



HAL
open science

Workforce Planning Models for Distribution Center Operations

Athul Gopala Krishna, Vittaldas V. Prabhu

► **To cite this version:**

Athul Gopala Krishna, Vittaldas V. Prabhu. Workforce Planning Models for Distribution Center Operations. IFIP International Conference on Advances in Production Management Systems (APMS), Sep 2016, Iguassu Falls, Brazil. pp.206-213, 10.1007/978-3-319-51133-7_25 . hal-01615728

HAL Id: hal-01615728

<https://inria.hal.science/hal-01615728v1>

Submitted on 12 Oct 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Workforce Planning Models for Distribution Center Operations

Athul Gopala Krishna and Vittaldas V. Prabhu

Pennsylvania State University, University Park, U.S.A.
aug276@psu.edu

Abstract. Customer order fulfillment at distribution centers (DC) is increasingly necessitated by innovative strategies to maximize operational performance that are primarily driven by cost and service level under supply chain variability. In order to better understand the trade-offs, in this paper, a generic computational model is developed to estimate forklift travel times for DCs with any arbitrary floor space and loading docks. In particular, travel times are modelled as random variables and the moments of the probability distribution of travel times are estimated and used as inputs to analytical queueing model and discrete event simulation model. Results show that the analytical and simulation models are within 3% under different demand scenarios. These models are used to determine the impact of work-force capacity on key performance measures such as Truck Processing Time (TPT) and Labor Hours Per Truck (LHPT). The workforce capacity for different demand scenarios is determined using three different approaches - Target Utilization Level, Square Root Staffing (SRS) rule (adapted from call center staffing) and Optimization. The result from these models indicate that adapting workforce capacity to match varying demand can reduce cost by 18% while maintaining de-sired service level.

Keywords: Distribution center · Workforce capacity · Simulation.

1 Introduction

Supply chain variability can be caused by product seasonality, batch production and transportation, product consolidation or value added processing. DCs buffer the material flow in supply chains to accommodate this variability. Incoming items brought to the DC are unloaded at the receiving docks (receiving) and put into storage (storing). Outgoing items are retrieved from storage (order picking), processed and shipped to customers through the shipping docks (shipping). Resources such as space, labor, and equipment are allocated to different DC functions following organizational policies to achieve desired operational performance in terms of capacity, throughput and service at minimum cost. DCs can adapt to varying demand by adjusting workforce capacity to meet desired service level [1,2].

Several authors had investigated the effects of warehouse design and control on operational performance and developed analytical and simulation models for

performance analysis. Pandit and Palekar investigated the effects of warehouse design on response time and suggested a method for optimal design based on response time [3]. Chew and Tang presented a travel time model that evaluates performance of an order picking system with consideration to order batching and storage allocation strategies [4]. Graves et al. evaluated warehouse performance for several sequencing and class based storage policies using continuous analytical models and discrete evaluation procedures, [5]. Bozer and White modelled the performance of an end-of-aisle order picking system by deriving analytical expressions and developed a design algorithm to determine the optimal configuration [6]. Koster proposed a modeling and approximate analysis method for a pick-to-belt order picking system based on Jackson network modeling and analysis [7]. Lee presented an analytical method for the stochastic analysis of a unit load AS/RS [8]. Hur et al. presented an analytical model for the stochastic analysis of a unit load AS/RS without assuming any specific distribution for the travel time of the S/R machine. [9]. The existing literature has predominantly focused on warehouse design & control using travel time models and performance analysis in terms of throughput, resource utilization and storage strategy for order picking systems. As labor cost can be a significant component of variable cost in DCs, it is evident that there is a need for research to explore the impact of workforce capacity on DC performance by considering cost and service level trade-off for varying demand scenarios. This work has been motivated by paucity of models and tools that can be used by DC operations managers for workforce planning in practice. Hence, a decision model that integrates critical operational performance measures to evaluate DC performance and workforce capacity policies under varying demand scenarios is developed.

