Operationalizing Demand Forecasts in the Warehouse

By

Dan Li
PhD in Genetics, University of Rochester, 2008
M.S. Genetics, Fudan University, 2003
B.S. Biology, Wuhan University, 2000

and

Kyung Kim
B.S. Electrical Engineering, Tufts University, 2004

SUBMITTED TO THE ENGINEERING SYSTEMS DIVISION IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ENGINEERING IN LOGISTICS

at the

Massachusetts Institute of Technology

June 2015

Signature of Author: .... Signature redacted ....
Master of Engineering in Logistics Program, Engineering Systems Division
May 10, 2015

Signature of Author: .... Signature redacted
Master of Engineering in Logistics Program, Engineering Systems Division
May 10, 2015

Certified by: ..... Signature redacted ....
Dr. Bruce Arntzen
Executive Director, Supply Chain Management Program
Thesis Supervisor

Accepted by: .... Signature redacted ....
Dr. Yossi Sheffi
Director, Center for Transportation and Logistics
Elisha Gray II Professor of Engineering Systems
Professor, Civil and Environmental Engineering

Operationalizing Demand Forecasts in the Warehouse

by

Dan Li

and

Kyung Kim

Submitted to the Engineering Systems Division
on May 10, 2015 in Partial Fulfillment of the
Requirements for the Degree of Master of Engineering in Logistics

Abstract

Demand planning affects the subsequent business activities including distribution center operational planning and management. Today's competitive environment requires distribution centers to rapidly respond to changes in the quantity and nature of demand. For the distribution center, accurate forecasts will help managers to accordingly plan operational activities.

In the present thesis, we evaluate the plausibility of leveraging the SKU level forecast to predict equivalent operational activities in the warehouse. Through literature review, we identified the key drivers in distribution center operation and management. We further chose outbound shipment picking time as our measurement to perform the evaluation between demand forecast and actual warehouse shipments. The thesis concludes with the presentation of results of the evaluation discussions regarding the rolling horizon based forecast and the potential areas to improve the accuracy. This work will help warehouse managers to perform root cause analysis to examine the discrepancy between the units/labor forecast and actual units/labor. Our work will also help warehouses achieve greater operational efficiency.

Thesis Supervisor: Dr. Bruce Arntzen
Title: Executive Director, Supply Chain Management Program

Acknowledgments

I would like to thank my family, friends, and fellow SCM students for their support and encouragement.

- Dan Li

I would like to thank my family, friends, MIT faculty, and fellow SCM students for supporting me through this experience.

- Kyung Kim

We both wish to thank Dr. Bruce Arntzen, Dr. Stephen C Graves, and CVS Health for generously assisting us.

Contents

List of Figures

List of Tables

1. Introduction

The present thesis explores the feasibility of translating SKU-level demand forecast into distribution center operations metrics to facilitate work planning. The following section discusses the overall topic, underlying research motivations and the research goal.

1.1. Research Motivation

CVS Health (CVS) uses sixteen distribution centers (DC) to replenish all of its 7,600 stores. Store-SKU level demand data is collected on a daily basis. The data is used to develop weekly demand forecasts for over 160,000 SKUs in terms of number of units and product values. The forecasts are updated weekly and used to drive ordering and replenishment operations. The demand forecasts are aggregated at the DC by aggregating the stores that each DC supports. Despite having the DC-level aggregated demand forecasts, CVS has not yet refined a tool to forecast the labor requirements in the distribution center at the SKU level for a forward looking multi week forecast.

1.2. Research Goal

Like many other companies in the industry, CVS uses time series-based forward-looking forecasting for sophisticated inventory replenishment systems. 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 need to thoroughly examine the key drivers of distribution center operation and management. The ultimate goal of this thesis project is to evaluate the demand forecast data in terms of expected unit demand and expected picking labor hours in the warehouse.

2. Literature Review

This chapter provides a review of distribution center operations and planning and demand forecasting approaches. First, we briefly discuss forecast accuracy versus the time horizon in Section 2.1. Then we review in Section 2.2, the application of activity based costing system in organizations. In Section 2.3, we describe a few of the factors found through our research that drive inefficiencies in warehouse activities. In Section 2.4, we discuss the distribution center planning methods. In Section 2.5, we examine multiple factors in distribution center operations which may be key components in the transformation system translating the demand data into operational planning metrics. In particular, we focus on order picking, cube size and workforce flexibility because they are measurable and can be included in a mathematical model. Last, in Section 2.6, we summarize the findings.

2.1. Forecast Accuracy and Time Horizon

It is a generally accepted truism in demand forecasting that one can predict what will happen next week with much more accuracy than what will happen next month (Inman, n.d.). For example, Metin Çakanyıldırım, professor of Operations Management at the University of Texas at Dallas, teaches that "forecast accuracy decreases as time horizon for forecasts increase (Çakanyıldırım, n.d.)." Prior research by Schoenmeyr and Graves examined this more quantitatively with results showing increasing correlation between forecast and actual for an electronic test system product as the weeks in between decreased (Schoenmeyr & Graves, 2009).

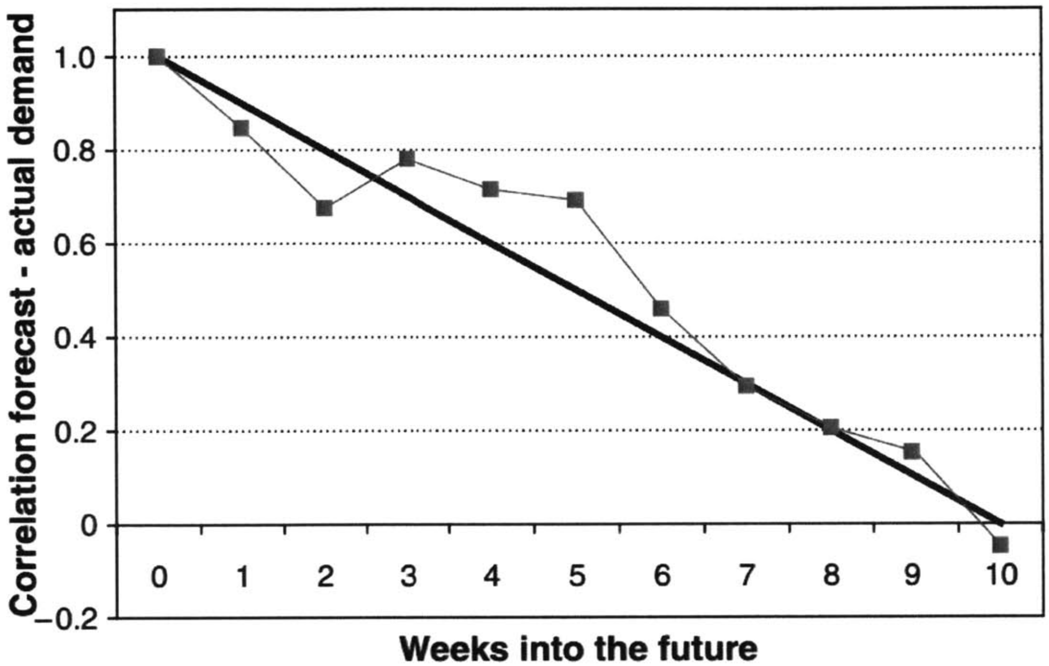

Figure 2-1 Correlation between Forecast and Actual versus Weeks into the Future (Schoenmeyr & Graves, 2009)

Since more demand volume equates to more units moved, it makes sense that demand forecasts could be used to derive equivalent forecasts of warehouse activities. The rest of the literature review summarizes the prior work in translating demand into warehouse activities.

2.2. Activity Based Costing

The sections below discuss the prior work in activity based costing as applied to supply chain management and more specifically the warehouse.

2.2.1. Activity Based Costing in Supply Chain Management (SCM)

Activity based costing (ABC) is a costing methodology that identifies activities and assigns the cost of each activity with resources to products and services. Based on the relationship between costs, activities and products, this system uses cost drivers to trace costs of the activities to the products that consume the resources used in these activities. This system assigns indirect costs to products less arbitrarily than traditional methods, and hence offers product costs more accurately than traditional cost systems.

Supply chain management (SCM) is a key component of competitive strategy to enhance organizational productivity, performance and profitability. Many techniques, including just-in-time (JIT), lean production (LP), computer generated enterprise resource planning schedule (ERP), and activity based costing, are widely implemented to make better use of SCM. Since every aspect of decision making process in SCM requires cost data, activity based costing can help managers make important strategic business decisions.

Since activity based costing was first introduced about 30 years ago (Kaplan, 1986), a long list of literature indicate that activity based costing can contribute to SCM and organizational performance from many different perspectives. Satoglu et al uses the activity based costing model to measure the cost of non-value adding activities for both central and decentralized mini-storage facilities, providing a useful approach and decision support for the companies to convert to decentralized mini-storages (Satoglu et al., 2006). Baykasoğlu and Kaplanoğlu demonstrate how effective the activity based costing method can be, combined with business process modeling and analytical hierarchy approach, in costing services of the land transportation company (Baykasoğlu and Kaplanoğlu, 2008).

However, despite of the critical role of activity based costing in improving the organizations' performance, the adoption of activity based costing is not highly prevalent. One explanation of the relatively low implementation is that decision makers remain unconvinced that whether ABC's advantages over traditional accounting techniques are high enough to pursue them to implement ABC in practice (Askarany, Davood and Yazdifar, Hassan, 2007).

