MASTER

Optimizing Workforce Size and Workload Distribution in Order Fulfillment Operations
An Aggregate Planning Approach

Azoum, Elias

Award date:
2024

Department of Industrial Engineering & Innovation Sciences

Elias Azoum (0860047)

Optimizing Workforce Size and Workload Distribution in Order Fulfillment Operations: An Aggregate Planning Approach

Supervisors:

First supervisor	Dr. K. (Karel) van Donselaar
Second supervisor	Dr. ir. M. (Melvin) Drent
Company supervisor	T. (Ted) Vertuin

Zeist, 04-07-2024

Abstract

As distribution centers face increasing capacity constraints, aligning the workforce with fluctuating demand is essential. This research explores an optimized aggregate planning approach to balance workforce size and workload distribution in order fulfillment operations. Focusing on a large wholesaler, the study examines the importance of demand accuracy and the systemic role of workload distribution throughout the day. An analysis of over 18 million order lines revealed distinct demand patterns, with workload peaking in the afternoon time intervals and exerting considerable pressure on the final workforce shift, amounting to expensive overtime and productivity inefficiencies. To address this, a mixed-integer quadratically constrained programming (MIQCP) model was developed to determine the optimal workforce size for each time interval. Testing the model with four different demand approaches showed significant reductions in overtime and operational costs, while increasing order completion rates. Impact analysis further suggested that shifting a portion of the peak workload to the next day or implementing a 48-hour delivery policy could eliminate overtime and reduce costs beyond the best demand approach. Consequently, this research presents a planning framework for managing resources in operations where building inventory or allowing backorders is not feasible. Future research could explore the workload distribution adjustments, along with factors such as employee satisfaction, long-term sustainability, and environments with greater demand variability.

Keywords: *Aggregate planning, workforce optimization, workload distribution, stochastic demand, order fulfillment, mixed-integer programming, MIQCP, operations management*

Executive summary

Problem statement

Distribution centers globally face the challenge of capacity expansion to meet increasing market requirements. A key issue is managing fluctuating demand with the existing workforce as effectively as possible, given the scarcity of resources. In this context, aggregate planning becomes crucial. It involves balancing the workforce size with incoming workload over a given period. However, the irregular pattern in which workload arrives throughout the day—i.e., the workload distribution—makes allocating workforce based on daily, weekly, or even monthly demand levels a complex task. This complexity is especially pronounced in order fulfillment environments, where building up inventory or allowing backorders is not an option. In these settings, the workforce can only commence production as the workload arrives, unlike in make-to-stock environments where production can occur continuously. To address this issue, the optimization potential of workforce planning at a large wholesaler has been investigated. The study initially focused on the importance of demand accuracy, followed by an examination of the systemic role of workload distribution throughout the day. This study can be encapsulated with the following main research question:

How can an aggregate planning approach effectively optimize workforce size and workload distribution in response to stochastic demand in order fulfillment operations?

Process and data analysis

The analysis of the current situation provided valuable insights into workload, workforce, and productivity rates. By examining the workflow of order fulfillment departments, the interdependency between activities, and resource utilization, the current state could be accurately quantified using relevant datasets.

<u>Workload Insights</u>

The company handled over 18 million order lines. Several demand patterns emerged:

1. *Seasonal Patterns*: Demand is lower during summer and holiday periods, while it remains stable during the rest of the year, with a coefficient of variation of 0.21.
2. *Weekly Patterns*: The highest demand occurs on Mondays and gradually decreases throughout the week, with the lowest on Fridays. There is no demand on weekends, as the company operates on a business-to-business model.
3. *Daily Patterns*: The daily order window, which closes at 8pm for next-day delivery, shows a clear workload distribution throughout the day. Figure 1 illustrates the percentual share of daily workload per hourly interval. For example, approximately 5% of daily demand arrives at 11, denoting the time interval between 11:00 and 12:00. This pattern is consistent across weekdays and departments.

<u>Workforce Insights</u>

Daily workforce planning must align with the workload distribution. The day starts with ample workload due to orders placed after the previous day's deadline. However, this workload quickly diminishes, leaving employees underutilized. In the afternoon, orders surge, peaking around 4pm. This late surge overwhelms the workforce, often necessitating overtime at premium costs to complete all orders.

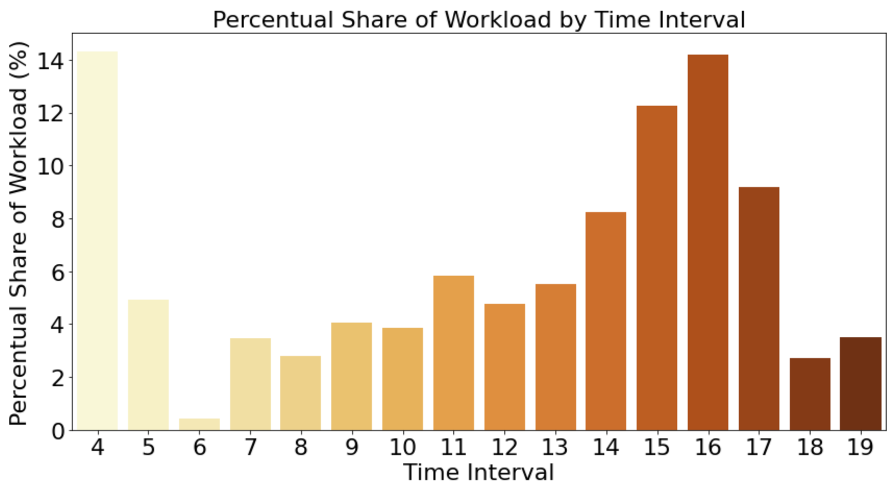

Figure 1: Percentual share of workload per time interval

<u>Productivity Rate Insights</u>

The analysis highlighted the need for a balance between workload and workforce. Allocating too many employees in a given interval can quickly reduce the workload, leaving the workforce idle during less busy times. Conversely, if the morning workload is not completed before the afternoon peak, overtime may be required. Overstaffing can also increase the likelihood of malfunctions, impacting productivity. Understaffing, on the other hand, can cause inefficiencies due to the nature of order fulfillment at the wholesaler. Thus, finding the ideal balance across all time intervals is imperative. The data analysis established a relationship between workload, active employees, and productivity rate, which can be utilized for mathematical optimization.

Optimization model

The goal is to determine a workforce size for each time interval such that the total daily workload is met at minimum cost, subject to specific constraints. Figure 2 illustrates the hierarchical structure of the planning assignment. For each weekday, a plan must be created for four order fulfillment departments. Each department can be staffed by four types of employees: regular full-timers, regular part-timers, subcontracting full-timers, and subcontracting part-timers. Employees work in either the morning, afternoon, or evening shift, with each shift consisting of specific time intervals corresponding to the department and employee type. Individual employees are assigned to these intervals, and the sum of these assignments determines the workforce size for each time interval. Since employees are unique, overlap must be prevented and working requirements must be respected.

The mathematical model aims to determine whether each employee $i \in I$ should be allocated in time interval $t \in T$ of shift $s \in S$ as employee type $e \in E$ for department $d \in D$ on weekday $wd \in WD$, as denoted by the binary variable $x_{d,e,i}^{wd,s,t}$. The workload is incorporated with auxiliary variables, and parameters have been added that represent the workload distribution. The relationship between workforce and workload, derived from the data analysis, is converted into piecewise linear relationships to ensure optimality. Cost parameters and constraints are further defined to minimize the following objective function:

$$\sum_{wd}^{WD}\sum_{d}^{D}\sum_{e}^{E}\sum_{s}^{S}\sum_{t}^{T}\sum_{i}^{I}(C_{d,e}^{t} * x_{d,e,i}^{wd,s,t})$$

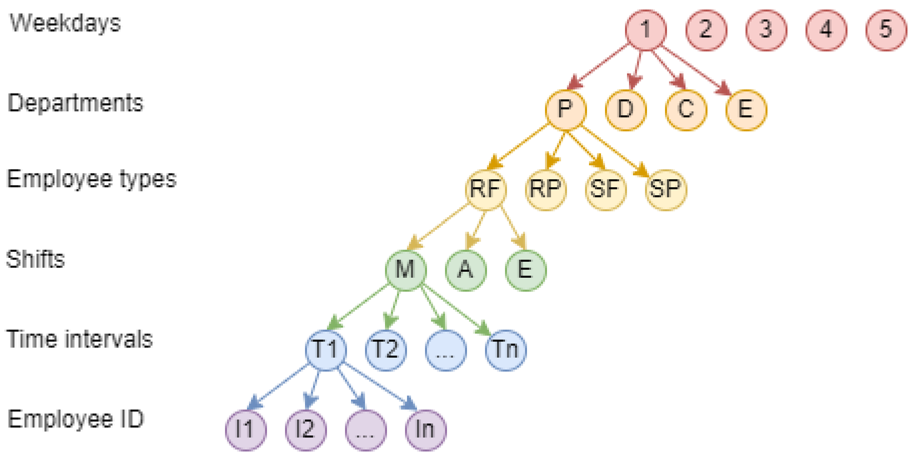

Figure 2: Hierarchical levels of the workforce planning

Results

The research was conducted against the backdrop of two critical elements: demand accuracy and the systemic role of workload distribution. An evaluation design was created for these elements and applied to the case study environment.