2 Model and Assumptions

A large scale, non-automated, rectangular DC is analyzed. A rectangular shape is the optimal geometrical shape for storing rectangular units such as pallets [10]. The storage locations are characterized by single-deep racks and drive-in racks. The racks are arranged back-to-back, to form a block, parallel to the dock of the DC, such that space between blocks form aisles. The blocks are arranged in a rectangular grid to form a network of aisles through which material handling devices such as forklifts travel as shown in Figure 2 (a) [3]. A generic computational model with length (L), width (W), aisle-width (A) and number of docks (N) as parameters is developed to estimate forklift travel times in DCs. The model is capable to evaluate DCs with a size between 60,000 sq.ft (small scale) and 500,000 sq.ft. (large scale) and up to 45 docks. A probability distribution is fit to the forklift travel times and moments of the distribution is provided as input to the analytical queueing model of DC. The skewness profile of forklift travel time distribution against distance from center dock and L & W of the DC is investigated. It is observed that for a given W , skewness profile is represented by a series of peaks and valleys with decreasing value as L increases and for a given L , skewness profile remains nearly constant as W increases as shown in Figure

2 (b). The skewness for forklift travel time distributions fluctuate significantly from dock 1 (end dock) and becomes negligible towards dock 23 (center dock) and remains negligible from the center dock to dock 45 (end dock) as shown in Figure 2.

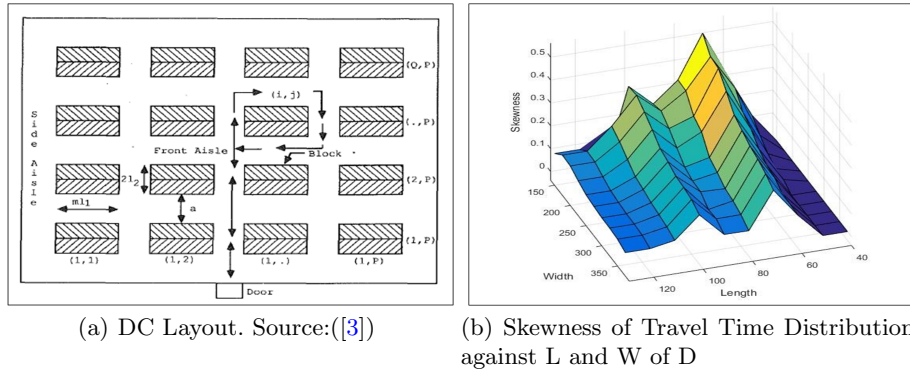


Fig. 1. DC Layout and Skewness of Travel Time Distribution against L and W of DC

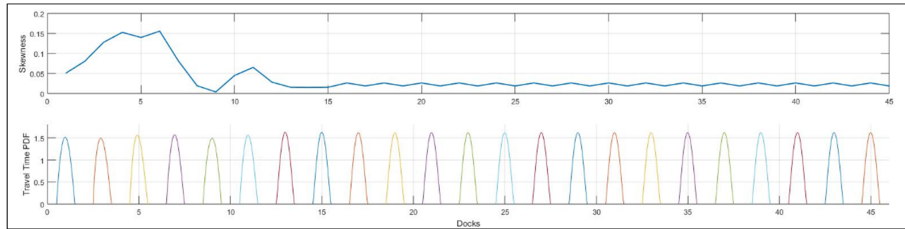


Fig. 2. Skewness for Travel Time Distribution with Distance from Center Dock

2.1 DC Queueing System

The outbound operations in the DC begins by “accumulation” of customer orders from storage area as pallet loads in the dock to form a truck load. The pallet loads then undergo a “wrapping” process to secure items in the pallet after accumulation. The pallet loads are then “inspected” to verify that items in the pallets have completed all processes before loading. Pallets are then “loaded” onto the truck for shipping. The process time associated with accumulation, wrapping, inspection and loading are denoted by Accumulation Time (AT) or Forklift Travel Time, Wrapping Time (WT), Inspection Time (IT) and Loading

Time (LT) respectively. The process time for a truck load, denoted by Truck Processing Time (TPT), is the sum of AT, WT, IT and LT. The operational productivity is expressed by a metric called Labor Hours Per Truck (LHPT). The movement of pallets are done using forklifts and pallet jacks. It is assumed that forklifts perform only accumulation operations and pallet jacks are used by wrapping team, inspection team and loading team for wrapping, inspection and loading respectively. The queuing model is shown in Figure 3. The analytical queuing model facilitates development and evaluation of workforce capacity policies for the DC for several demand scenarios by providing a performance estimate for each policy-scenario combination.

