximum (Hou...	130.3624
					Minimum (Hou...	0.0997
					Observations	5,420.0000
			Content	NumberInSystem	Average	405.7213
					Maximum	796.0000

Figure 9

As from the above table, each order line has basically reached 100%, which means that it has a lot of work to do, has never stopped, has a surprising amount of cache, but there were not processing all incoming products, resulting in more work-in-progress and accumulating in the system. For order line 3, this conveyor is the bottleneck of order line 3 because it takes up time, space and money. We had regarded it would be the fastest line, but it was contrary to our expectation and accumulated plenty of products in line. Since the packaging time of maintaining node is very short, the utilization rate is very low, resulting in a waste of working capacity. These are the two main problems of the whole system, and we will improve them in the future.

PART 6 & 7 SHOW IN MODEL

PART 8

The recommendation we assume to improve the system:

Potential improvement 1

Consider redesigning the three maintaining station to increase the number of processing time and reduce the pressure on packing station.

Process Logic	
Capacity Type	Fixed
Initial Capacity	1
Ranking Rule	First In First Out
Dynamic Selection Rule	None
▸ Transfer-In Time	0.0
Process Type	Specific Time
▸ Processing Time	**60**
Off Shift Rule	Suspend Processing
▸ **Other Processing Options**	

Process Logic	
Capacity Type	Fixed
Initial Capacity	1
Ranking Rule	First In First Out
Dynamic Selection Rule	None
▸ Transfer-In Time	0.0
Process Type	Specific Time
▸ Processing Time	**60**
Off Shift Rule	Suspend Processing

Process Logic	
Capacity Type	Fixed
Initial Capacity	1
Ranking Rule	First In First Out
Dynamic Selection Rule	None
▸ Transfer-In Time	0.0
Process Type	Specific Time
▸ Processing Time	**60**
Off Shift Rule	Suspend Processing

Figure 10

We reset three maintaining station processing time to the same, 60 minutes, it was 70 min, 60 min, 55 min. Although the move likely increases the cost, where it will lead the process to reduce overstocking at packing space.

Potential improvement 1

Since the product backlog of packing 3 at order line 3 is severe and it is a bottleneck workstation, so we changed its working capacity from 1 to 2

