Operationalizing Demand Forecasts in the Warehouse

By Dan Li and Kyung Kim
Thesis Advisor: Dr. Bruce Arntzen

Topic Area: Forecasting

Summary: This research evaluates the plausibility of leveraging the SKU level forecast to predict equivalent operational activities in the warehouse. We analyze rolling base forecasts of outbound shipment with actual picking data from the warehouse. The thesis concludes with the presentation of results of the evaluation discussions regarding the rolling base forecast and the potential areas to improve the accuracy. This work will help warehouse managers to understand the potential of utilizing demand forecasts to plan warehouse activities.

KEY INSIGHTS

1. The forecast accuracy and bias gets better as we approach closer to the actual observed week.

2. Forecasting errors of Unit was not highly correlated with those of Picking Labor Hours, even though the trend was similar between the two.

Introduction

Like many other companies in the industry, CVS Health uses time series-based, forward-looking forecasting for demand planning and inventory management. Such methods usually forecast demand at the retailer store and SKU level before aggregating demand to the supplier and distribution center level. The aggregated demand is a key indicator for supply chain operations: it helps distribution centers plan activities, and make financial decisions. To leverage the store and SKU level forecast, we examined the key drivers of distribution center operation and management. The ultimate goal of this thesis project is to evaluate the potential use of translating the demand forecast into expected activities in the warehouse.

Forecasted Units and Hours

With the numerous combination of stores, DCs, and SKUs available, we scoped the effort in order to reduce the amount of data required as well as the potential complexities in attempting to model the entire network. For the purpose of the model, we limited the efforts in the following way:

- A single DC
- Regular items only (non-seasonal, non-promotional)
- Full Case and Split Case picking activities
- Two SKU categories (Allergy Remedies and Laundry)
The major CVS DC was selected by the team because of the team's familiarity with the DC and access to data. The scope was limited to regular items in order to remove the higher demand variability that typically exists with seasonal or promotional items. There are seven distinct activities occurring in a CVS DC: receiving, machine put away/replenishment, manual replenishment, split-case (SC) picking, full-case (FC) picking, shipping, and support. Of these seven activities, SC picking and FC picking were selected as part of the scope, because these activities are some of the most labor intensive in the DC. The other reason is because the costs data available for these activities had only a limited amount of overhead included. The limited amount of overhead ensured that the model results would be more in line with the actual costs of the activities themselves exclusive of allocated overhead.

The Allergy Remedies and Laundry SKU categories were selected because they aligned well with the SC and FC picking activities. Most of the products in the Allergy Relievers categories were split-case while most of the products in the Laundry category were full-case. This was important such that from an aggregate category level, the picking activities are consistent. A mix of SC and FC picking activities within a category would have required extensive and time-prohibitive manual data clean up in order to separate the SC and FC SKUs and pickings from the warehouse labor management system data.

Both of the SC and FC pick rates were provided by our partner, and their values were as follows:

- **Split Case Pick Rate = 920 pieces/hour**
  (applied to Allergy Remedies)

- **Full Case Pick Rate = 851 pieces/hour**
  (applied to Landry category)

These forecasted pick rates were divided by the forecasted demand in units to calculate the total forecasted labor hours.

**Actual Units and Estimated Actual Hours**

The available actual picking data from the WMS was limited to the actual number of units picked per SKU on each day. In order to evaluate the cost model in predicting the number of warehouse labor hours, we needed actual picking time for comparison. However, this specific measure was not tracked in the WMS, and we needed to develop an estimation based on the number of units picked and other measure provided by our partner. This estimation was based on several additional data points provided by our partner:

- Actual picking time for all the orders containing the specific SKU (A) (the total actual amount of time that it took to fulfill all the pick orders that contained the SKU)
- Estimated picking time for all the orders that contained the specific SKU (B) (the total estimated amount of time that it should have taken to fulfill all the pick orders that contained the SKU, based on a “level seconds” metric developed for each SKU by the partner)
- Estimated picking time to pick all the units of the specific SKU (C) (the total estimated amount of time that it should have taken to pick only the specific SKU across all orders, also based on the “level seconds” metric developed for each SKU by the partner)