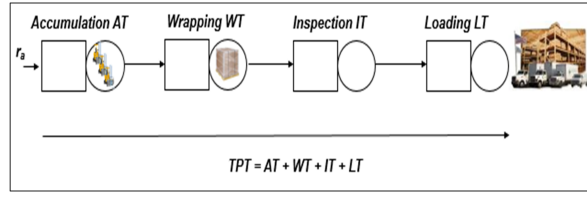


Fig. 3. DC Queuing System

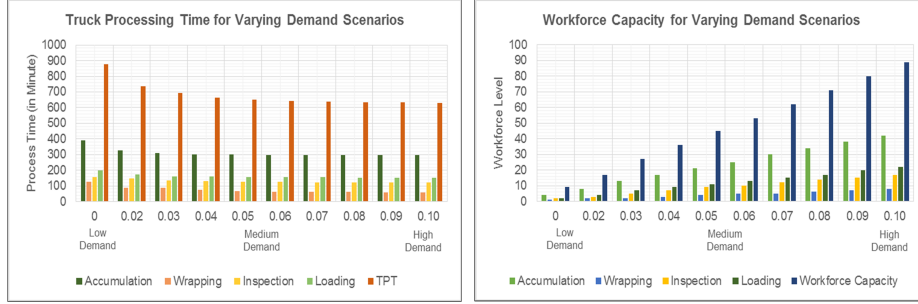
2.2 DC Analytical Queuing Model

This system can be modelled as an open network queueing system with a First Come First Serve (FCFS) queue discipline and an infinite queue space. The customer orders for accumulation arrive at a rate r_a in the form of palletized truck loads. Forklifts, wrapping team, inspection team and loading team act as servers. The system is characterized by a series of G/G/m queues for accumulation, wrapping, inspection and loading with general interarrival and process time distributions. The following parameters are provided as input to the analytical queueing model: r_a - rate of arrivals in truck load per unit time; u - Workforce Utilization; t_a - Average time between arrivals (in minute); c_a - Arrival CV; c_d - Departure CV; t_e - Mean effective process time (in minute); c_e - CV of effective process time. The performance of the queueing system is characterized by the following parameters and are considered as the output from the analytical queueing model: m - Workforce level in a process; CT_q - Expected waiting time in queue for a process (in minute); CT - Expected time for a process (in minute); TPT - Truck Processing Time (in minute); WIP - Average work-in-process level at process (in truck load); WIP_q - Expected WIP in queue (in truck load); ATL - Average Truck Load; LHPT - Labor Hours Per Truck [11].

3 Performance Analysis using Analytical Models

The workforce capacity of the DC is analyzed for several demand scenarios that range from a low demand to a high demand for a given process time at 70% target

workforce utilization for all the process using the analytical queueing model. It is concluded that TPT decreases steadily as workforce capacity increases and LHPT increases linearly with workforce capacity as shown in Figure 3 (a) and (b). This establishes that as more workforce is deployed to process a truck load, the time to process the truck load decreases (service level improves) but labor hour investment on a truck load increases (cost increases).



(a) Truck Processing Time for Varying Demand Scenarios (b) Workforce Capacity for Varying Demand Scenarios

Fig. 4. Truck Processing Time for Varying Demand Scenarios and Workforce Capacity for Varying Demand Scenarios

The Square Root Staffing (SRS) rule is applied in call centers to determine the appropriate staffing levels for an offered load (R) and Quality of Service, (β) and is approximated as $R + \beta\sqrt{R}$, where, R is the amount of work that arrives in the system in unit time. The value of β signifies the operational regime of the system and can be a Quality Driven (QD) regime with emphasis on service level over efficiency, Efficiency Driven (ED) with emphasis on efficiency over service level or Quality Efficiency Driven (QED) regime with trade-off between service level and efficiency. Larger the β value, better the service level, [12,13].