<u>Demand Accuracy</u>

The first part examined how different demand inputs could affect workforce planning performance, with the workload distribution fixed to represent the current situation. Four approaches were tested:

1. *Approach 1*: Used the company's current forecast and workforce allocation method as the baseline.
2. *Approach 2*: Used the same forecast method but optimized workforce allocation using the mathematical model.
3. *Approach 3*: Applied the Newsvendor model to account for stochastic demand with optimized workforce allocation.
4. *Benchmark*: Assumed demand was known beforehand, representing a deterministic scenario as the upper limit of the optimization framework.

The results for the four approaches are summarized in Table 1, showing the mean performance during the review period of 2023.

Table 1: Performance of the tested approaches

	Workforce size	**Overtime hours**	**Costs**	**Productivity rate**	**Percentile completed**
Approach 1	737 hrs	47.8 hrs	€21,167	109.41 picks/he	94.69%
Approach 2	740 hrs	14.4 hrs	€20,884	110.91 picks/he	98.41%
Approach 3	739 hrs	16.8 hrs	€20,882	111.04 picks/he	98.01
Benchmark	735 hrs	11.8 hrs	€20,715	111.77	98.61%

Key Findings:

- *Optimization Impact*: The difference between the current situation (Approach 1) and the optimized workforce planning (Approach 2) is substantial, with improvements across all metrics. The optimized model slightly increased overall hours used but significantly reduced overtime by almost 70%, leading to cost savings and higher productivity rates (in picks per hour per employee, picks/he).

- *Stochastic Demand*: Approach 3, incorporating stochastic demand, performed marginally better than Approach 2, likely due to already stable demand.

- *Benchmark Performance*: The benchmark approach showed the best performance, though the improvements were not vastly superior. Persistent overtime hours suggest systemic issues in workload distribution, even with perfect demand knowledge.

<u>Systemic Role of Workload Distribution</u>

While the initial results assumed a fixed workload distribution parameter, the impact analysis tested the effects of altering this parameter through two policies:

1. *Policy 1*: Shifted a percentage of the workload arriving after 17:00 to the next day. The results, using Approach 1 as the baseline, showed improvements in workforce size, daily costs, and workload completion percentages as more workload was shifted. Significant gains were observed up to a 30% shift, after which returns diminished, as visualized in Figure 3.

2. *Policy 2*: Shifted the entire day's workload to the next day, effectively implementing a 48-hour delivery policy. This policy eliminated overtime, further reduced costs, and increased productivity rates, achieving nearly 100% workload completion before midnight, as shown in Table 2.

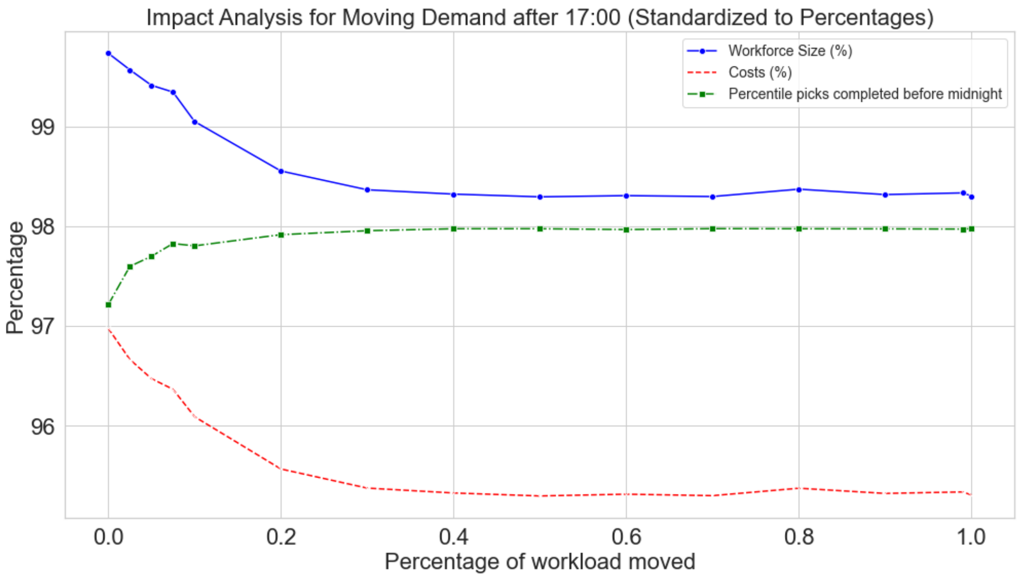

Figure 3: Impact analysis of policy 1

Table 2: Performance of policy 2 compared to baseline

	Workforce size	Overtime hours	Costs	Productivity rate	Percentile completed
Approach 1	737 hrs	47.8 hrs	€21,167	109.41 picks/he	94.69%
Policy 2	726 hrs	0 hrs	€20,321	112.39	99.99%

Conclusions and recommendations

The research concluded that an optimized aggregate planning approach, which incorporates accurate demand forecasting and strategic workload distribution, significantly enhances workforce efficiency and reduces operational costs. To achieve these improvements, it is recommended to replace existing planning methods with an optimization model that accounts for weekdays and specific time intervals. Additionally, investing in advanced forecasting tools to provide more precise demand input for workforce planning could lead to substantial improvements.

However, systemic issues related to workload distribution indicate that part of the problem may persist. Therefore, exploring practical methods for shifting a portion of peak workload to the next day could further enhance performance and result in larger yearly cost reductions. Another recommendation is to consider shifting the entire day's workload to the next day, effectively implementing a 48-hour delivery policy, which could completely eliminate overtime.

Future research should explore how desired workload distributions can be effectuated and assess the potential corresponding costs. Furthermore, it would be worthwhile to include factors such as employee satisfaction, sickness leave, and the long-term sustainability of workforce practices to provide a more holistic view of workforce planning.

Preface

First and foremost, I would like to take this moment to express my gratitude to my parents.

Secondly, these past few months have been a whirlwind of experiences, filled with challenging moments, growth, intellectual stimulation, and excitement. I am happy to have embarked on this project surrounded by a network of diverse and talented individuals.

Karel, Ted: thank you for your guidance and supervision. Your expertise and cooperation have been invaluable in shaping the outcome of my research. I also appreciate the little things, such as Karel responding to my countless short-notice Teams requests, or Ted acknowledging my preference for irregular office hours. It has helped me find my rhythm throughout the project.

Allison: thank you for being so conscientious about data, and for your active participation and interest in my progress.

Melvin: thanks for facilitating this project in the first place, and the useful feedback you provided.

My colleagues: thank you for the kind reception and the generous time you dedicated to supporting me.

Arsenal FC: thank you for an amazing year of football. Next season we go again. 2025 PL. 2026 CL.

Elias Azoum,
Zeist, July 2024

Contents

Abbreviations

AS/RS	Automated Storage and Retrieval System
CoV	Coefficient of Variation
DC	Distribution Center
DOL	Distribution Order Line
DPS	Dynamic Picking System
iLPN	Inbound License Plate Number
MAPE	Mean Absolute Percentage Error
MIQCP	Mixed-integer Quadratically Constrained Programming
MIP	Mixed-item Pallet
OCB	Order Consolidation Buffer
oLPN	Outbound License Plate Number
P&D	Pick & Deposit
R2	R-squared
RMSE	Rooted Mean Squared Error
SKU	Stock-keeping Unit
SIP	Single-item Pallet
TSP	Turret Stock Picker

List of Figures

List of Tables

Chapter 1 – Introduction

1.1 Problem background

The recent expansion of distribution centers (DCs) is a prominent trend fueled by e-commerce growth, globalization, and technological advancements. An industry report from 2023 valued the global distribution market at $7.72 trillion in 2022, projecting it to reach $11.93 trillion by 2032 with an annual growth rate of 4.5% (Benchmark International, 2023). Companies like Amazon exemplify this trend by continually opening new DCs to meet online retail demands (World Trade Organization, 2023). This expansion clearly highlights not only the growth of individual companies but also a broader industry shift towards rapid processing and delivery to satisfy market demands.

Amid this expansion, DCs globally encounter several operational challenges, particularly in order-picking processes. The diverse product range and the expectation of next-day, error-free deliveries increase the complexity, which affect productivity and cost-efficiency (Törn, 2022). Additionally, supply chains are evolving towards mass customization and personalized goods, posing major hurdles to traditional bulk handling and shipping methods (Törn, 2022).

One critical challenge centers around labor capacity constraints, a persistent issue exacerbated by various factors. Notably, historically low unemployment rates, a shrinking job seeker pool, an aging workforce, and pandemic-related health concerns have all contributed to labor shortages in order-picking roles (Romaine, 2021). Demographic shifts, changing market dynamics, and evolving work practices in distribution setups further compound these constraints. A pressing concern surrounding these challenges is the widening gap in skilled labor, stemming from an aging workforce and the diminishing appeal of manual labor roles among younger generations (McRary, 2023). The physically demanding and repetitive nature of the work makes it difficult to attract and retain talent, particularly in a competitive labor market (EAM Mosca, 2022).

Another major challenge facing DCs is the substantial rise in operational costs, driven primarily by external factors. The war in Ukraine has spiked energy prices, and China's stringent COVID-19 policies have disrupted the global supply chain, leading to extended financial pressures (Accenture, 2022). These global dynamics have increased the costs of running DCs. To tackle this financial challenge, DCs must optimize workforce productivity and make the most efficient use of existing resources. Strategic areas for improvement include staff deployment, reduction of superfluous expenses, and better utilization of facilities and equipment (Sunol, 2023). By maximizing resource use and workforce efficiency, DCs can strive to reduce overhead and streamline operations, thereby maintaining competitiveness.