2.2.2. Activity Based Costing in Warehouse Management

Another fact is that even though activity based costing has been used in SCM field, a great majority of the reported activity based costing applications have been limited to manufacturing. The explanation is that the main benefit of activity based costing lies in its ability of providing more accurate cost information through effective cost allocation for products which consume different amount of overhead costs. Such overhead costs can be at different levels of cost hierarchies such as unit level, batch level, product level and facility level. The implementation of these cost hierarches are more in line with complex manufacturing products rather than non-manufacturing environment.

Through literature review, we find very little report discusses how the ABC approach can be used for warehouse management. However, there is indication of prior work in applying ABC within a warehouse environment. In one of the literatures, the author studies how activity based costing principles could be applied to distribution logistics. In this paper, the authors designed an activity based costing system for the warehouse, by analyzing the resources cost (resources: space, equipment and personnel) and allocating costs to each warehouse activity such as receiving, high handling and low picking (Pirttilä and Hautaniemi, 1995). Varila et al examines the applicability of different drivers for assigning costs to activities in warehouse logistics environment (Varila et al., 2007).

2.3. Demand Forecast and Warehouse Operational Inefficiency

A few factors, including demand forecasts and push-pull boundary, drive inefficiencies in warehouse activities. The ability to accurately translate demand forecasts from a warehouses perspective plays a critical role in warehouse management planning. Warehouse activity forecast errors (or lack thereof) can have significant impacts on overall organizational cost and performance - 3.5% to 7.5 when out of line with demand forecasts (Sanders and Ritzman, 2004a).

According to a survey conducted by Dr. Raj Veeramani et al at the University of Wisconsin Madison, the most significant warehouse workforce planning challenge was no visibility to future demand with 22% prevalence (Veeramani, Paulson, & Peschel, 2007). Inaccurate demand forecast information was also a significant challenge with 14% of the responses.

Strategies to mitigate forecast errors include improving the forecast (the purpose of this literature review), keeping safety stock, and workforce flexibility as studied by Sanders and Ritzman.

2.4. Distribution Center Planning Methods - Warehouse Management Systems

Warehouse management systems (WMS) aid in numerous warehousing activities that are critical to overall warehousing operations. Specific functionalities may include a combination of the following:

- Inbound processing – aids in the management of shipments inbound from suppliers, captures information as shipments are received, and directs the warehousing of products

- Outbound processing – aids in the planning and management of scheduling, picking, packing, and outbound shipping of products from the warehouse

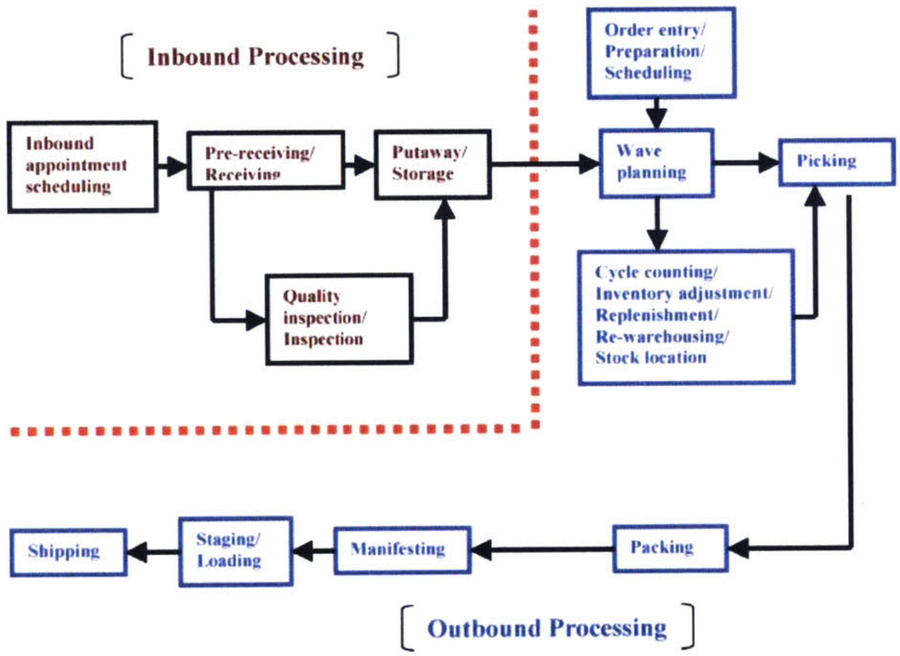

Figure 2-2 WMS Functionalities (Min, 2006)

According to Min (2006), there are numerous benefits to utilizing a WMS including increased processing accuracy, reduced labor costs, and more efficient workload management. Oftentimes, these systems will be integrated with the organizations enterprise resource planning systems (ERP) to automate the ordering process. Through these mechanisms, the demand forecast data can be directly integrated into the WMS in order to drive warehouse operations.

2.5. Distribution Center Operations

The main warehouse activities include receiving, transfer and put away, order picking/selection, accumulation/sortation, cross-docking, and shipping. As early as the 1970's, Lynagh presented a maximum performance model for evaluating the physical distribution activities of a warehouse. He identified total 37 factors to measure the effectiveness of warehouse performance. The 37 factors fall into the following three areas: order processing (10 factors), warehouse handling (10 factors), and transportation (17 factors) (Lynagh, 1971). Although these factors are for warehouse operation and out-of-date to some extent, they still shed light on the direction on which we should focus in the study.

We review multiple drivers of improvements in DCs necessitating needs for advanced DC operations planning. These key drivers, including order picking, cube size, and degree of workforce flexibility, will be the top targets to translate into DC operational metrics, because the other factors can't be quantified and therefore can't be included in a mathematical model.

2.5.1. Order Picking

Order picking is one of the largest labor components in a distribution center. The majority of warehouses and DCs employ humans for order picking. Recently, more DCs employ integrated systems, which reduces human travel and the number of touches associated with picking. For instance, pick system automation solutions have been deployed effectively in Europe for decades (Shemesh, 2010). However, order picking automated systems are very capital-intensive, even though they cut down human labor ("Book," 2002). Koster et al conduct an extensive review of order picking and their work suggests that nowadays order picking is still the highest priority area for productivity improvement in a distribution center (de Koster et al., 2007). Therefore, order picking is one of the top drivers to translate in the present thesis project.

2.5.2. Workforce Flexibility Degree

Workforce flexibility is the use of workforce arrangements that help the organization respond to change. Studies find that workforce flexibility is a highly effective tool in addressing variation in operations planning and reducing scheduling costs (Kum-Khiong Yang et al., 2002). Workforce flexibility can be achieved by cross-training, hiring part-time workers, using "floaters" (multi-skilled workers who are directed to points of greatest need), and adjusting length of workdays (Sanders and Ritzman, 2004b). The degree of workforce flexibility is an important indicator of DC performance, and thus deserves a close-up examination.

2.6. Literature Review Conclusion

The research clearly indicated the relationship between demand and warehouse operations. Lapide states that demand forecasts can be translated into supply-side views, including the warehouse, although without any specifics as to how (Lapide, 2006). Even more so, with the complexity and costs involved in warehouse operations, accurate demand forecasts can translate to enormous efficiencies in the warehouse. Despite research and technological advancements bridging the gap between the demand forecast and warehouse operations, there are some limitations found in the research. First, is that the simulated model by Sanders and Ritzman can only suggest potential efficiency gains with workforce flexibility. More extensive research in comparing with actual data in real warehouse environments could help in the validation and usability of such a model. In the case of the more automated WMS solutions, these provide a wealth of functionality in terms of inventory management and process automations. As effective as they are, there still exists potential gaps in the organizations effective use of such tools, and the large amount of effort necessary to integrate with external systems to fully enable much of the functionality. Consequently, WMS is wholly dependent on the quality of data fed from other systems (Min, 2006) and the overall total cost of ownership of implemented such complex systems is not well understood.

In summary, there has been significant progress in both the research and practical application of methods for translating demand forecast data into actionable warehouse operation plans. However, there still remains areas of opportunity in fully understanding the full impacts of these available methodologies that merit further research.

3. Methodology

The following section outlines the methodology utilized. First, we specify the scope of the modelling effort and the logic behind the specifications. Next, we introduce the different data sets we gathered for analysis. Then, we explain the forecast factors that were utilized to translate the demand forecast into forecasted picking hours. We continue by explaining how the actual picking hours were estimated due to the lack of actual detailed picking hours data available. We explain the different aggregation techniques utilized in preparation for comparison with the actuals. Finally, we explain the specific tools and analysis techniques involved in determining the performance of the results.

3.1. Scoping Demand Data

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 (one of major CVS DCs)
- 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 it is a DC due to 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.

Finally, we compare the forecast to actual. We focus on comparing the rolling horizon based forecast to the actual for a single week within the fiscal year, in order to access the time-changing performance of subsequent forecasts as the actual dates approach.

3.2. Data Collection

In order to access the performance of the demand forecast, we needed to collect multiple weeks of both demand forecasts as well as actual target weeks. Unfortunately, due to accessibility constraints, we ultimately collected the system-generated shipment forecasts for only two different weeks, as well as actual picking data from the warehouse management system for several weeks.

It is important to note that the first week forecast data in each of the forecast files may not include the full forecast data for that week. For example, the Week 11 forecast file may exclude forecast data from the first one or two days. This is because the forecast file was generated during the actual week. However, we were unable to confirm whether or not this was indeed the situation, and conducted our final analysis with all the available data. All data sets follow CVS' fiscal calendar convention of using week numbers. **Table 3-1** shows the data sets that were ultimately collected. **Table 3-2** shows the data sets that were analyzed.