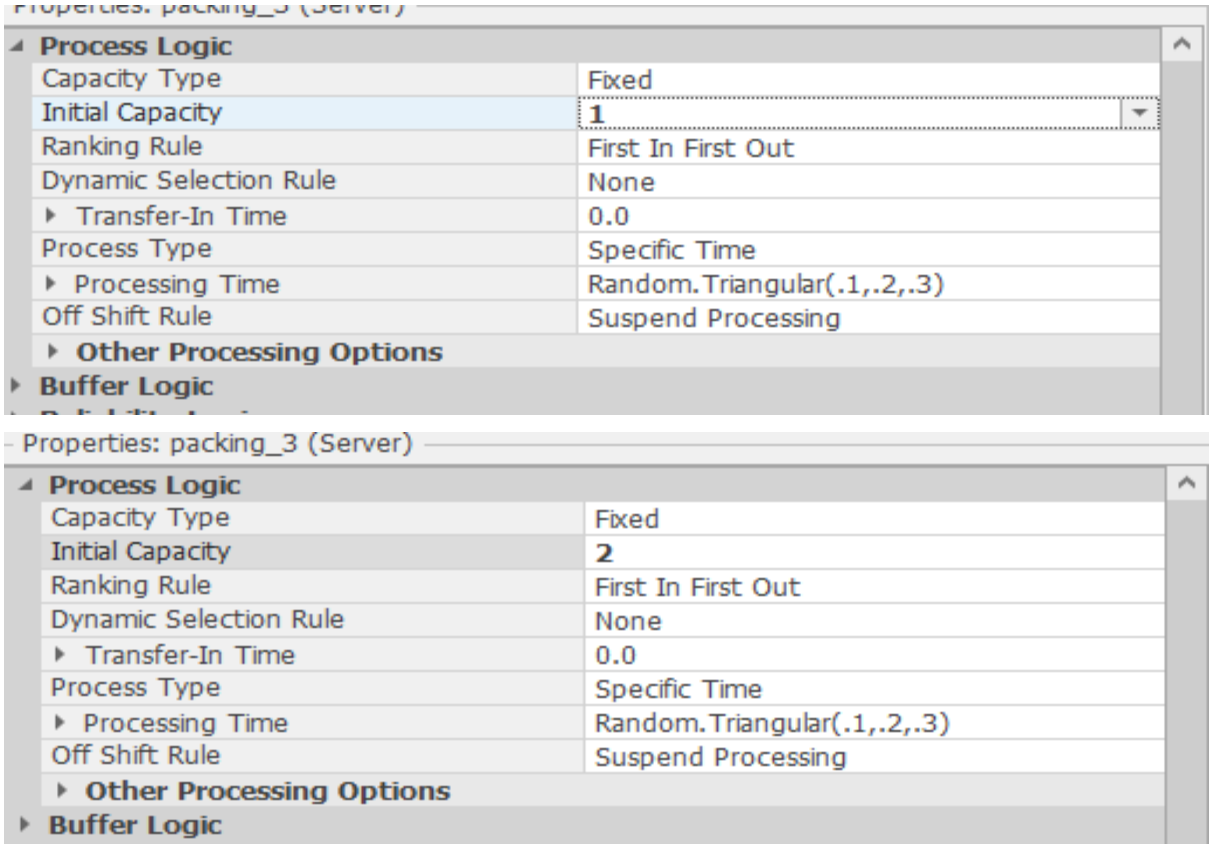

Figure 11

This way, the pressure from the checkpoint to get a substantial link, although will reduce the utilization rate, but can greatly reduce the product backlog, thus increasing productivity, faster capital turnover and system efficiency. The label expression showed from 9 go down to 4.