Managing varying workloads in DCs, however, is a complex endeavor that requires careful planning and optimization (Deloitte, 2016). The fluctuating nature of demand, including seasonal peaks, promotional events, and market trends, causes problems in workforce and resource allocation that are difficult to solve. DCs must balance the need to meet peak demand periods without the cost of surplus resources during slower periods (Myers, 2021). Finding the ideal balance therefore necessitates flexible staffing models and a scalable approach to workforce planning. In this problem area, aggregate production planning (APP) emerged as the research field that offers a framework to harmonize these elements, providing a strategic approach to workforce and operational planning in DCs.

1.2 Problem statement

The company at the center of this master thesis is a wholesale distributor in the Netherlands. The company's unique selling point is its extensive product availability, including hard-to-find items in various sizes and variations, with a promise of next-day delivery. This strategy has driven consistent annual growth, resulting in a vast network of suppliers and customers. To accommodate its daily customers as best as possible, the company operates a well-distributed network. The company has three DCs in the country: one for large items, and two for smaller products, which constitute the majority of orders. After processing and order fulfillment, goods are dispatched to regional transfer points nationwide and delivered to customers in smaller vans. The company also operates physical stores for direct-to-consumer sales and a general sales office, totaling over 40 selling locations across the country.

With its annual growth, the company faces new challenges. One primary concern is the increasing need for warehouse space due to the expanding variety and quantity of products, prompting the company to initiate expansion plans. Additionally, managing the surge in customer orders presents urgent logistical complexities. The company's policy of next-day delivery for orders placed up to 8pm adds to these challenges. The reason is that the pattern in which orders arrive is highly imbalanced. The order distribution is very uneven during the day, with the majority of orders arriving in the late afternoon. Since the peak order arrivals vastly exceed the capacity, it becomes difficult to complete all orders in time. This uneven distribution of orders is illustrated in Figure 4.

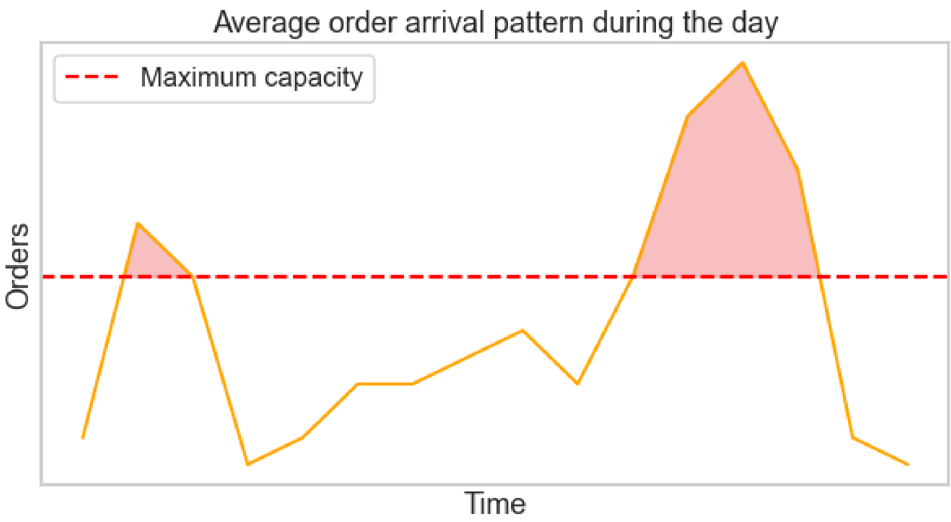

Figure 4: Order arrivals during the day

In an ideal scenario, the order arrivals are constant, which would allow for uniform workload distribution and deterministic planning. However, at the company, over 40% of orders arrive after 3pm, with the majority concentrated between 3pm and 6pm. Moreover, the absolute value of the order arrivals is largely stochastic, fluctuating across weekdays and seasons. With a nearly constant workforce throughout the day, this misalignment results in inefficiencies, as there are too many or too few employees working during certain time intervals. These inefficiencies lead to multifaceted issues affecting several aspects of the operation.

First, morning capacity cannot be reserved for later use. This principle of factory physics means excess morning resources are wasted, as employees cannot pick orders that have not yet arrived. Second, the afternoon peak in order arrivals places substantial strain on the evening shift, exceeding the workforce's capacity. Although employees are scheduled to work until midnight, they often extend beyond this time to complete remaining orders. Third, the partial mechanization of the order-picking system introduces a capacity constraint, which prevents the company from simply adding more employees to increase productivity linearly. When mechanized processes operate at full capacity, especially in the evening, the likelihood of system failures and other malfunctions also increases, further pressuring the final shift of the day for timely order completion. When overtime is required, premium costs will be incurred.

In short, the company's operational issues can be categorized into three main areas: inefficient resource allocation results in unproductive expenses; overtime past midnight incurs additional costs due to higher wage rates; and, finally, overtime can lead to employee dissatisfaction, increasing the likelihood of sick leave. The company is therefore encouraged to find the ideal workforce planning and workload balance that minimizes costs, eliminates overtime, and maximizes the employee experience.

1.3 Literature study

1.3.1 Classical aggregate planning

The concept of aggregate planning was pioneered by Holt, Modigliani, Muth, and Simon in a seminal 1955 paper. This work significantly influenced the field, with their HMMS model laying the groundwork for subsequent research. The central question they addressed was how a factory should balance its production rate and workforce in response to order fluctuations, aiming to minimize overall costs. Holt et al. (1955) highlight that although stable order arrivals would simplify planning, reality often displays large variability in order patterns. To manage these fluctuations, Holt et al. (1955) suggest three approaches: modifying workforce size, adjusting production rates, or using inventory and backlog strategies.

In the first approach, workforce adjustments are made to align with order fluctuations, leading to hiring when orders increase and layoffs when they decrease. However, expanding the workforce often requires training and reorganization, while reducing it can lead to costs like terminal pay and diminished worker morale. In contrast, Holt et al. (1955) suggest a second approach that maintains a constant workforce where order fluctuations are absorbed through overtime. With sufficient personnel, all orders can be met if the workforce remains active long enough, but the limit to overtime means peak orders dictate the workforce size. Naturally, a drop in orders eliminates overtime but eventually leads to undertime where the workforce is not fully utilized. The final approach manages fluctuating demand by varying the inventory of finished goods or, if inventory is unavailable, allowing for backorders.

In the HMMS model, a quadratic cost function is employed, resulting in several constants that must be established (Holt et al., 1956). These constants form the basis for a Linear Decision Rule (LDR) that consists of two decision-making steps applied at each time interval. These steps, based on the constants and demand forecasts, determine the aggregate production rate and workforce size. The inventory level follows implicitly from the two decision rules. While the model incorporates sales forecasts, its design minimizes sensitivity to forecast errors because each decision mainly impacts the immediate period. Decisions for future periods are subsequently made when more information is available. Van de Panne & Bosje's (1962) reinforced this aspect as their sensitivity analysis showed that errors in the cost function underlying LDRs did not critically affect costs.

In their case study of a paint factory, Holt et al. (1955) demonstrated that applying the decision rules and a simple moving averages forecasting method could result in a minimum cost saving of 8.5% compared to the actual outcomes. This paper established the HMMS model as an effective baseline in the field for determining workforce size and production rate in an easily interpretable manner for each chosen period. Building on this, Schild (1959) generalized the inventory component of the HMMS model by introducing an interest rate factor, considering setup costs, and allowing for the independent aggregation of individual items.

Despite its relative ease of use, the HMMS model has not been widely adopted in the industry. One reason is the quadratic cost assumption, which can lead to frequent workforce adjustments as these functions tend to undervalue minor changes. Additionally, doubts about the LDR's ability to accurately capture a company's true cost have been raised (Singhal & Adlakha, 1989). These concerns have spurred a new branch of aggregate planning research focused on linear programming (LP) models, which assume costs to be linear or piecewise linear (e.g., Hanssmann & Hess, 1960; Bowman, 1956; Charnes, Cooper & Farr, 1953).

LP models in the context of APP typically assume deterministic demand and linear costs, and users are required to define all variables related to production for each period. Unlike LDRs, it is possible to add production or workforce constraints in the form of inequalities. The primary goal remains cost minimization, but the original production and workforce factors become slack variables within the LP framework (Hanssmann & Hess, 1960). These problems can subsequently be solved by the Simplex method with the aid of a computer.

Multiple variations of the LP application can be found in existing literature. Orbeck, Schuette & Thompson (1968), for instance, extended the model of Hanssmann & Hess (1960) by incorporating productivity rates. Furthermore, Bowman (1956) suggested the LP problem of aggregate planning could be placed into a transportation-method framework, treating production decisions as sources for supply or input and production requirements as a destination or output. Posner & Szwarc (1983) further developed this framework by adding inventory backlog into the model.

LP models indeed offer a versatile approach to aggregate planning. However, like the strictly quadratic functions in LDRs, the assumption of linear costs may not always be accurate. Another considerable limitation is the treatment of demand as deterministic. While this simplifies computation, demand in reality is often highly uncertain. Hanssmann and Hess (1960) therefore suggested using probability distributions to estimate total expected costs, which they found to yield satisfactory results with moderately reliable forecasts. Nonetheless, the applicability of this approach may vary depending on specific situations and conditions.