Table 3-1 Data Sets Collected

Data Set	Description	Source
Week 11 Forecast	The demand forecast that begins forecasting from fiscal week 11, ends with week 20 (may be missing some week 11 demand)	Demand forecasting system
Week 15 Forecast	The demand forecast that begins forecasting from fiscal week 15, ends with week 24 (may be missing some week 15 demand)	Demand forecasting system
Week 17 Forecast	The demand forecast that begins forecasting from fiscal week 17, ends with week 26 (may be missing some week 17 demand)	Demand forecasting system
Actual Picking Activity Data	The actual picking activity data for weeks 11 through 17	Warehouse Management System

Table 3-2 List of Forecast Weeks and Target Weeks

Forecast Week	Target Week	Weeks In Advance
Week 11	Week 11	0*
Week 11	Week 12	1
Week 11	Week 13	2
Week 11	Week 14	3
Week 11	Week 15	4
Week 11	Week 16	5
Week 11	Week 17	6
Week 15	Week 15	0*
Week 15	Week 16	1
Week 15	Week 17	2
Week 17	Week 17	0*

*Note – the forecast data from the same week may be missing demand some data

3.2.1. Forecast File Data

The forecast files consisted of the following fields:

Table 3-3 Forecast Data Fields

Field	Description	Example
SKU	Stock keeping unit number	477069
CAT_NBR	The major product category number	87
SKU Description	A description of the SKU	Allergen relief tablets 50mg
Week Number	The week number during which the demand is forecasted to be experienced in YYYYWW format	201511
Units	The number of units forecasted	500
ID	A column that identifies the item type (we are only interested in 'regular' items)	'regular'

3.2.2. WMS Actuals Data

The actual picking data from the WMS system consisted of the following fields:

Table 3-4 WMS Actuals Data Fields

Field	Description	Example
Item Number	The stock keeping unit number	477069
Major/Minor	The major category number and minor category number combined into a four-digit number	8714
Description	A description of the SKU	Allergen relief tablets 50mg
Work Date	The date when the picking occurred	2015-05-11
Order Type	A column that identifies the item type (we are only interested in regular items indicated by STD and PSE)	STD
Units to Pick	The number of units picked	32
Level Seconds	The expected number of seconds to pick all units of the SKU (based on an engineered standard).	15
Assg Level Seconds	The expected number of seconds to fully pick all the pick orders containing the SKU	400
Assg Actual Seconds	The actual number of seconds elapsed to fully pick all the pick orders containing the SKU	395

3.3. Forecasted Pick Rate and Forecasted Picking Hours

CVS currently maintains a cost model for seven DC activities. A cost-per-piece (CPP) value is calculated for each activity (SC and FC picking) by the following formula:

$$CPP = APR \div PR$$

Where:

- CPP = Cost per piece-pick in $/piece-pick
- APR = Average Labor Pay Rate in $/hour
- PR = Pick rate in piece/hour

The APR is the average labor pay rate of all the dollars paid for the activity (SC or FC picking) in the selected CVS DC divided by the number of total hours required for the activity in all of 2014.

The end result were two CPP rates (one for SC and one for FC) that will be used to transform the demand forecast data for the SKUs in the Allergen Relievers and Laundry categories into costs.

As part of this cost model, a forecasted pick rate is calculated for each activity/product category (SC and FC picking).

For SC, the pick rate is a calculation based on the aggregate average pick rate for the DC across all categories and adjusting it by a category factor. The category factor is based on data from the labor management system in the Fredericksburg DC which more accurately tracks the pick rates of individual categories.

For FC, the pick rate is a calculation based on the total number of full cases picked in December 2014 across all product categories in the selected CVS DC divided by the total number of labor hours required to pick all the cases resulting in a pick rate in cases/hour. At this point, the pick rates for all full-case picks in the DC across all categories are equal, and an adjustment is necessary in order to more accurately model the pick-rates for different categories. Each product category had a weighted pieces/case metric provided by our partner. Multiplying the two numbers provided a pick rate (pieces/hour) for each product category.

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, category 87)

Full Case Pick Rate = 851 pieces/hour (applied to Landry category, category 94)

Figure 3-1 depicts the rick-rate analysis provided by CVS and the final pick-rate values utilized to transform the demand data into hours.

Category		ALLERGY REMEDIES (87)	LAUNDRY (94)		
SC Pick (SC, DPS)	**Metric**	100%	0%		
Pick Rate	*ps/hr*	920	633	*Vol*	229,919,552
Piece Adjustment		1	1	*Hrs*	285,877
Pieces per Hour	*pc/hr*	920	633	*TP*	804
FC (Belt & Pallet)	**Metric**	0%	100%		
Pick Rate	*ps/hr*	138	138	*Vol*	1,165,566
Piece Adjustment		51.06	6.15	*Hrs*	8,421
Pieces per Hour	*pc/hr*	7,067	851	*TP*	138

Figure 3-1 Partner-provided pick-rate inputs and analysis

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

3.4. Estimated Actual Picking 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)*
    - This is the total actual amount of time that it took to fulfill all the pick orders that contained the SKU. For example, 10 units of SKU A may have been split between two orders (each order also containing multiple other SKUs to pick). If completing each of the two pick orders took 30 minutes, then the total picking time would have been 60 minutes.
- *Estimated picking time for all the orders that contained the specific SKU (B)*
    - This is the total estimated amount of time that it *should* have taken to fulfill all the pick orders that contained the SKU. This is 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)*
    - This is the total estimated amount of time that it *should* have taken to pick only the specific SKU across all orders. This is 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:

$$[Actual\ picking\ time\ for\ all\ orders\ containing\ the\ specific\ SKU]$$
$$\times\ [Proportion\ of\ the\ estimated\ picking\ time\ attributable\ to\ the\ specific\ SKU]$$

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

$$[Estimated\ picking\ time\ to\ pick\ all\ the\ units\ of\ the\ specific\ SKU]$$
$$\div\ [Estimated\ picking\ time\ for\ all\ the\ orders\ that\ contained\ the\ specific\ SKU]$$

Making the final equation for **estimated actual picking time**:

$$[Actual\ picking\ time\ for\ all\ orders\ containing\ the\ specific\ SKU]$$

$$\times\ [Estimated\ picking\ time\ to\ pick\ all\ the\ units\ of\ the\ specific\ SKU]$$

$$\div\ [Estimated\ picking\ time\ for\ all\ the\ orders\ that\ contained\ the\ specific\ SKU]$$

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

3.5. Comparing Forecasts and Actuals

Once the above data transformations were implemented, we were able to compile all the forecast data for each of the target weeks into a single view, allowing us to continue with our analysis.

Wk11		Wk11 Actual		Forecast from Wk11				Forecast of Wk11 at Wk11 (UNITS)					Forecast of Wk11 at Wk11 (HOURS)				
SKU	CAT	Units	Hours	Units	Hours	Variance from Actu	Variance from	et = At-Ft	et^2	Et/At	Absolute Et	ABS Et /At	et = At-Ft	et^2	Et/At	Absolute Et	ABS Et /At
210285	87	170	0.338	146	0.16	16%	113%	24	576	14%	24	0.14	0.18	0.03	53%	0.18	0.53
210287	87	264	0.264	253	0.28	4%	-4%	11	121	4%	11	0.04	-0.01	0.00	-4%	0.01	0.04

Figure 3-2 Example of Consolidated Data View (Week 11)

4. Data Analysis

The following sections describe the data analysis conducted. First, we describe the statistical methods utilized in our evaluations. Then, we compare the forecast to actual by comparing the rolling horizon based forecast to the actual a single week at a time within the fiscal year, in order to access the time-changing performance of subsequent forecasts as the actual dates approach. This was limited to actual weeks 16 and 17 because these are the actual weeks for which we have two confidently complete forecasts. We conclude the analysis by compiling all the forecast and actual comparisons to understand how a rolling time horizon affects the accuracy of forecasts at different time intervals in advance.

4.1. Parameters to evaluate the accuracy of demand forecast versus actual

The mean deviation (MD) measures model accuracy. The ideal value of MD is zero. If MD > 0, the model tends to under-forecast. If MD < 0, the model tends to over-forecast. The mean absolute deviation (MAD) is another common parameter for forecasting evaluation. While MD is a measure of forecast model bias, MAD indicates the absolute size of the errors. MAD is a linear score: it means that all the individual differences are weighted equally in the average.

The mean squared error (MSE) and the mean absolute percent error (MAPE) both measure model accuracy. MSE measures total model error and does not normalize the error terms for the size of the observed values. The MAPE term, on the other hand, gives the average absolute error as a proportion of the observed value. In other words, MAPE measures the size of the error in percentage terms. The mean percent error (MPE) measures the bias of the model. A MPE of zero means the model is perfectly unbiased. A non-zero MPE, however, means that the model has either positive or negative bias.

The RMSE is a quadratic scoring rule which measures the average magnitude of the error. Since the errors are squared before they are averaged, the RMSE gives a relatively high weight to large errors. This means the RMSE is most useful when large errors are particularly undesirable.

The MAD and the RMSE can be used together to diagnose the variation in the errors in a set of forecasts. The RMSE will always be larger or equal to the MAD; the greater difference between them, the greater the variance in the individual errors in the sample. If the RMSE=MAD, then all the errors are of the same magnitude. Both the MAE and RMSE can range from 0 to ∞. They are negatively-oriented scores: Lower values are better.