The workforce capacity the DC queueing system is computed by the application of SRS rule for ED ($\beta = 0$), QD ($\beta = 1$) and QED ($\beta = 0.5$) operational regimes. It is observed that workforce capacity increases steadily as DC operates from an ED regime towards a QD regime as shown in Figure 5. This implies that as more workforce is deployed to process a truck load, the service level of DC improves.

4 Performance Analysis using Simulation Model

The DC queueing system is modelled using Simio simulation software for a given interarrival & process time distribution and workforce capacity. The DC simulation model is run for 500 hours with 30 replications after a warm-up period

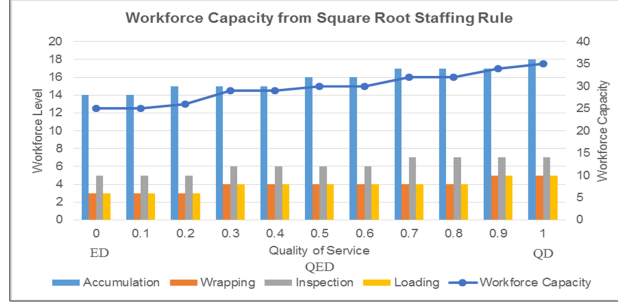


Fig. 5. Workforce Capacity from SRS Rule for ED, QED & QD Operational Regimes

of 100 hours. The results from simulation model is compared with that of the analytical queueing model. The analytical and simulation model results for cycle time and workforce utilization are highly comparable as shown in Figure 4 (a) and (b). The TPT and LHPT are determined with -3.13% and -6.49% error respectively from the simulation model with reference to the analytical queueing model.



(a) DC Cycle Time Comparison: Analytical v/s Simulation (b) DC Workforce Utilization Comparison: Analytical v/s Simulation

Fig. 6. DC Cycle Time Comparison: Analytical v/s Simulation and DC Workforce Utilization Comparison: Analytical v/s Simulation

A simulation-based multi-objective optimization model is developed using OptQuest for Simio add-in. The optimal scenarios for workforce capacity that minimizes TPT and operating cost subject to constraints for a target workforce utilization policy is determined using pattern frontier optimization.

Minimize: TPT and Operating Cost subject to:

- $1 \leq \text{Forklift} \leq 25$
- $1 \leq \text{Wrapping Team} \leq 6$

- $1 \leq \text{Inspection Team} \leq 8$
- $1 \leq \text{Loading Team} \leq 10$
- $u \leq 0.70$

OptQuest optimizes across all responses and finds the set of scenarios that are optimal, rather than a single optimal solution based on weights, [15]. The model is run for 500 hours after a warm-up period of 100 hours with 30 replications for 100 scenarios. It is observed that there are six optimal scenarios with TPT, LHPT and operating cost that range between 628.24 minutes, 340 labor hours & 5915.95 and 689.66 minutes, 279 labor hours & 5000.07 respectively corresponding to a workforce capacity that range between 43 to 33 as shown in Figure 7. It can be concluded from the optimal scenarios that the desired service level can be maintained by adapting workforce capacity to match demand and thereby reducing operating cost by up to 18%.