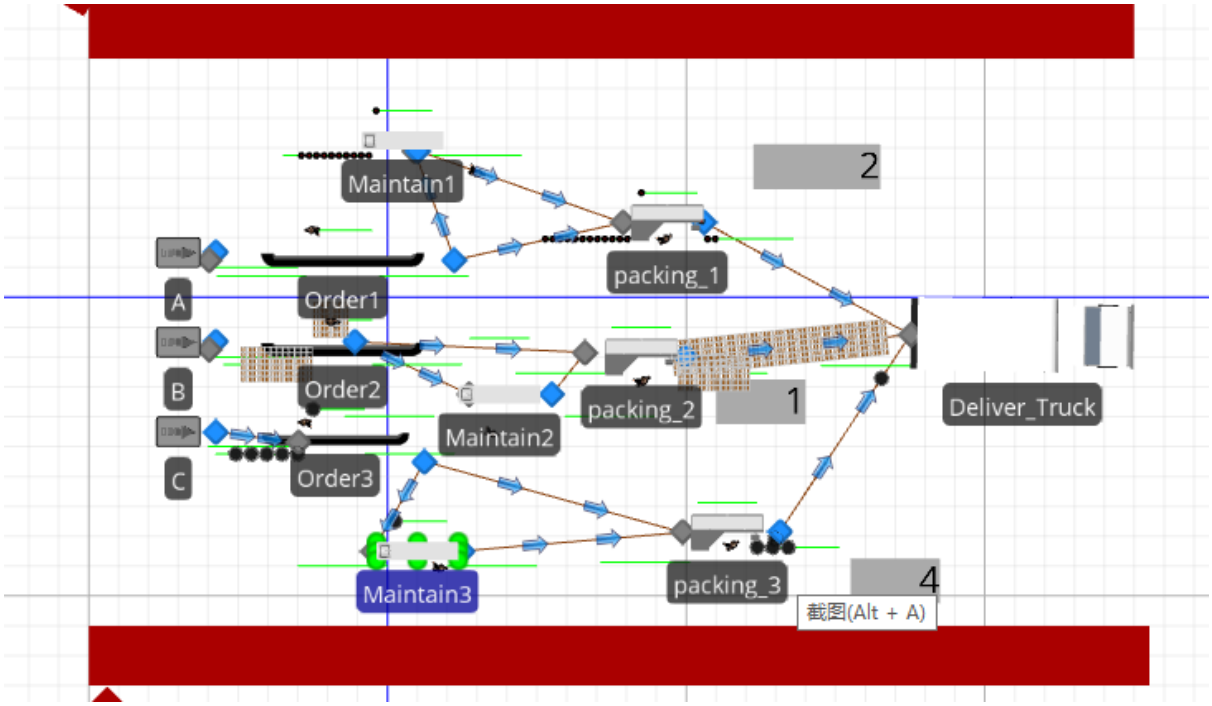

Figure 12

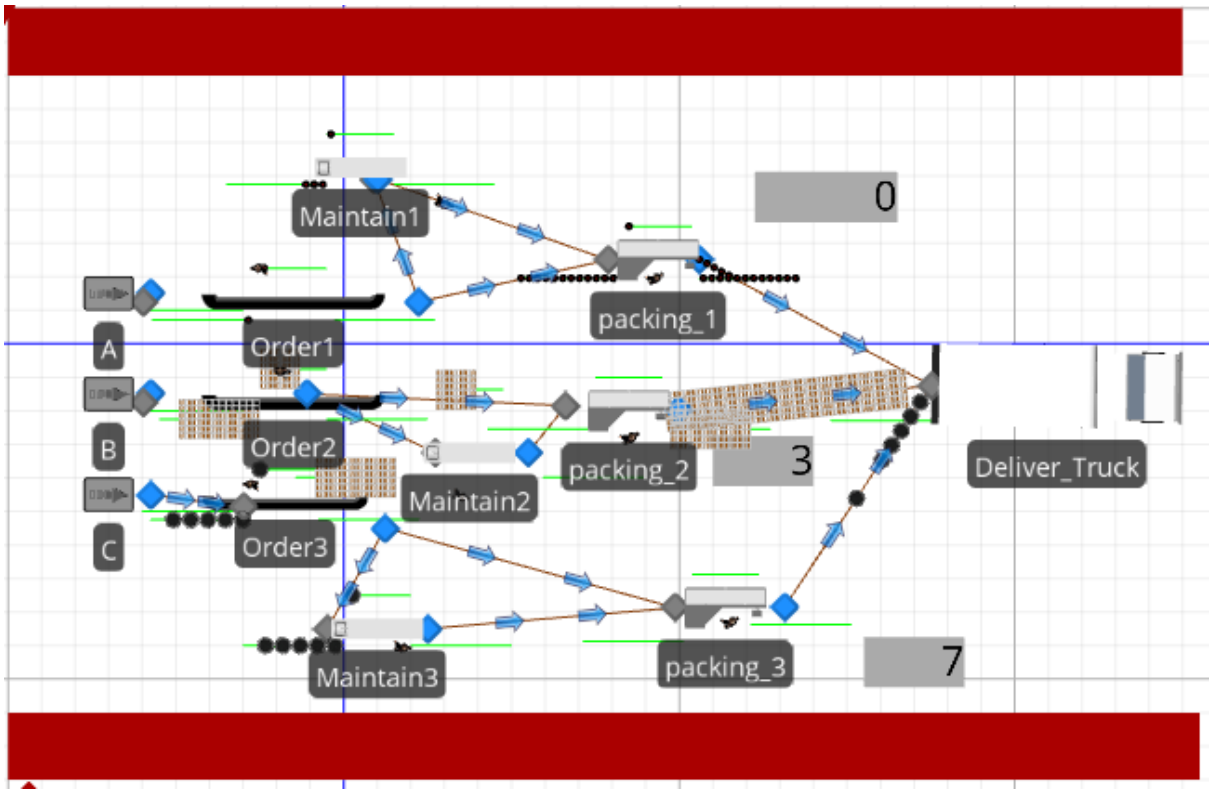

Figure 13

Compared with the original model, it solves the problem of product backlog as far as possible, solves the problem of low utilization rate of packaging station through the way of changing the time to the same processing time, and reduces the packing cost and improves the output.