The following equations are for MD, MAD, MSE, MAPE, MPE and RMSE, respectively.

$$\textit{Mean Deviation (MD)} = \frac{\sum_{t=1}^{n} e_t}{n}$$

$$\textit{Mean Absolute Deviation (MAD)} = \frac{\sum_{t=1}^{n} |e_t|}{n}$$

$$\textit{Mean squared error (MSE)} = \frac{\sum_{t=1}^{n} e_t}{n}$$

$$\textit{Mean absolute percent error (MAPE)} = \frac{\sum_{t=1}^{n} \frac{|e_t|}{a_t}}{n}$$

$$\textit{Mean percent error (MPE)} = \frac{\sum_{t=1}^{n} \frac{e_t}{a_t}}{n}$$

$$\text{Root Mean Squared Error (RMSE)} = \sqrt{\frac{\sum_{t=1}^{n} e_t^2}{n}}$$

Here, a_t represents an actual value for observation t from the dataset, e_t represents the error for observation t, or the difference between the actual value and the forecast value F_t ($e_t = a_t - F_t$), and n represents the number of observations in the dataset.

4.2. Accuracy Analysis of Subsequent Rolling Horizon Based Forecasts

In this section, we focused on comparing the rolling horizon based forecast to the actual for single weeks within the fiscal year, in order to access the time-changing performance of subsequent forecasts as the actual dates approach. We limited these comparisons to actual weeks 16 and 17 because these were the weeks for which we were completely confident that we had two full weeks of forecast data.

4.2.1. Actual Week 16 Forecast Analysis

Two forecasts were made at week 11 and week 15, respectively, for the actual week 16 (discussed in this section) and actual week 17 (discussed in Section 4.2.3).

To begin, we examined how accurate, and unbiased the SKU level unit demand forecast is. We checked the mean absolute percent error (MAPE) and the mean percent error (MPE). MAPE refers to the accuracy of the forecast, and MPE refers to the bias in the forecast. The Unit forecast at week 11 for week 16 had a MAPE of 668 percent and a MPE of negative 635 percent. These values mean that this forecast had negative 568 percent accuracy with a bias of as high as 635 percent. Interestingly, MAPE and MPE values both decreased significantly at week 15 for week 16 forecast. As we can see from **Table 4-1**, MAPE decreased from 668 percent at week 11 forecast to 93 percent at week 15 forecast. This result suggests that the accuracy of Unit forecast improved from week 11 to week 15 forecasts, as the rolling base forecast approaches to the actual week 16.

Meanwhile, MPE of Unit forecast increased from negative 635% (week 11 forecast) to negative 58% (week 15 forecast). This result suggests that: 1) the bias of the Unit forecast improved, 2) the forecast was too high at week 11 (actual less than forecast) and later became closer to the actual even though the actual was still less than forecast.

RMSE's for the forecasts of Unit and Hours at week 11 and at week 15 did not change much, although RMSE at week 15 was slightly higher than RMSE at week 11. This result suggests that the errors remained at the similar scale.

Table 4-1 Forecast Accuracy Parameters for the Actual Week 16

Parameters	Unit Variance		Hours Variance	
	Forecast of Wk 16 at Wk 11	Forecast of Wk 16 at Wk 15	Forecast of Wk 16 at Wk 11	Forecast of Wk 16 at Wk 15
MD	-73.0	-91.4	-0.06	-0.08
MAD	146.1	147.5	0.20	0.21
RMSE	250.6	279.7	0.35	0.39
MSE	62804	78242	0.12	0.15
MAPE	668%	93%	765%	114%
MPE	-635%	-58%	-722%	-65%

4.2.2. Actual Week 16 Forecast Bias Distribution

To examine how the forecast is distributed, we divided the MPE range into different buckets and counted how many SKUs lie in each range bucket. **Figure 4-1** and **Figure 4-4** shows that the majority SKUs had Unit MPE and Picking Hours MPE around negative 100% for both forecast at week 11 and week 15.

Another interesting result is that the Unit MPE pattern was not highly correlated with the Labor hours MPE pattern, even though the trend was similar between the two (**Figure 4-1** and **Figure 4-4**). This result indicated that the Unit/Labor conversation factor determined by the cost model is not fully representative the engineer factor provide by the sponsor company.

When looking at the two product categories 87 and 94, we observed that their distribution patterns of Unit MPE were not the same (**Figure 4-2** and **Figure 4-3**), nor were the distribution patterns of Picking Hours MPE (**Figure 4-5** and **Figure 4-6**).

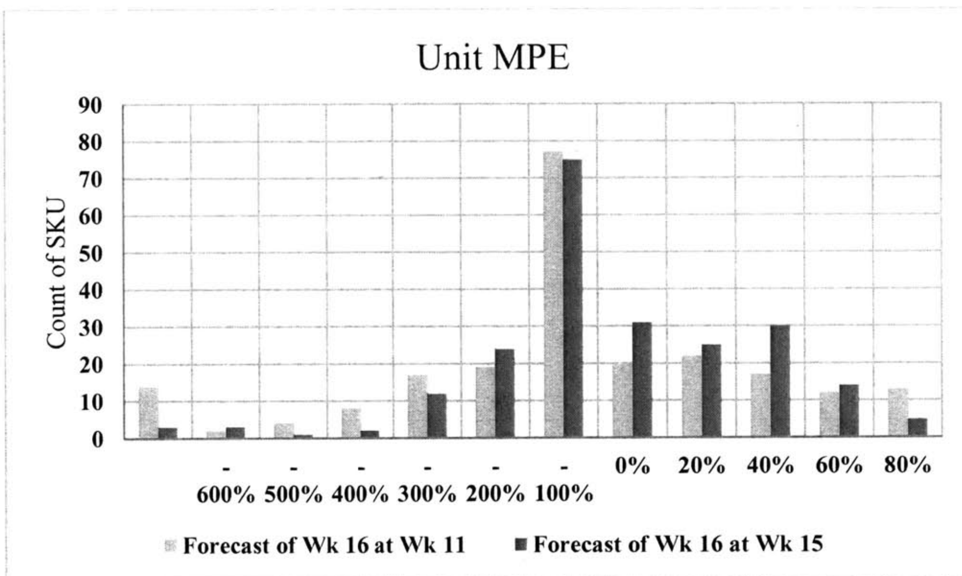

Figure 4-1 Actual Week 16 Distribution of Unit MPE Variance from Both Category 87 and 94