The limitations of the LDR and LP models have prompted research into alternative models. Nam & Logendran (1992) categorize linear and quadratic function methods as "optimal solution methodologies", as these methods invariably yield an optimal solution. However, these methods could fall short in practical applications. In contrast, "near-optimal solution techniques" strive to closely approximate an optimal outcome. Though not achieving exact optimality, these methods are more sophisticated in their replication of real-world conditions.

One such technique is the Search Decision Rule (SDR), which blends branch-and-bound and search methodologies, as explored by Taubert (1968) and Flowers & Preston (1977). This method addresses non-linear problems—problems that were previously difficult to solve—by segmenting feasible solutions for production rate and workforce size into simpler aggregate scheduling issues. While the resulting solutions may not always be optimal, the SDR method offers increased realism, broader applicability, and the ability to add dimensions like multiple products and subcontracting. Results showed comparable costs to the LDRs of Holt et al. (1955), with the added benefit of greater reliability in practical scenarios (Taubert, 1968).

Other heuristic approaches involve similar computationally intensive methods. Jones (1967) introduced Parametric Production Planning where, after having established two decision rules, the corresponding ideal parameters were searched across a wide range of possibilities. This approach allows for flexibility in cost structure, not being restricted to linear or quadratic forms. Similarly, Vergin (1966) applied simulation to assess various production and workforce schedules, attempting to identify optimal parameters in complex cost scenarios.

1.3.2 Modern aggregate planning

As the field of aggregate planning has evolved, it has increasingly integrated concepts from Operations Research, Supply Chain Management, and Industrial Engineering, especially focusing on uncertainty (Jamalnia et al., 2019). Unlike earlier models that often assumed deterministic demand and singular objectives, recent approaches encompass a broader spectrum, including warehouse space, multi-echelon networks, multiple sites, carbon emissions, and workload management. This newer phase of APP literature emphasizes the inherent uncertainty in demand, costs, and other critical planning parameters.

One of the focal research areas that has garnered considerable attention in this respect is the application of stochastic mathematical programming. Thompson, Watanabe & Davis (1993), for example, developed LP models to evaluate different APP policies where customer demand was modeled with empirical probability distributions. The performance of the most effective strategies proved to be only a small percentage below the situation with perfect information. Leung & Lai (2006) advanced this approach by creating a stochastic LP model for optimizing multi-site production plans. This model considered various costs, including production, subcontracting, and labor, with a notable inclusion of penalty costs for underutilized resources under different economic prosperity scenarios. In their study, uncertain demand was accounted for by a discrete probability distribution, assigning a likelihood of occurrence to scenarios ranging from ‘booming’ to ‘poor’ (Leung & Lai, 2006).

A nonlinear variant tailored to APP across multiple products, sites, and periods in a green supply chain was proposed by Mirzapour, Baboli & Sazvar (2013). Demand uncertainty here was modeled by specifying a distribution prior to the analysis, which by means of demonstration was chosen to be Normal. After applying numerical techniques, the nonlinear mixed integer problem was converted into a linear model, leading to a global solution. In the realm of stochastic multi-objective programming, real-world production systems often juggle multiple objectives across several functional areas. Nam and Logendran (1992) noted that these objectives typically include cost minimization, profit maximization, inventory reduction, shortage level minimization, stabilization of production output or workforce levels, and maximizing the utilization of workforce, machines, and equipment.

Jamalnia et al. (2017) proposed a different stochastic, multi-objective optimization model for APP that focuses on a "mixed chase and level" strategy in uncertain markets. This model recognizes demand as the primary source of uncertainty. In a chase strategy, workforce adjustments directly respond to demand fluctuations without building inventory. Conversely, a level strategy maintains constant output, relying on inventory or overtime to address any discrepancies when demand exceeds this constant production level.

The model of Jamalnia et al. (2017) involves objectives such as total revenue, total production costs, total labor costs, optimal utilization, and customer satisfaction, validating it with real-world data from a beverage manufacturing company. Uniquely, it handles the APP problem on a daily level as it distinguishes between two 8-hour shifts, highlighting the wage cost differences that are dependent on the time of the day. This approach is a departure from regular APP models where longer time horizons are more common (Demirel, Özelkan & Lim, 2018). Moreover, by using pricing and advertising tactics, the model enables business managers to shift demand from peak to off-peak periods, adding flexibility to manage periods when demand is forecasted to exceed workload capacity (Jamalnia et al., 2017).

1.3.3 Research gap

The field of aggregate planning, while extensively explored, reveals notable research potential in the application of APP to environments where inventory building is constrained. Such order fulfillment environments typically operate on a make-to-order basis, contrary to the manufacturing contexts discussed in traditional literature that allow for production to commence ahead of incoming demand. Given this distinction, exploring how demand varies *within* a period becomes a crucial area of investigation, extending research beyond the usual emphasis on demand fluctuations *across* periods.

Contemporary methods in APP predominantly address uncertainties through stochastic mathematical programming, robust optimization, and multi-objective approaches, assuming a set workforce level that perfectly matches production output. However, they often overlook the dynamic nature of variables during the operative period that can diminish production output. In this regard, Jamalnia et al. (2017) initiated a refined focus within APP by considering different policies for various shifts within a day and attempting to alter production levels by influencing day-specific demand. Nonetheless, there is still a need for more specialized strategies tailored to operational contexts where production is dictated by the moment orders arrive and inventory cannot be pre-built. In such settings, workforce levels deemed optimal by conventional APP methods do not account for inefficiencies that may arise due to irregular order arrivals throughout the day, potentially rendering the solution suboptimal.

Building on that premise, this thesis research aims to investigate the effects of APP approaches on a more granular level (i.e. per hour instead of per day or week). The research focuses on analyzing and optimizing workforce size and production plans not only across days but also within individual hourly time intervals of a day, acknowledging that production output in practice is often nonlinear. By integrating this granular perspective within the broader scope of aggregate planning, the thesis introduces a novel approach to workforce and production planning in order fulfillment environments.

1.4 Research questions

The primary objective of this master thesis project is to assess the design of the company's current operations and explore ways to alleviate the end-of-day strain in the order-picking process, such that the identified issues can be minimized or resolved entirely. Given the company's strict promise of next-day delivery and its product model where order fulfillment commences only after receiving an order, building inventory or allowing backorders is not possible. More generally, from an aggregate planning perspective, the focus is set on determining how many of the order arrivals should be fulfilled at what time and by how many employees. The goal is to optimally balance workforce size and workload distribution. The essence of this research is encapsulated in the following main research question:

Main research question*: How can an aggregate planning approach effectively optimize workforce size and workload distribution in response to stochastic demand in order fulfillment operations?*

To address the main research question, it has been divided into a series of smaller research questions, which will be discussed step by step below and form the outline of the remainder of the thesis report.

Process analysis

The company's order fulfillment operations are complex due to their large scale, involving numerous processes and departments that are interlinked. Understanding this network requires examining the process from the moment a customer places an order online up until it leaves the DC. This process involves various entities, stakeholders, and employees, each contributing in their unique roles. It is important to first map all activities and variables that affect costs and performance, along with other key indicators. This leads to the following two research questions, which will be answered in Chapter 2:

Research question 1: How is the current order fulfillment workflow structured from order arrival to delivery?

Research question 2: What are the critical KPIs, parameters, and processes relevant to optimizing this workflow?

Data analysis

Understanding customer behavior and the stochastic nature of order arrivals is a critical aspect of the order fulfillment operation. Analyzing historical data to identify possible seasonal and weekday patterns can help predict demand fluctuations. Further data analysis seeks to unravel patterns in the order fulfillment process because, in addition to longer-term trends, it is imperative to dissect demand trends throughout the day. This way, the distribution of the workload over shifts and departments can be captured. By examining workforce performance across various time intervals, insights into productive output, potential inefficiencies, and encountered challenges can be obtained. Taken together, these analyses provide a full view of existing workload demands and production capabilities, prompting the following research questions:

Research question 3: What trends and patterns are observable in order arrivals and fulfillment over different periods?

Research question 4: Are there any unique or irregular patterns in order arrivals and fulfillment that could impact workload distribution?

These research questions are at the center of Chapter 3.

Model design and optimization

The first couple of research questions establish a foundation to identify the key factors that drive costs and other performance indicators. The next phase comprises a detailed formulation and quantification of the cost structure. This step encompasses defining an objective function, decision variables, and constraints for the aggregate planning model. Additionally, understanding the role of accurate forecasting and workforce optimization reveals the potential for improvement given the systemic nature of the current workload distribution. This forms the basis for the following research questions:

Research question 5: What cost function can be defined to quantitatively measure the performance of the current process?

- *What are the key decision variables and constraints in this context?*
- *How does the current process perform in terms of the predefined KPIs?*

Research question 6: How does the current situation compare to:

- *An optimized scenario using the current forecast method?*
- *An optimized scenario incorporating stochasticity?*
- *An optimized scenario with a perfect forecast?*

Chapter 4 will be devoted to answering Research Question 5. Then, in Chapter 5, the case study design will be elaborated. Finally, Chapter 6 reports the results of the case study, thereby answering Research Question 6.