Using the metrics above, the estimated actual picking time was calculated by the following formula:

\[
\text{Estimated actual picking time} = \left( \frac{\text{Actual picking time for all orders containing the specific SKU}}{\text{Estimated picking time to pick all the units of the specific SKU}} \right) \times \text{Estimated picking time for all the orders that contained the specific SKU}
\]

The proportion of the estimated picking time attributable to the specific SKU is:

\[
\text{Proportion of the estimated picking time attributable to the specific SKU} = \frac{\text{Estimated picking time to pick all the units of the specific SKU}}{\text{Estimated picking time for all the orders that contained the specific SKU}}
\]

Making the final equation for estimated actual picking time:

\[
\text{Estimated actual picking time} = \left( \frac{\text{Actual picking time for all orders containing the specific SKU}}{\text{Estimated picking time to pick all the units of the specific SKU}} \right) \times \text{Estimated picking time for all the orders that contained the specific SKU}
\]

The estimated actual hours were used to compare to the forecasted hours.
Results

Figure 1 summarizes the results of the units and hours comparisons organized by the number of weeks in advance.

<table>
<thead>
<tr>
<th>Actual Week</th>
<th>Forecast Week</th>
<th>Weeks in Advance</th>
<th>Units</th>
<th>Hours</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>RMSE</td>
<td>MAPE</td>
</tr>
<tr>
<td>11</td>
<td>11</td>
<td>0</td>
<td>76.97</td>
<td>59%</td>
</tr>
<tr>
<td>12</td>
<td>11</td>
<td>1</td>
<td>174.93</td>
<td>84%</td>
</tr>
<tr>
<td>13</td>
<td>11</td>
<td>2</td>
<td>270.00</td>
<td>220%</td>
</tr>
<tr>
<td>14</td>
<td>11</td>
<td>3</td>
<td>236.55</td>
<td>160%</td>
</tr>
<tr>
<td>15</td>
<td>11</td>
<td>4</td>
<td>264.22</td>
<td>150%</td>
</tr>
<tr>
<td>16</td>
<td>11</td>
<td>5</td>
<td>281.14</td>
<td>54%</td>
</tr>
<tr>
<td>17</td>
<td>11</td>
<td>6</td>
<td>352.71</td>
<td>312%</td>
</tr>
<tr>
<td>18</td>
<td>11</td>
<td>7</td>
<td>482.02</td>
<td>103%</td>
</tr>
<tr>
<td>19</td>
<td>11</td>
<td>8</td>
<td>371.98</td>
<td>66%</td>
</tr>
</tbody>
</table>

Figure 1 Statistical Results

Figure 2 charts the results of comparing the demand forecast and the actual picking data for units aggregated across all the SKUs in our data set. The x-axis represents the number of weeks in advance of the forecasts. The mean absolute percent error (MAPE) is a measure of the accuracy of the forecast, while the mean percent error (MPE) is a measure of the forecast bias. As the number of weeks in advance gets closer to zero, we see that each of the metrics converge towards zero. This suggests that the accuracy of the forecast improved as the target week approached.

Figure 2 Results of Units Comparison

Figure 3 charts the results of comparing the labor hours forecast and the estimated actual hours data aggregated across all the SKUs in our data set. The x-axis represents the number of weeks in advance of the forecasts. Again, as the number of weeks in advance gets closer to zero, we see that each of the metrics converge towards zero. This suggests that the accuracy of the forecast improved as the target week approached.

Figure 3 Results of Hours Comparison

Conclusions

The purpose of this thesis was to evaluate the demand forecast data in terms of expected unit demand and expected picking labor hours in the warehouse. Our study indicated that the demand forecast is weakly correlated with the actual unit number or actual picking labor hours. For picking higher than forecast, the deviation could come from additional order, either placed by store or corporate. For forecast higher than picking, the deviation could be due to inventory insufficient on mainframe, or because the warehouse do not fulfill total quantity. Our research warrants a more thorough examination of the warehouse activities to identify the root cause accounting for the high deviation between the forecast and the actual.

For further information on the research contact Dr. Bruce Arntzen, Executive Director, MIT Supply Chain Management Program, at: barntzen@mit.edu.