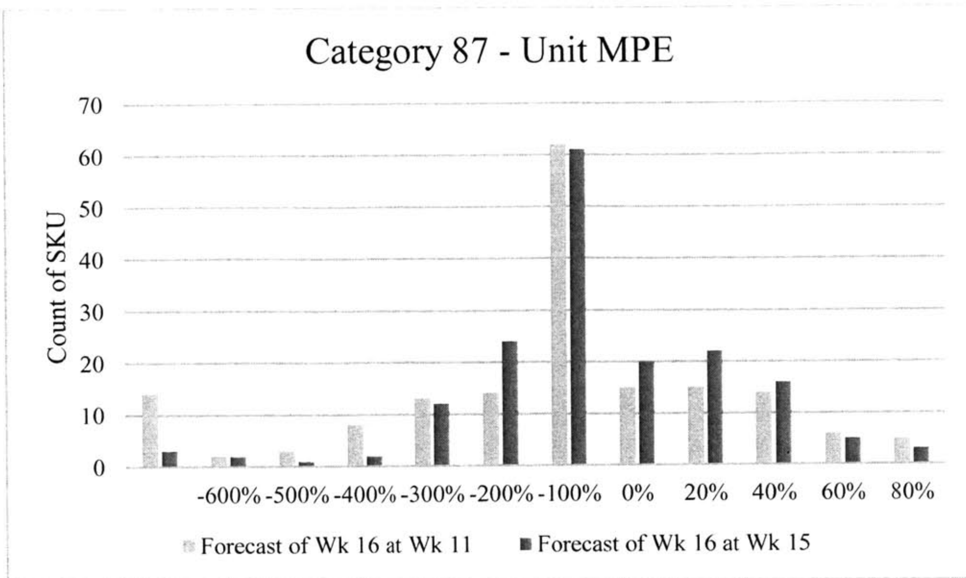

Figure 4-2 Actual Week 16 Distribution of Unit MPE from Category 87

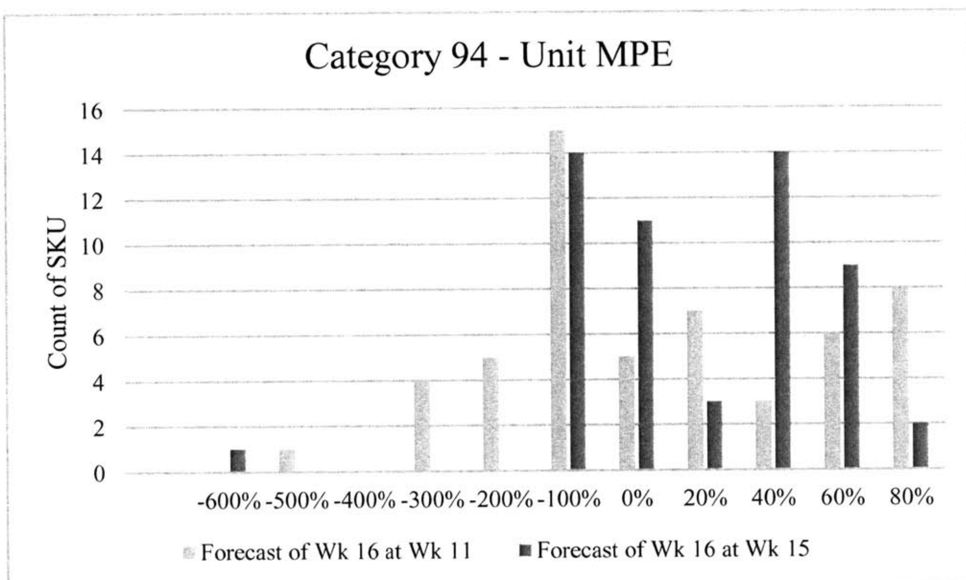

Figure 4-3 Actual Week 16 Distribution of Unit MPE from Category 94

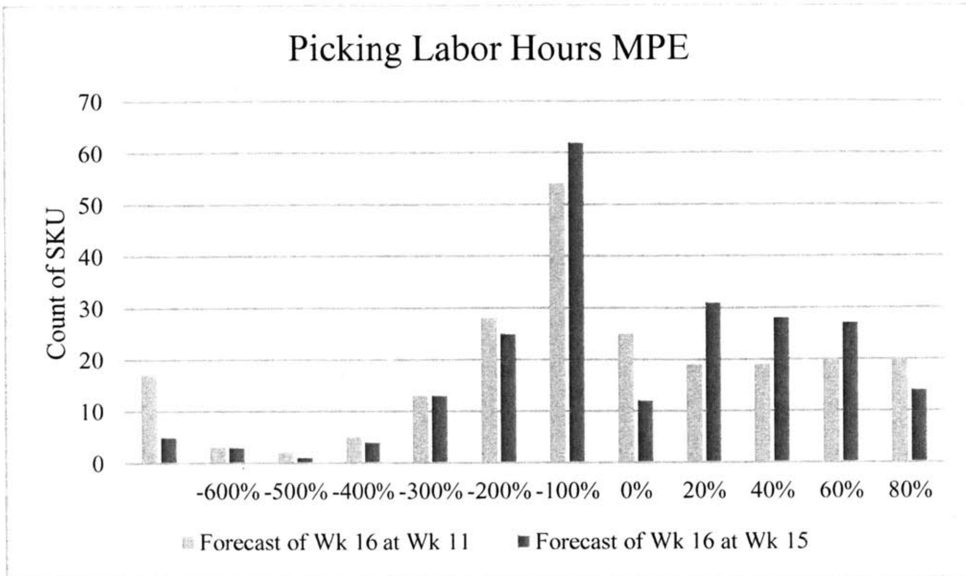

Figure 4-4 Actual Week 16 Distribution of Picking Hours MPE from Both Category 87 and 94

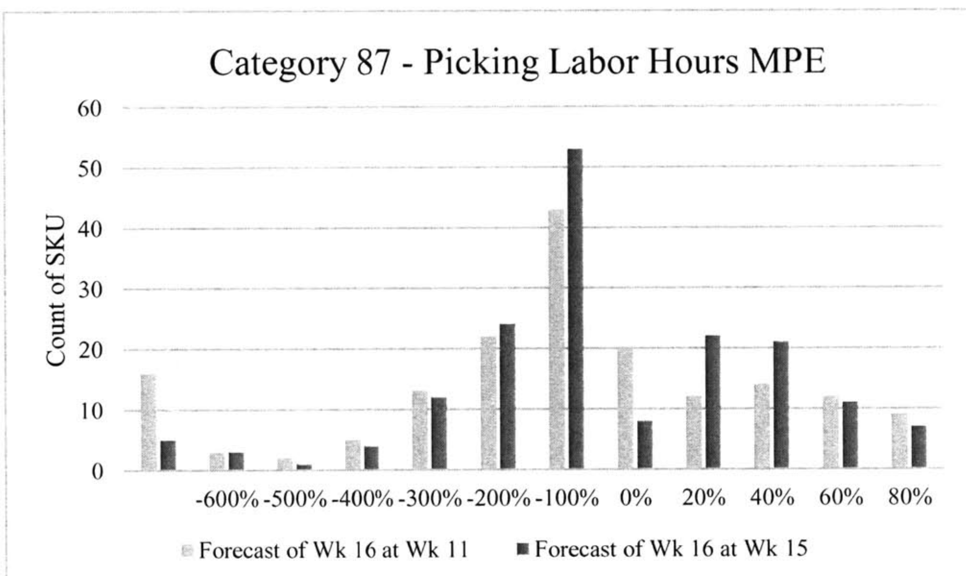

Figure 4-5 Actual Week 16 Distribution of Picking Hours MPE from Category 87

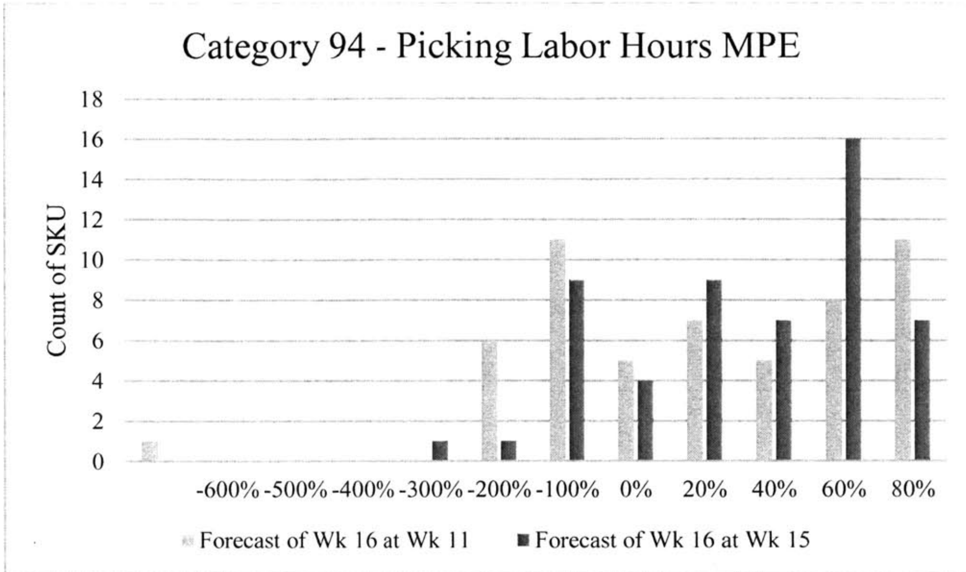

Figure 4-6 Actual Week 16 Distribution of Picking Hours Variance from Category 94

4.2.3. Actual Week 17 Forecast Analysis

The Unit forecast at week 11 for week 17 had a MAPE of 312 percent and a MPE of negative 276 percent. These values mean that this forecast had negative 212 percent accuracy with a bias of as high as 276 percent. Interestingly, MAPE and MPE values both decreased significantly at week 15 for week 17 forecast. As we can see from

Table 4-2, MAPE decreased from 321 percent at week 11 forecast to 103 percent at week 15 forecast. Meanwhile, Unit MPE increased from -276% (week 11 forecast) to -23% (week 15 forecast). These results are in line with the trend we observed for the actual week 16.

Table 4-2 Forecast Accuracy Parameters for the Actual Week 17

Parameters	Unit Variance			Hours Variance		
	Forecast of Wk17 at Wk11	Forecast of Wk17 at Wk15	Forecast of Wk17 at Wk17	Forecast of Wk17 at Wk11	Forecast of Wk17 at Wk15	Forecast of Wk17 at Wk17
MD	12.6	113.9	48.6	-0.01	0.10	0.0
MAD	188.7	156.9	114.6	0.24	0.19	0.2
RMSE	552.7	482.0	372.0	0.46	0.39	0.4
MSE	305487	232342	138370	0.21	0.15	0
MAPE	312%	103%	66%	357%	129%	88%
MPE	-276%	-23%	-31%	-310%	-44%	-37%

4.2.4. Actual Week 17 Forecast Bias Distribution

We also did the same MPE distribution analysis to evaluate how many SKU lie in each range bucket. **Figure 4-7** shows that just like the forecasts for actual week 16, the majority SKUs in the week 11 and week 15 forecasts for actual week 17 had a MPE around negative 100% for both Unit and Picking Hours. However, for the forecast made at week 17, the majority SKUs had MPE close to 0%, suggesting that this forecast was unbiased. This result was expected because the forecast was made in the same week as the actual.

The Unit MPE pattern was again not highly correlated with the Labor hours MPE pattern, even though the trend was similar between the two (**Figure 4-7** and **Figure 4-10**). When looking at the two product categories 87 and 94, we observed that their distribution patterns of Unit MPE were not the same (**Figure 4-8** and **Figure 4-9**), nor were the distribution patterns of Picking Hours MPE (**Figure 4-11** and **Figure 4-12**). Overall, we observed similar results as we saw in Section 4.2.2.