Impact analysis

The frequency of order arrivals per time interval directly determines the workload in said interval, and when aggregated over a whole day, a distribution across time intervals can be derived. To assess the role of the current workload distribution in order fulfillment, the effects of altering this distribution compared to the current situation must be investigated. The outcome of this analysis reveals whether adjustments to the workload distribution are beneficial. If positive outcomes are found, it would be logical to determine the optimal extent of these changes and the point at which the benefits begin to diminish, culminating in the final research question:

Research question 7: What impact would adjusting the distribution of workload across time intervals have on overall efficiency and costs?

The final research question will also be answered in Chapter 6.

1.5 Research scope

This thesis project aims to accurately replicate the operational environment of the company to address the three issues outlined in the problem statement. To achieve this goal, the project involves documenting the internal processes, analyzing the interdependencies between activities and departments, and measuring the impacts resulting from proposed optimizations or adjustments. Given the time constraints of this project, it is not feasible to cover every aspect in-depth. Therefore, clear boundaries and assumptions are necessary.

The project will focus on the relationship between workforce size and workload distribution within one main DC and its respective order fulfillment departments. An order is considered to have arrived at a DC as soon as it is released by the warehouse management system, typically soon after the customer has placed the order. The order is fulfilled once it has been picked and prepared for transport to one of the regional transfer points across the country. The transport itself to the transfer points and the subsequent delivery to the customer's location are excluded from the scope of this project because these processes do not influence the workforce size and workload within the DC. Moreover, since all orders are required to be fulfilled by a set nightly deadline, the departure time of the drivers is a known and fixed variable and is not relevant to this analysis.

While the problem statement highlights issues related to employee and customer satisfaction, this study will not delve into the psychological or mental impact of policy changes. Instead, it aims to provide quantitative justifications for how the company in particular, and the industry in general, might approach workforce sizing and workload distribution long-term. Therefore, exploring the mental effects of policy changes falls outside the scope of this study.

Chapter 2 – Process analysis

2.1 DC Overview

The DC at the center of this master thesis (hereafter DCA) is segmented into four large halls, housing the departments Arrivals, Returns, Production, DPS, Cable, and DCE, each being responsible for a specific set of tasks. Figure 5 shows a schematic overview of these distinct areas within the DC. To gain a better understanding of the processes, this chapter will elaborate on each department and the areas in which they operate. The emphasis is set on the flow of goods through all areas, starting from the point a supplier shipment arrives at the DC on a pallet as inventory until the moment a good has been picked as part of an order, leaving the DC to be transported to the regional transfer point. Moreover, the workflow that emerges from the analyses will be contrasted with the resources that are utilized to execute the tasks of the workflow.

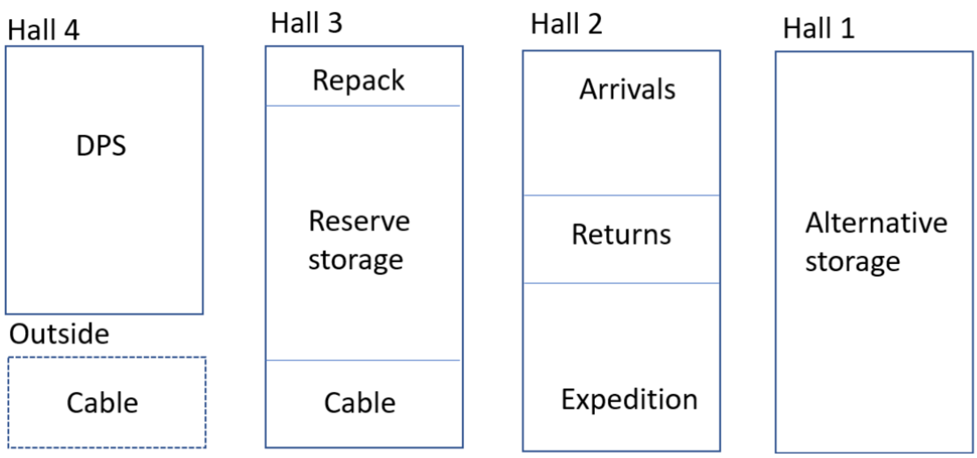

Figure 5: Layout of DCA

2.2 Arrivals

The arrival area of the DC, located in hall 2, is where incoming goods from the suppliers are handled. The department responsible for operations here is called Arrivals and is one of the two departments not part of order fulfillment activities, together with Returns. In the arrival area it is made sure newly supplied goods are received correctly, checked upon completeness, and that these goods are transported to the correct place within the DC. Once a supplier arrives at the DC, they announce themselves at the reception, after which the unloading can begin. The goods are delivered on pallets on which either a mixture of stock-keeping units (SKU) or a single SKU are placed.

Based on the incoming license plate number (iLPN), which is a label assigned to an item once it has entered the DC, and the appearance of the pallet, the employee can swiftly deduct whether there are multiple products on a pallet or whether it is just a single product type. Depending on the homogeneity of a pallet, two things are possible. If a pallet consists of a single item, these type of pallets are denoted as a 'single-item pallet' (SIP) and do not need further processing. This pallet is placed on the conveyor belt, which moves directly to the reserve storage & repack area and takes place at the pallet queue in hall 3. Further processing is handled in this area and will be discussed in section 2.4.

If the pallet consists of multiple product types, it is classified as a 'mixed pallet' (MP) and requires manual processing at the MP workstation. A computer-aided system directs these goods to one of three possible destinations: reserve storage, alternative storage, or direct replenishment of the inventory at the DPS. For items designated for reserve storage, they are placed on the same conveyor belt as the previously discussed SIPs. However, each product type must be placed on a separate pallet before being sent on the belt, effectively converting them into SIPs. Items destined for alternative storage are manually transported there. Lastly, if an item on a pallet is meant for DPS, it bypasses reserve storage and is added directly to the DPS inventory from the MP workstation. In this case, the item will be removed from the pallets, and each SKU of this item will be placed in green totes, which is the smallest bin used in the DC for storing and moving individual, conveyable items. This process is called 'repacking', which means preparing units to be picked as separate entities, unlike the full-sized cases or pallets stored in reserve storage. The green totes are then transported on a separate conveyor belt directly into the DPS inventory system.

2.3 Returns

When customers place an order, they have the option to return items at no cost. These returns are collected at the regional transfer points and brought back to the DC by truck. This truck always carries a load: on the way out, it delivers customer orders; on the way back, it transports returns, including empty green totes, grey containers, pallets, and items customers wish to return. Returns may occur if the customer is not home, does not want the product they ordered, or is simply dissatisfied with the item. In such cases, all or part of the order will be returned.

At the DC, the returns area in hall 2 is where all returns are processed. Operators in this area receive returns from all trucks coming from various transfer points. Depending on the type of return, items are redirected to different areas. Conveyable SKUs are replenished into DPS by being placed on the conveyor belt. Non-conveyable SKUs are temporarily stored in grey containers, which are the largest bins used for storing and moving the biggest items, and later moved to the alternative storage area in hall 1. The returns department ensures that all returned items from the trucks are reintegrated into the inventory. Unlike the high-pressure timelines of the order fulfillment departments, the Arrivals and Returns department operates with more flexibility, allowing them to manage daily fluctuations without the urgent need to complete all tasks by the end of the day.

2.4 Production

The Production department is a multifaceted entity within the DC, and its activities span across three halls, being responsible for the reserve storage in hall 3, the alternative storage in hall 1, and the sorting belt of which the activities take place in hall 2. Each will be elaborated on step by step below.

The reserve storage area in hall 3 of the DC oversees two vital subareas for storing new inventory from SIP or MP stations, which is later repacked for DPS. This area is crucial not only for order replenishment but also for order picking for non-DPS managed orders. The workflow and interactions within this area are depicted in Figure 6. Pallets arriving from the workstations of hall 2 wait in the pallet queue. These pallets follow one of two paths: repacking or storage. Pallets destined for repacking are meant to replenish DPS inventory and either get delivered to a repack workstation or temporarily placed in a drop zone if the workstation is full. Items are repacked into green totes at these stations, then conveyed to DPS. If no repack is needed, pallets are moved to reserve storage by being moved to a pick and deposit (P&D) station, which are handled by one of seven Turret Stock Pickers (TSP) for storage in one of the 26 pallet bays.

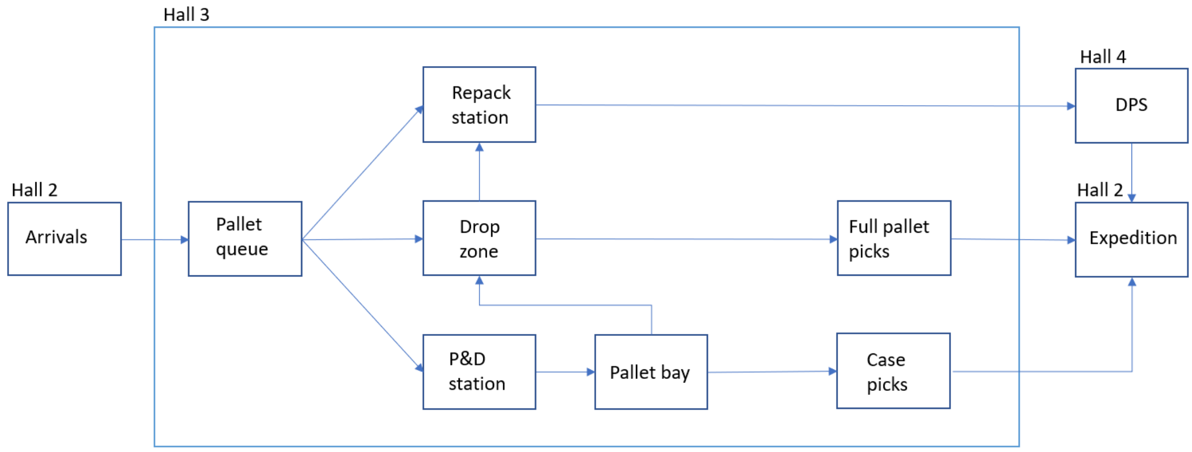

Figure 6: Workflow at the Production department

The movement of incoming pallets is done by a transporter who operates a reach truck. In total there are three reach trucks available that can perform this task. Generally, the transporter handles all traffic from and towards the repack station, P&D station, and drop zone. Next to the transport, a reserve storage handler is tasked with storing and retrieving full pallets using the TSP truck. The reserve storage's pallet bays are organized such that a TSP truck operates in the 14 even-numbered aisles, designed for straight-line movement and turning into the next aisle for swift storage, retrieval, and dropping of pallets.