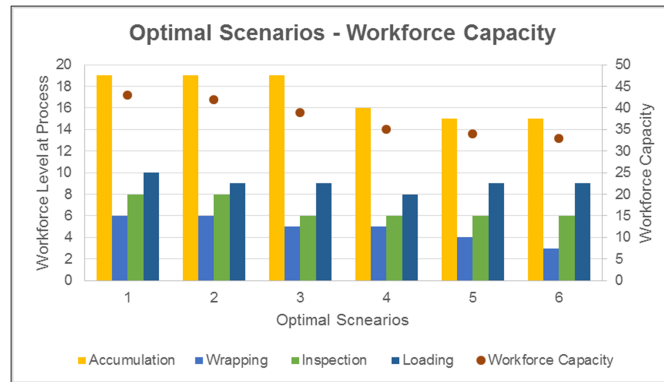


Fig. 7. Optimal Scenarios - Workforce Capacity

5 Conclusions

A decision model is developed to evaluate the impact of workforce capacity on DC performance in terms of Truck Processing Time and Labor Hours Per Truck by considering cost and service level trade-off for varying demand scenarios. Industrial practitioners can use this model for tactical and operational decisions in DCs by application and integration of travel time computation model, analytical and/or simulation model based on the DC geometry considered. The desired service level can be maintained and opportunities to reduce operating cost can be identified by adapting workforce capacity to match demand. Delays in truck processing significantly impacts detention time, the time spent waiting for a truck to be loaded. Increased detention time implies lost revenue by paying late fees to the shipper or missing an opportunity to secure another load. The models

can facilitate industrial practitioners to negotiate for better pricing and terms of contract by reducing detention time substantially. The models presented in this work can be extended to automated DC with conveyors and the resulting energy footprint can be estimated [14,15]. Ongoing work is focusing on implementing these models in a userfriendly spreadsheet tools for industrial practitioners.

References

1. Gu, J., Goetschalckx, M., McGinnis, L.F.: Research on Warehouse Operation: A Comprehensive Review. *European Journal of Operational Research* 177(1), 1–21 (2007)
2. Rouwenhorst, B., Reuter, B., Stockrahm, V., van Houtum, G.J., Mantel, R.J., Zijm, W.H.M.: Warehouse Design and Control: Framework and Literature Review. *European Journal of Operational Research* 122(3), 515–533 (2000)
3. Pandit, R., Palekar, U.S.: Response Time Considerations for Optimal Warehouse Layout Design. *J. Eng. Ind* 115(3), 322–328 (1993)
4. Chew, E.P., Tang, L.C.: Travel Time Analysis for General Item Location Assignment in a Rectangular Warehouse. *European Journal of Operational Research* 112(3), 582–597 (1999)
5. Graves, S.C., Hausman, W.H., Schwarz, L.B.: Storage-Retrieval Interleaving in Automatic Warehousing Systems. *Management Science* 23(9), 935–945 (1977)
6. Bozer, Y.A., White, J.A.: Design and Performance Models for End-of-Aisle Order Picking Systems. *Management Science* 36(7), 852–866 (1990)
7. Koster, R.: Performance Approximation of Pick-to-belt Order Picking Systems. *European Journal of Operational Research* 72(3), 558–573 (1994)
8. Lee, H.F.: Performance Analysis for Automated Storage and Retrieval Systems. *IIE Transactions* 29(1), 15–28 (1997)
9. Hur, S., Lee, Y.H., Lim, S.Y., Lee, M.H.: A Performance Estimation Model for AS/RS by M/G/1 Queuing System. *Computers & Industrial Engineering* 46(2), 233–241 (2004)
10. Berry, J.R.: Elements of Warehouse Layout. *International Journal of Production Research* 7(2), 105–121 (1968)
11. Hopp, W.J., Spearman, M.L.: *Factory Physics*. Waveland Pr Inc, Long Grove, Ill., 3 edition edn. (2011)
12. Technion Israel Institute of Technology), f.U.: <http://ie.technion.ac.il/serveng/>
13. Mandelbaum, A., Zeltyn, S.: Service Engineering in Action: The Palm/Erlang-A Queue, with Applications to Call Centers. In: Spath, P.D.I.D., Fähnrich, P.D.I.h.D.M.K.P. (eds.) *Advances in Services Innovations*, pp. 17–45. Springer Berlin Heidelberg (2007)
14. Krishna, A.G.: *Workforce Planning Models for Distribution Center Operations*. Ph.D. thesis, The Pennsylvania State University, State College (2009)
15. Anand, V., Lee, S., Prabhu, V.V.: Energy-Aware Models for Warehousing Operations. In: Grabot, B., Vallespir, B., Gomes, S., Bouras, A., Kiritsis, D. (eds.) *Advances in Production Management Systems. Innovative and Knowledge-Based Production Management in a Global-Local World*, pp. 390–397. No. 439 in *IFIP Advances in Information and Communication Technology*, Springer Berlin Heidelberg (2014)