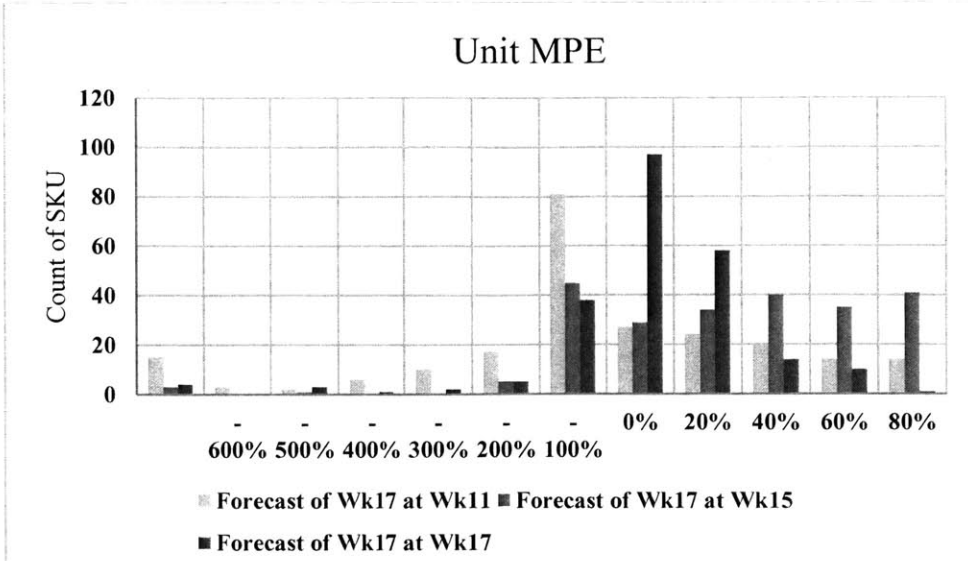

Figure 4-7 Actual Week 17 Distribution of Unit MPE Variance from Both Category 87 and 94

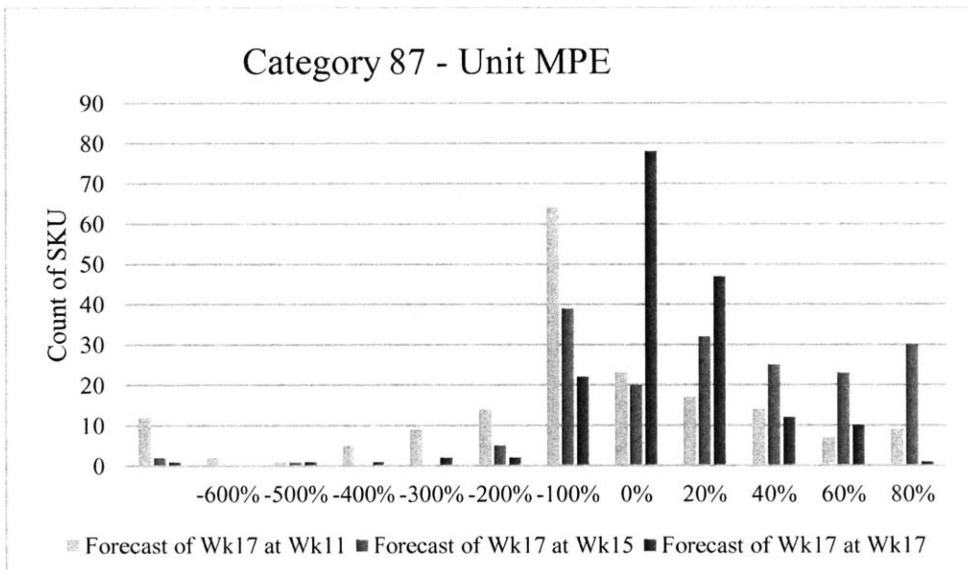

Figure 4-8 Actual Week 17 Distribution of Unit MPE from Category 87

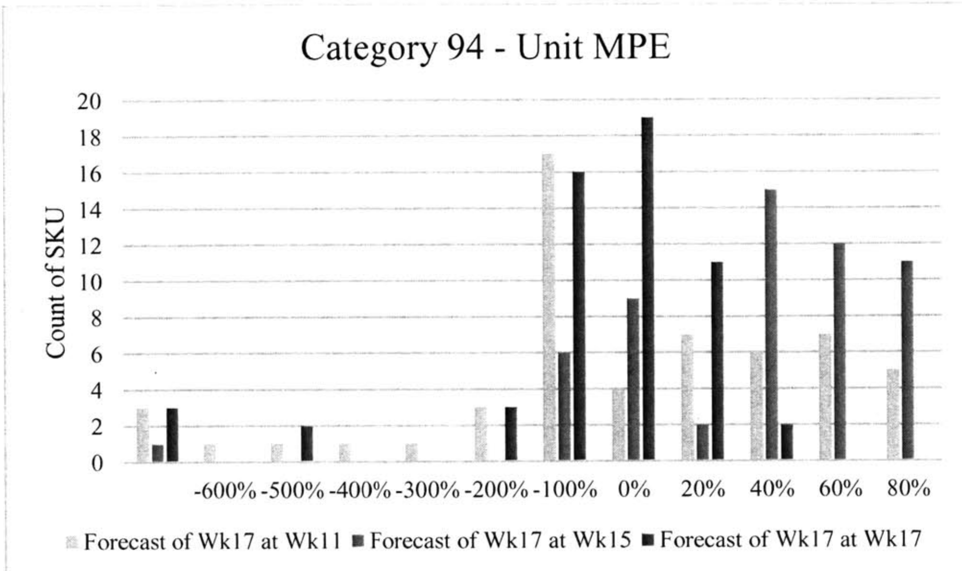

Figure 4-9 Actual Week 17 Distribution of Unit MPE from Category 94

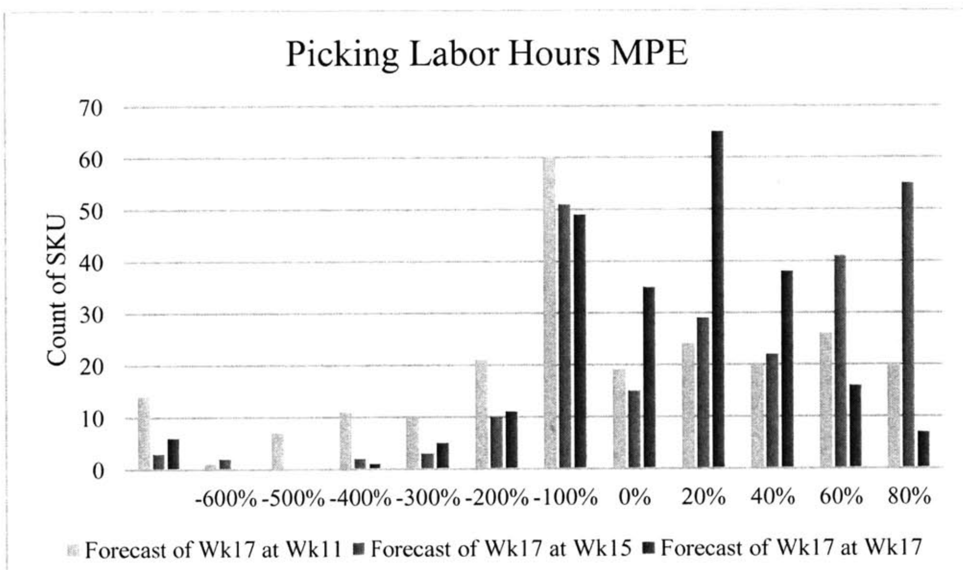

Figure 4-10 Actual Week 17 Distribution of Picking Hours MPE from Both Category 87 and 94