Pallets in reserve storage are eventually retrieved for orders through repack, case-picking, or cross-docking. For repack, TSPs retrieve pallets and move them to a repack area drop zone, where reach trucks place them at available repack stations. Secondly, case-picking allows for direct retrieval of large quantity orders from reserve storage using special cranes, bypassing the need for repacking. Cranes, operating in even-numbered aisles, pick the required number of cases from pallets, placing them in grey boxes for sorting and dispatch to the Expedition area. Grey boxes are medium-sized bins used for storing and moving a larger set of items at once. Lastly, cross-docking handles orders requiring entire pallets, retrieved by TSPs and transported by reach trucks to full pallet pick spots. Large orders exceeding pallet quantities may involve multiple full pallet retrievals, with any remaining quantities fulfilled through case-picking or DPS.

The alternative storage area, located in hall 1, is similar to the reserve storage in that it contains inventory for SKUs that customers can order. The big difference is that items in alternative storage do not end up in DPS because the products may be too large or too heavy, also denoted as non-conveyable SKUs. All items that are larger than a specific dimension or heavier than a set weight are stored in this area, which has 18 pallet bays. Such items cannot be placed on the conveyer belt and must be stored and retrieved manually. Here, a GPC truck is used, of which there are six available. The same applies to hazardous products, which are stored in a bunker within the alternative storage area, isolated from the rest of the DC.

The order picker navigates the aisles with a flexible picking car equipped with two large grey containers, each designated for a specific destination. There are seven picking cars in use. The full list of orders for a destination is referred to as a picking task, and each task is compiled to fill the grey containers to 90% capacity, or an expected load carrier of 0.9. Tasks are released when a load carrier of 1.8 is reached, allowing two containers to be picked simultaneously, thus minimizing inefficient space usage and the frequency of aisle traversal. This strategy requires awaiting sufficient order arrivals, often until the 20:00 order deadline. The grey containers are then dropped at the Expedition area, highlighting interdepartmental dependency. Releasing tasks as late as possible ensures most incoming orders are included and picking tasks are optimized, but releasing too late can delay the production shift and extend the Expedition department's end time. Conversely, releasing too early results in minimal container usage and way more frequent picking runs. For that reason, picking tasks are rarely released in the early time intervals.

Lastly, Production is responsible for sorting, which happens in the Expedition area in hall 2. The grey boxes on the sorting belt that have arrived from alternative storage or reserve storage have been sorted automatically by the sorting lane based on destination. However, there are 40 destinations, divided over 16 lanes. Therefore, the contents of the grey boxes must be sorted a second time. At each lane, two to four larger grey containers are prepared to receive the contents of the boxes, each representing a different destination. The first lane, for instance, may have grey containers for destinations such as ‘Eindhoven’ or ‘Zwolle’. The process of unloading the grey boxes into the appropriate destination container is performed manually. At all times, 16 sorters can be present simultaneously, but in practice a maximum of three is common.

2.5 DPS

The DPS area is located in hall 4 and is the most sophisticated area in the DC. The majority of all employees work at this department and they also perform the bulk of the order picking process. This is made possible by the DPS (Dynamic Picking System) which is an automated storage and retrieval system (AS/RS). Such a system makes it possible to transport a very high volume of storage effectively and accurately. The product density in the DPS area is therefore also very high, containing thousands of green totes each containing tens to hundreds of SKUs. The DPS area consists 16 aisles, each containing two important factions, namely a storage and picking faction. An overview of the DPS workflow is given in Figure 7.

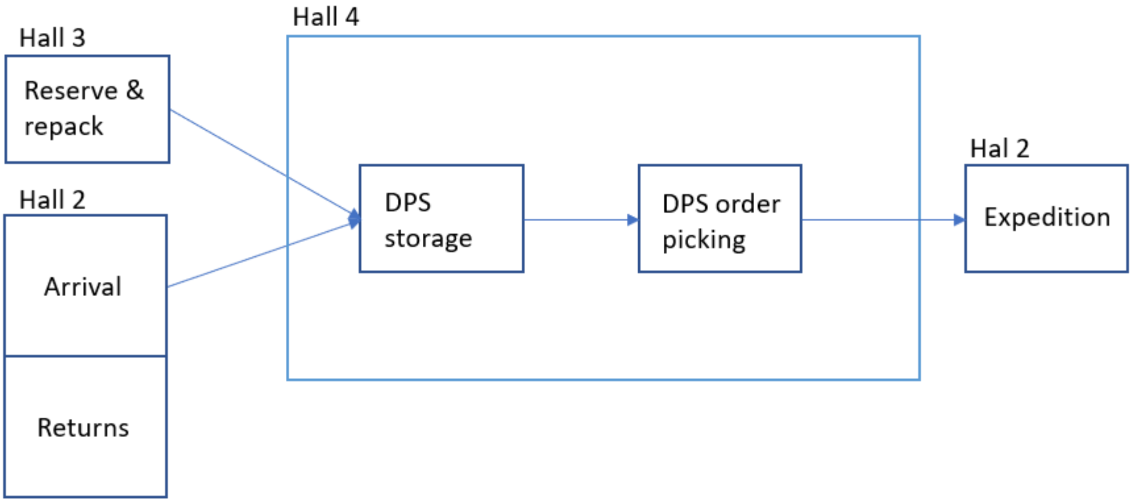

Figure 7: Workflow at the DPS department

There are three possible ways to replenish the DPS which is through the repack station, MP station, or Returns. In any case, the SKUs will have been placed in green totes that will arrive at the DPS through the conveyor belt. Here, the AS/RS will intelligently store the green tote in one of the empty locations. For the storage, an elevator and mechanic arm store the tote in the designated location. As every green tote is supplied with a barcode, the DPS will know exactly which SKUs are stored in what green tote. To understand which tote must be placed where, the DPS system employs an algorithm that optimizes the storage. The system also knows which totes must be relocated at what time such that an optimal order picking process can be achieved.

This leads to the second area of the DPS area. Next to the storage area, there is the picking area where orders can be picked directly from the green totes. These order pickers are assigned to a work station. The DPS system will know which item they will have to pick and will therefore make sure that green totes that contain that item will be relocated to the aisle that is closest to the work station of that employee. This minimizes walking and increases order picking efficiency. The order picker will then pick the necessary items and the correct quantity, and place the picked order in an empty green tote that will have arrived through the conveyor belt. The worker proceeds with a different product from the same customer order until completion, after which the green tote moves along the conveyor belt to the next picking location, where a worker repeats the same process until all orders for the customer have been picked. If no more orders are required, the green tote moves from the conveyor belt to the Expedition area, where it will be processed for outbound.

The core of the DPS department is the DPS itself which due to its mechanized AS/RS is the most automated part, but it is also very scalable and therefore the largest out of all departments. At the same time, the complexity and scale of operations make it heavily prone to errors. Spanning two floors, the DPS comprises 16 picking aisles, each with 5 workstations and a clearing station, totaling 80 workstations and 16 clearing stations. While this setup could theoretically accommodate 96 workers simultaneously, several factors limit this number. Firstly, while the volume of work in certain intervals justifies full staffing, it is not practical due to the speed of the AR/AS, which can handle a high pace but only up to a certain level. Secondly, a higher number of order pickers working simultaneously increases the risk of system malfunctions, further decreasing productive output.

Efficiency is not just about the number of active workstations but also their distribution. Ideally, only three workstations per aisle should be active at any time to prevent bottlenecks and maintain smooth operations. Additionally, each aisle has a key user at the clearing station, tasked with resolving issues flagged by the system such as weight discrepancies or irregular shapes, which indicates potential errors in order picking. Key users often work on different aisles simultaneously, as the percentage of picked orders that need controlling is relatively low, enabling them to move around aisles where they can clear the flagged issues.

A key characteristic of the DPS department is that they align their operations with the departure schedules of trucks from the Expedition area, operating in batches related to these times. For example, if a set of trucks has a departure time of 21:00, then the orders that should be part of those trucks must be picked before 20:00. What this means is that the orders that fall within a certain batch will have priority in the order picking. The AS/RS will handle this automatically and the operators at the work station will not notice anything from this priority setting.

The way the DPS aisles are structured is reminiscent of a loop, where the conveyor belt traverses through all the aisles, and the output ending up at Expedition. Figure 8 displays the integral parts of the DPS. Each aisle is equipped with two conveyor belts: a main belt that interconnects all aisles and an aisle-specific belt for intra-aisle transport. This dual-belt system allows for a green tote to be selectively routed within an aisle for picking at specific workstations before being reintegrated into the main flow towards Expedition or moved to the next aisle.

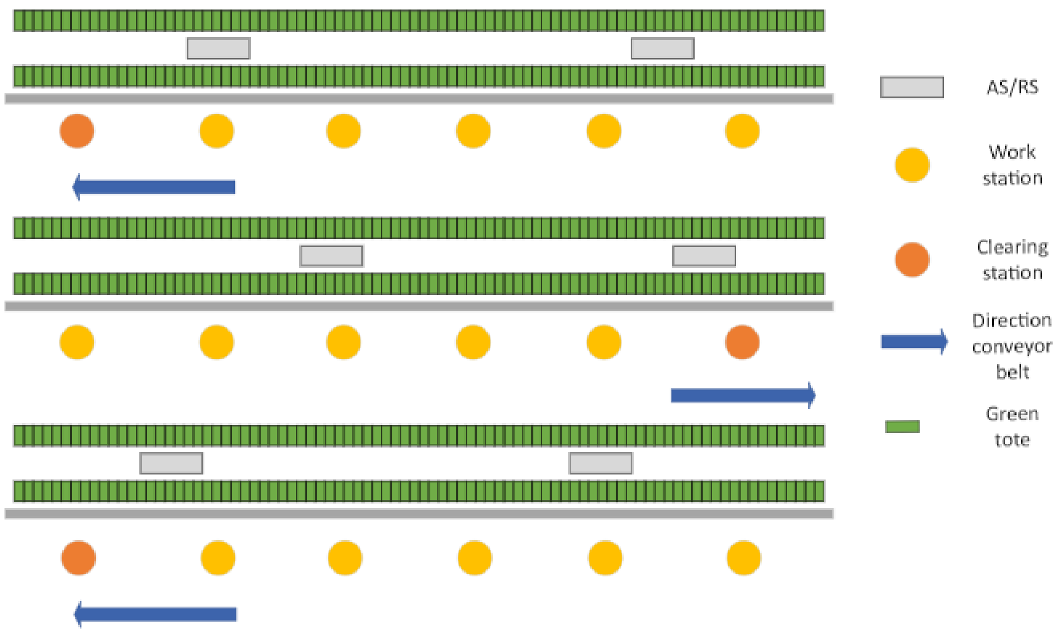

Figure 8: Overview of DPS aisles

The two-belt setup also offers an advantage in mitigating the impact of malfunctions. If an aisle-specific belt encounters an issue, it restricts the affected workstation without compromising the entire aisle's productivity. Conversely, a malfunction in the main belt poses a significant challenge, potentially halting operations across multiple aisles and necessitating the relocation of workers to maintain productivity. Evening shifts, which typically experience higher activity levels, are especially vulnerable to disruptions from main belt malfunctions since other aisles are often already occupied, leading to lost productivity with the impeded workers. Additionally, as the AS/RS cannot enter this aisle, it will store inventory elsewhere, which would be a suboptimal allocation. It is therefore crucial to minimize malfunctions and keep the work flowing.

2.6 Cable

Cable is a unique department within DCA, solely focused on one product. However, this product comes in a wide variety of versions and sizes, often requiring custom craftsmanship. The cable section is located in hall 3 of DCA. Many of the cables are stored outside in racks, where reels are kept. Four reach truck drivers operate in this area, preparing and transporting the reels to the cutting machines. The reach truck driver retrieves the reel from the rack and places it on a roller conveyor, which automatically delivers it to the cutting machine. However, this conveyor only serves four machines. For the other machines, the reach truck driver positions the reels manually and collects them once the cutting is completed.

Within the department, various cutting machines are used to tailor the cables. Different types of machines are designated for each cable size, with the primary distinction being between ring machines and reel machines. Ring machines unwind the cable from the reel and coil it into a ring for delivery to the customer. Reel machines, on the other hand, transfer the cable from one reel to another, cutting it to the required length for the customer. The latter method is used for heavier cables that are impractical to coil. Additionally, cable cutting is conducted directly from the racks. Here, the operator maneuvers a flexible machine through the cable aisles to cut the cables to size. Once a reel in the rack is empty, a new one is placed. This method is mainly used for thin cables. After cutting, the cable is placed in a full grey container, which is sorted according to the destination routes. When the container is full, it is taken to quality control within the department. A final inspection is performed, and the cable is sorted for dispatch. This process applies only to cable rings. Reels follow a similar process but are stacked on a pallet and brought to quality control. Eventually, all grey containers and pallets are collected by Expedition.

2.7 Expedition

The Expedition department, situated in hall 2 directly across the arrival area, plays an integral role in managing outbound logistics, as it oversees the Order Consolidation Buffer (OCB) and dock areas. This area serves as a consolidation point for orders originating from all previously discussed areas within the DC. Here, orders are grouped, readied for shipment to the designated regional transfer point, and finally dispatched via the assigned truck. This final stage is depicted in Figure 9.

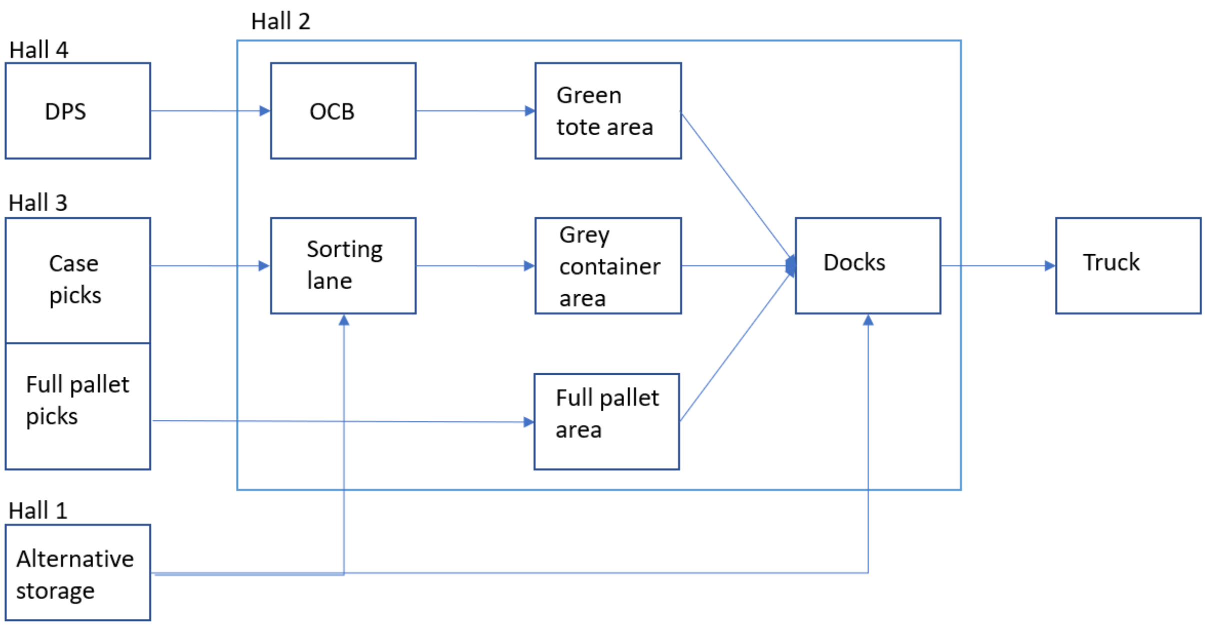

Figure 9: Workflow at the Expedition department

The primary inflow for outbound shipments originates from the DPS, where order pickers place selected items into green totes that are conveyed from hall 4 to the order consolidation buffer (OCB) in hall 2. Here, totes are stored until a complete set of twelve is assembled, then packaged onto a dolly and sent to a designated area for further Expedition processing. The second inflow arrives in grey boxes via a sorting belt, containing full cases or non-conveyable SKUs from reserve storage. These boxes are sorted directly into lanes by destination, emptied into larger containers, and sent back for reuse. The grey boxes from Cable have already been sorted. The final inflow consists of full pallets from cross-docking, placed in the full pallet picks area in hall 3.

In total, 40 different destinations are served each day, divided over 17 docks. Outbound logistics involves several key responsibilities. Firstly, one group of workers handles the output of the OCB in hall 1, where finished dollies are organized in a designated buffer area. This area acts as an intermediate stage between the OCB and the docks in hall 2, allowing operators to better manage the packaging of green totes. Specifically, the operators wait for a second package of the same destination such that those dollies can be doubled up for transport, which is done by a low truck. The doubled up package is subsequently stored in its designated area, until it gets picked up for outbound. At most, six employees are working at OCB. This is due to the output of the three OCB stackers, which is constrained to a limited dollies per hour. It follows that using more employees at OCB is redundant.

The second group of operators, known as internal transport employees, use small, agile vehicles to quickly transport dollies from the OCB to the designated docks. This division of labor minimizes long-distance travel for OCB operators and reduces internal traffic. Internal transport also collects loads from other DC areas, including grey containers from the sorting belt in hall 2 and full pallet picks from reserve storage in hall 3. They constantly circulate the DC to gather loads.