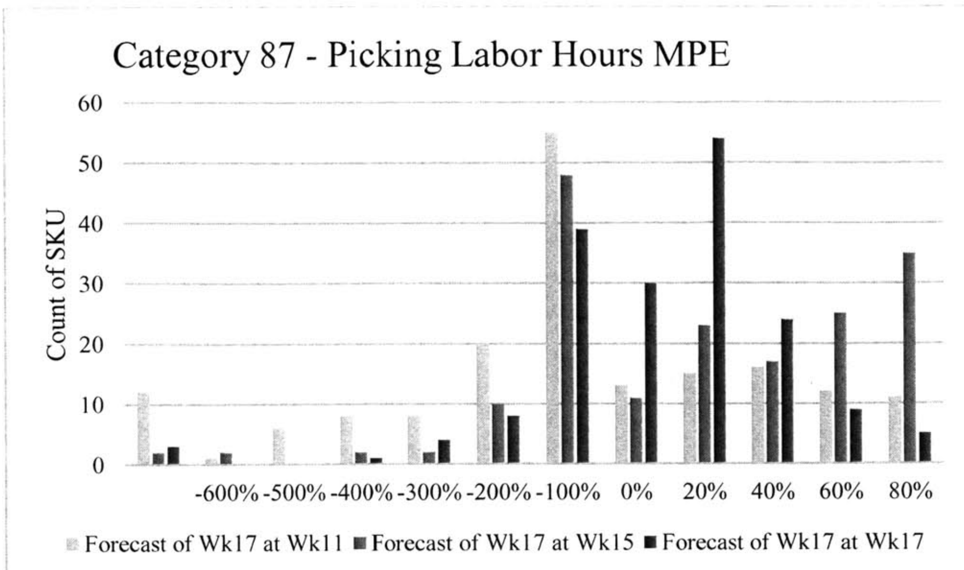

Figure 4-11 Actual Week 17 Distribution of Picking Hours MPE from Category 87

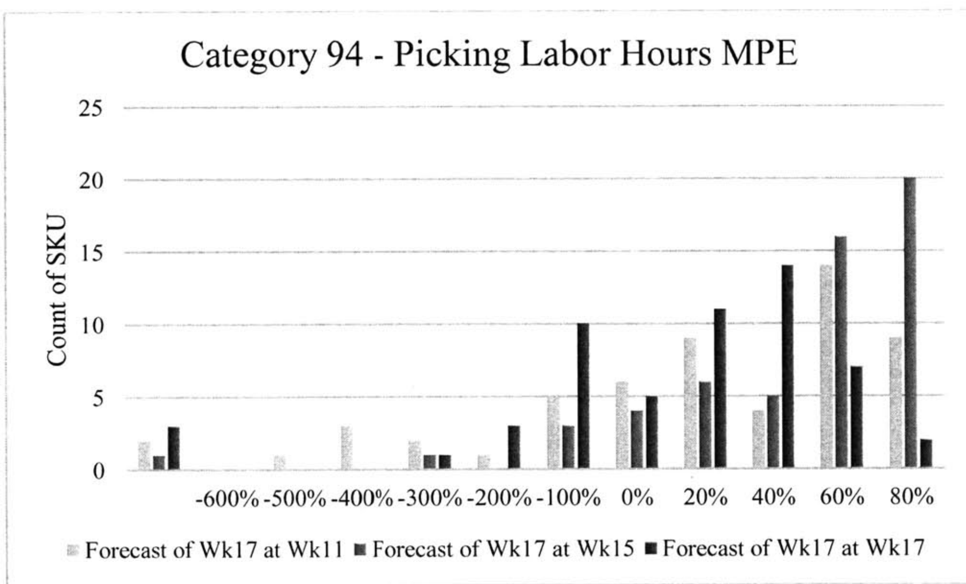

Figure 4-12 Actual Week 17 Distribution of Picking Hours Variance from Category 94

4.3. Overall Combined Comparison of Forecasts and Actuals

Next, we examined the overall change of all the parameters. **Table 4-3** lists all the calculated parameters for Unit Forecasting. **Table 4-4** lists all the calculated parameters for Picking Hours labor forecasting. To visualize the change of forecast parameters, we plotted three parameters against the number of week between the forecast time point and actual time point (**Figure 4-13** to **Figure 4-18**). RMSE and MAPE both tended to get smaller as the number of week in advance decreased from 6 week in advance to 0 week in advance. MPE became less negative as the number of week in advance decreased. This result suggested the forecast accuracy and bias gets better as we approach closer to the actual observed week.

Table 4-3 Forecast Accuracy Parameters for All Units

Units								
Actual Week	Forecast Week	Weeks in Advance	MD	MAD	RMSE	MSE	MAPE	MPE
11	11	0	-1.92	37.74	76.97	5924.98	59%	-31%
15	15	0	132.63	135.20	248.14	61571.37	54%	52%
17	17	0	48.63	114.62	371.98	138370.39	66%	-31%
12	11	1	70.97	100.81	174.93	30599.39	84%	11%
16	15	1	-91.45	147.52	279.72	78242.05	93%	-58%
13	11	2	50.74	113.69	270.00	72901.76	220%	-141%
17	15	2	113.95	156.92	482.02	232342.04	103%	-23%
14	11	3	13.96	126.00	236.55	55955.00	160%	-101%
15	11	4	8.12	128.71	264.22	69812.42	150%	-106%
16	11	5	-73.03	146.13	250.61	62804.19	668%	-635%
17	11	6	12.60	188.65	552.71	305486.92	312%	-276%

Table 4-4 Forecast Accuracy Parameters for All Picking Hours

Hours								
Actual Week	Forecast Week	Weeks in Advance	MD	MAD	RMSE	MSE	MAPE	MPE
11	11	0	0.02	0.10	0.21	0.05	231%	-184%
15	15	0	0.16	0.17	0.38	0.15	55%	51%
17	17	0	0.03	0.17	0.38	0.14	88%	-37%
12	11	1	0.11	0.15	0.38	0.15	99%	1%
16	15	1	-0.08	0.21	0.39	0.15	114%	-65%
13	11	2	0.07	0.16	0.40	0.16	603%	-523%
17	15	2	0.10	0.19	0.39	0.15	129%	-44%
14	11	3	0.02	0.16	0.29	0.08	182%	-116%
15	11	4	0.02	0.20	0.42	0.18	172%	-121%
16	11	5	-0.06	0.20	0.35	0.12	765%	-722%
17	11	6	-0.01	0.24	0.46	0.21	357%	-310%

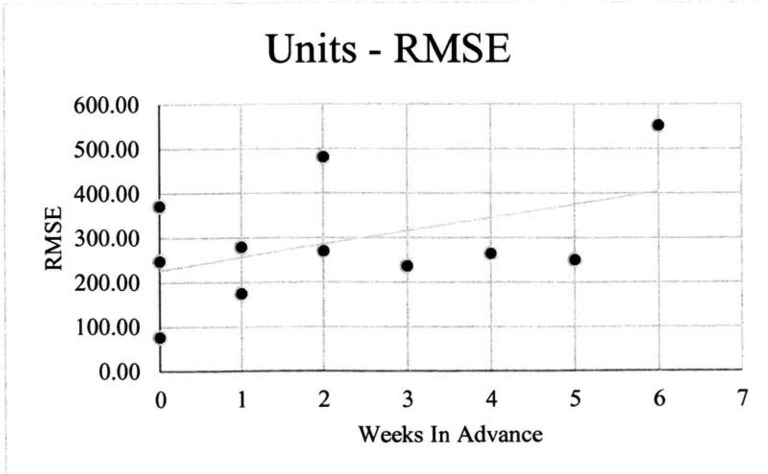

Figure 4-13 Unit RMSE of All Units

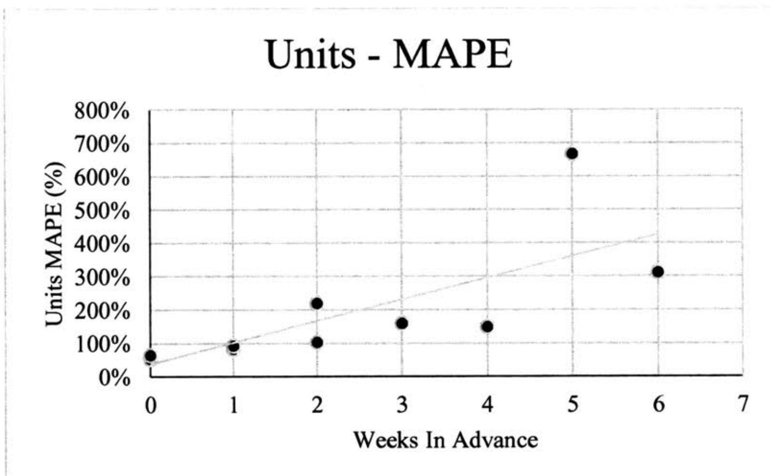

Figure 4-14 Unit MAPE of All Units

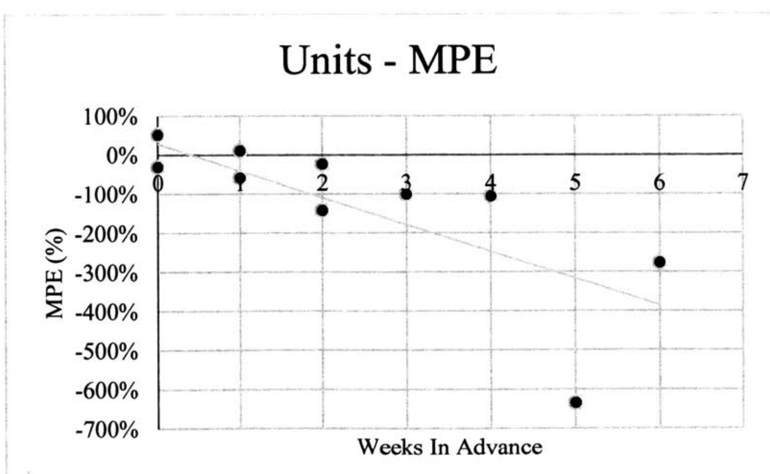

Figure 4-15 Unit MPE of All Units

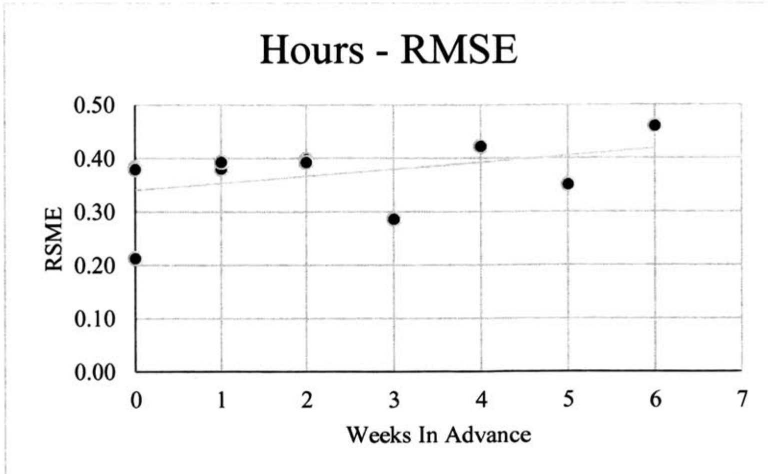

Figure 4-16 Hours RMSE of All Picking Hours

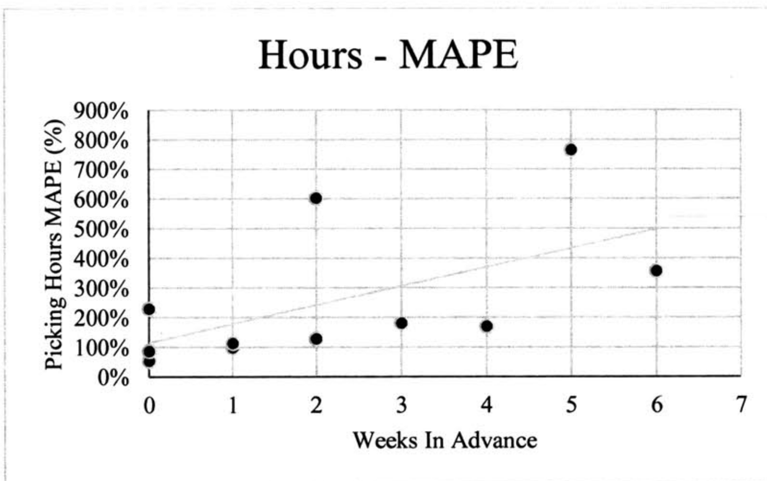

Figure 4-17 Hours MAPE of All Picking Hours

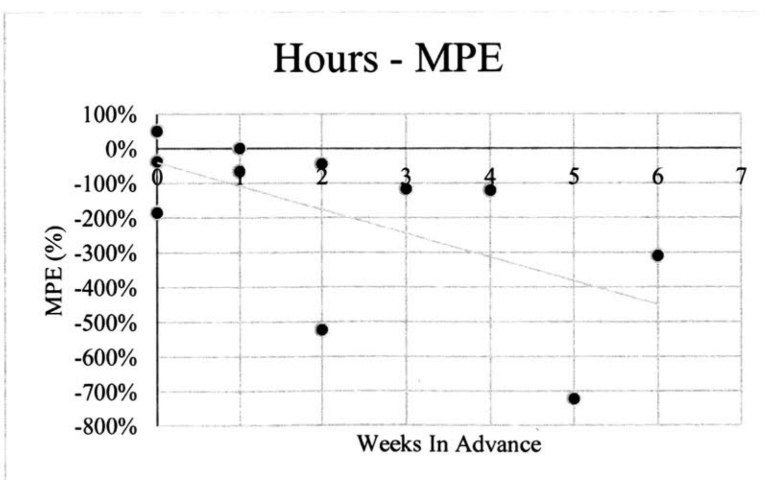

Figure 4-18 Hours MPE of All Picking Hours

5. Discussion

Our performance analysis of the demand forecasts in terms of units and picking hours suffers from multiple limitations. Each of these are briefly introduced below.

5.1. Forecast and Actual Data Availability

Our analysis limited to two demand forecast points and one actual point for each SKU. With more data points in both the forecasts and the actuals, we could implement a more comprehensive analysis with improved resolution spanning multiple weeks and multiple forecasts per week.

5.2. Forecasted Pick Rates in the Cost Model

The forecasted pick rates were based on overall average pick rates within the DC with category factors. This method resulted in pick rates per category. Since each SKU within a category likely has significantly varying real-world pick rates due to differences in size, weight, and location within the warehouse, improving the methodology for forecasted pick rates that addresses the differences between individual SKUs within a category may improve results. One potential avenue would be to utilize the same engineered level seconds that are currently used to estimate the actual picking hours. In that case, the process by which the engineered level seconds are produced will require continual monitoring and updating to that they represent changes in real-world activities.

5.3. Estimated Actual Picking Hours

The estimated actual picking hours were based on engineered level seconds provided by CVS. We are aware that there has been an ongoing project to capture picking time data at the individual SKU level within one of CVS' other DCs. Utilizing real-time, actual picking data to calculate the level seconds at the detailed SKU level may better represent reality. However, utilizing this data from a different DC may not be accurate for our selected DC due to innate differences including but not limited to warehouse size, warehouse layout, item mix, and number of stores served.

5.4. Process-Time Balancing

Our analysis aggregated data on a weekly level and attributed a data point to a fiscal week if the timestamp fell within the fiscal week. However, there are natural delays in the system between picking and shipments (e.g. an item picked on Saturday may not ship until the following week). Because we compared actual picking data with forecasted demand shipment data, we may find that applying a time-based offset to our week-based aggregation may improve results.

6. Conclusions

The purpose of this thesis was to explore the feasibility of translating SKU-level demand forecast into distribution center operations metrics to facilitate work planning. Our research suggests that the appropriateness of utilizing demand forecasts to predict warehouse picking generally improves as we approached to the actual week for both units and hours. However, there are still significant variations in the results that have no clear explanation. In particular, the results for five weeks in advance showed significantly worse results than the other data points. Unfortunately, with such a small sample set, it is impossible to determine if this is an outlier, or a particular bias that actually exists across the population. In a similar vein, there is also insufficient data to determine if a specific number of weeks in advance is more accurate than others in forecasting warehouse picking activity. Overall, it is the opinions of the researchers that forecasts could be utilized to forecast warehouse activities under additional conditions that address the stated challenges. These include:

- the collection and analysis of a much larger sample of forecasts and actuals, preferably over the span of a calendar year;
- the capture of accurate labor hours data for specific warehouse activities, and
- the utilization of a single standard – in the case of our thesis, utilizing either the pick rates from the cost model or the engineered level seconds standards. Because we consider the number units to be the primary driver of warehouse activities, it is our opinion that effectively using different pick rates for the forecasts and actuals is unmerited.

Additionally, with a year's worth or more of data, it is the opinions of the researchers that the comparative performance of the forecasts in terms of weeks in advance can be conclusively determined should the above challenges be addressed. Overall, our research warrants a more thorough analysis of additional data and warehouse activity metrics in order to identify more specific insights as to the appropriateness of utilizing demand forecasts to predict warehouse picking and other activities.

References

Askarany, Davood, Yazdifar, Hassan, 2007. Why ABC is not widely implemented? *Int. J. Bus. Res.* 7, 93–98.

Baykasoğlu, A., Kaplanoğlu, V., 2008. Application of activity-based costing to a land transportation company: A case study. Int. J. Prod. Econ. 116, 308 – 324. doi:http://dx.doi.org/10.1016/j.ijpe.2008.08.049

Book: Facilities planning, 3rd ed.. James A. Tompkins, John A. White, Yavuz A. Bozer, and J.M.A. Tanchoco, authors © 2003 John Wiley & Sons, hardcover, 744 pp., $117.90 ISBN: 0471413895, 2002. . J. Manuf. Syst. 21, 326. doi:10.1016/S0278-6125(02)80176-5

Çakanyıldırım, Metin. "Forecasting." University of Texas at Dallas. Web. 1 May. 2015. Retrieved from http://www.utdallas.edu/~metin/Ba3352/Slides/forecast.ppt.

Crane, D.B., Crotty, J.R., 1967. A Two-Stage Forecasting Model: Exponential Smoothing and Multiple Regression. *Manag. Sci.* 13, B–501.

De Koster, R., Le-Duc, T., Roodbergen, K.J., 2007. Design and control of warehouse order picking: A literature review. *Eur. J. Oper. Res.* 182, 481–501. doi:10.1016/j.ejor.2006.07.009

Inman, R. Anthony, "Forecasting" *Reference for Business.* Web. 1 May. 2015

Kaplan, R.S., 1986. Accounting Lag: the obsolescence of cost accounting systems. *Calif. Manage. Rev.* 28, 174–199.

Kum-Khiong Yang, Webster, S., Ruben, R. a., 2002. An evaluation of flexible workday policies in job shops. *Decis. Sci.* 33, 223–249.

Lapide, L., 2006. Top-down & Bottom-up Forecasting in S&op. *J. Bus. Forecast.* 25, 14–16.

Liebeskind, A., 2013. It's all about the CUBE. *Ind. Eng. IE* 45, 36–41.

Lynagh, P.M., 1971. Measuring Distribution Center Effectiveness. Transp. J. Am. Soc. Transp. Logist. Inc 11, 21–33.

Min, H. (2006). The applications of warehouse management systems: an exploratory study. International Journal of Logistics Research and Applications, 9(2), 111–126.

Pirttilä, T., Hautaniemi, P., 1995. Activity-based costing and distribution logistics management. *Int. J. Prod. Econ.* 41, 327 – 333. doi:http://dx.doi.org/10.1016/0925-5273(94)00085-9

Schoenmeyr, T., & Graves, S. C. (2009). Strategic Safety Stocks in Supply Chains with Evolving Forecasts. *Manufacturing & Service Operations Management*, *11*(4), 657–673. http://doi.org/10.1287/msom.1080.0245

Sanders, N.R., Ritzman, L.P., 2004. Using Warehouse Workforce Flexibility to Offset Forecast Errors. *J. Bus. Logist.* 25, 251–269.

Satoglu, S.I., Durmusoglu, M.B., Dogan, I., 2006. Evaluation of the conversion from central storage to decentralized storages in cellular manufacturing environments using activity-based costing. *Int. J. Prod. Econ.* 103, 616 – 632. doi:http://dx.doi.org/10.1016/j.ijpe.2005.12.003

Shemesh, L.D., 2010. Planning for picking? Pick your plan. *Mater. Handl. Manag.* 65, 28–32.

Varila, M., Seppänen, M., Suomala, P., 2007. Detailed cost modelling: a case study in warehouse logistics. *Int. J. Phys. Distrib. Logist. Manag.* 37, 184. doi:10.1108/09600030710742416