Lastly, dock workers manage both inbound and outbound logistics, handling the loading and unloading of trucks. They sort inbound items such as empty pallets, grey containers, green totes, and returns, while ensuring the timely dispatch of outbound loads. The dock area is divided into six subareas, each managed by an employee who oversees loading operations for specific destinations. Essentially, all Expedition employees form a relay system, moving items from the OCB to internal transport, and from internal transport to the docks.

The Expedition department, marking the final stage in the order fulfillment process, is inherently linked to the throughput of preceding departments. Its operations, while straightforward, are heavily reliant on the timely completion of tasks by other areas of the DC. Delays in picking processes, often caused by unexpectedly high demand, insufficient workforce, or malfunctions, directly impact Expedition and can lead to extended working hours and potential complications with truck departure schedules. These delays could again affect the truck drivers' ability to reach the regional transfer points on time, causing a propagation of negative effects. Ultimately, adequate levels of throughput at the Production, DPS, and Cable departments would also benefit the Expedition department.

2.8 DCE

The final department that belongs to DCA is DCE which carries a unique role within the whole operation. DCE is actually a standalone DC around the corner of DCA. The reason is that because of ground capacity constraints, space had to be freed within DCA and lots of items had to be moved. As a result, DCE has been created in order to store items with peculiar characteristics that do not fit within DCA. DCE also completes specific, tailor made tasks from special clients that cannot be done in the main DC. Additionally, DCE has its own arrival, returns, order picking, and Expedition operations, but much smaller in size. As DCE falls under the reporting umbrella of DCA, it has been included in the analysis. However, since its operations do not affect or depend on the other departments in any way, it will be left outside of the research scope.

2.9 Planning and employee types

To calculate the required resources for a given day, i.e. the desired workforce size, each order fulfilling department looks at their expected workload. The workload is defined by the number of orders for a given day and a department specific productivity norm, which is the standard number of picks per hour per employee. The required number of employee hours on a given day is then the expected workload divided by the norm. This way, the workforce size is set for those departments. The final department, Expedition, depends on the output of Production, DPS and Cable. However, since they are able to handle large volumes of orders in bulk (the orders will have been stored in large packages of green totes or grey containers when it reaches Expedition), the workforce size at Expedition can remain constant, whereas the end time of the shift is more of a concern.

Each department operates on different shift schedules, most commonly morning, afternoon, and evening. The order picking departments work throughout the day, with a particular emphasis on the afternoon and evening shifts. Similarly, Expedition workers are active all day, with a focus on the evening shifts. Each shift requires a specific number of employees, as calculated based on the expected workload. The planning lead of each department is responsible for filling these requirements. Initially, regular full-time employees are allocated to the shifts. However, due to capacity constraints, the departments often cannot meet demand with regular employees alone. This necessitates the hiring of subcontractors, who are flexible workers that can be allocated to unfilled shifts.

The cost structure and rules for subcontractors differ from those for regular employees. For instance, once a subcontractor is hired for a shift, they must be utilized for a minimum number of periods, making it crucial to balance immediate needs with potential future overstaffing. Regular employees typically work strict shifts and leave once their shift ends, whereas subcontractors can extend their shifts more easily and are often needed past midnight, where overtime is common. Thus, planning is a complex task that must integrate expected workload with capacity flexibility and specific requirements. Table 3 provides an overview of all jobs related to order fulfillment in each department, detailing capacity, average productivity norms, and possible shifts.

Table 3: Overview of resource utilization of the order fulfillment departments

Department	Component	Max capacity	Norm per hour	Shifts
Production	Reserve storage	7 TSP trucks and 3 Reach trucks 13 cranes	50 per hour	Three shifts
	Alternative storage	6 GCP trucks	60 per hour	Three shifts
	Sorting	16 belts	120 per hour	Only afternoon and evening shift
DPS	Order pickers	80 work stations	108 per hour	Three shifts, both 8-hour and 5.5-hour
	Key-user DPS	16 clearing stations	Not applicable	
Cable	Cable retrieval	3 machines	Not applicable	Three shifts
	Cable cutting	8 work stations	41 per hour	
	Forklift truck	6 trucks	Not applicable	
Expedition	OCB	3 low trucks	Not applicable	Three shifts, with evening shift extending until after midnight
	Loading	No more than 4		
	Unloading	No more than 4		
	Internal transport	3 cars		

2.10 Customer Order Management

The company's order fulfillment process relies heavily on automation and a series of interconnected information systems to manage the customer orders. Initially, a customer order is accepted by SAP, which converts it into a sales order (SO) containing details about the ordered items and their quantities. This SO is then processed by Manhattan Active Omni (MAO), which selects the most suitable DC for fulfillment and may result in splitting the SO between multiple DCs based on item and stock availability. Following this, MAO assesses if the chosen DC can meet the order by comparing the requested items' quantities against the inventory, eventually leading to a distribution order line (DOL). This DOL serves as an assignment for the DC to pick the order.

The exact allocation of the assignment is further handled by the warehouse management system (WMS) of the DC. This WMS receives all DOLs and automatically determines where the required items can be found and in which quantities, thus creating order picking tasks for respective departments. A DOL could, for instance, consist of multiple products that can be entirely picked in

DPS. However, it could also occur that a share of the DOL must be picked from alternative storage and another share from a different location. The WMS is the system that assigns the contents of the DOL's to the respective departments. Since one DOL could create order picking tasks in multiple departments, each order pick will get a different label for scanning and navigating the transport. This label is called an oLPN and is the unit used for measuring and processing workload at a department. The number of DOLs, the distribution across the departments, and the workforce that has been used to fulfill the orders will be analyzed and discussed in the next chapter.

Chapter 3 – Data analysis

3.1 Datasets

Workload

A dataset has been acquired of the demand in 2023, consisting of over 18 million entries. This dataset contains all the order lines of the year for DCA, accompanied with its precise date and time. An example of the columns and rows is shown in Table 4. The list begins with a purchasing order, which is tied to one or more DOLs, depending on the number of different items a customer has ordered. The dataset also shows the product group and the quantity that has been ordered.

Table 4: Example entries of the demand dataset

Purchasing Order	Customer Number	DOL	DOL Date	...	SHIP TO	ITEM	QUANTITY	Product Group
7529194	39-22192	752912110154642	2023-01-01 03:12:00	...	EIND1	752996	20	DPS
2821020	32-05044	282102110154714	2023-04-17 20:44:00	...	AMS3	282102	1	Bunker
1877295	21-55323	187722110154787	2023-11-28 16:59:00	...	UTR2	187727	450	KV Battery

Multiple preprocessing steps have been conducted to prepare the dataset for analysis. Firstly, the data has been sorted based on the DOL date. Subsequently, columns have been created to denote the weeknumber, weekday, and time interval of the DOL date. This step allows for the translation of demand into workload. Demand can namely occur at every moment of the day, but workload is defined differently. Workload for a given day is all DOLs from 20:00 the day prior till 20:00 the current day. Moreover, since the weekends are no working days, the DOLs made from Friday 20:00 onwards are carried over to the Monday as well. More precisely, the DOLs received past the order deadline of 20:00 will be carried to the 4:00 time window of the first upcoming working day, which is the first time MAO releases the DOLs to WMS.

Based on this information, a specific workload date and time interval is coupled to each DOL. This makes it possible to analyze the workload in different aggregate levels, namely a week, weekday, and time interval. A second means of aggregation is the type of product, which can be defined by the product group or by the corresponding department. Therefore, analysis can be performed on a time dimension and department dimension, and grouping can take place on several aggregate levels for more granular perspectives.

Workforce

A second dataset has been acquired that has stored an anonymized log of all employees' activities at each working day of 2023 of the whole company. For each working day, all activities of all employees are logged. A log consists of an employee ID, the task performed within a shift, the shift time, and the corresponding department. It is common for employees to perform multiple tasks within the same department during a shift. Both tasks are logged independently but fall under the same shift. For each employee, it is also registered whether it is a regular employee or subcontractor. A summary example of this dataset is provided in Table 5.

Table 5: Example entries of the workforce dataset

Employee ID	Employee type	Date	Shift duration	Task duration	...	Code	Task	Shift time	Department
129	Regular	25-09-2023	8	8	...	7150113	Production DPS	07:00 – 15:30	DPS
188	Subcontractor	25-09-2023	5.5	2	...	7150114	Key-user DPS	07:00 - 12:30	DPS
188	Subcontractor	25-09-2023	-	3.5	...	7150113	Production DPS	07:00 - 12:30	DPS

In total, the dataset initially contains over 1 million entries of the employee logs from all 41 selling locations of the company. After filtering for DCA and retaining only the relevant departments, the dataset was reduced to 80,000 entries. Further preprocessing steps mirror those of the workload dataset. With the workload, it was imperative to know how many DOLs had to be handled in a certain time interval. Similarly, with the workforce dataset it must be known how many employees have been active within a certain time interval. This can simply be deducted from their task duration and shift start time, taking into account breaks.

As a result, 24 columns have been added, denoting the time intervals from 4am in the morning when the first employees have logged a task until 3am at night when the last employees finish their tasks in case of overtime. For each task that has been logged, the corresponding time interval columns are filled with the (decimal part of an) hour an employee has spent on this task in this time interval. In the example above, employee 129 would have one full hour in the column denoting the time interval from 7am till 8am, and half an hour from the time interval starting at 15:00. The total hours used in a time interval, i.e. the workforce size in a certain time interval, is then the sum of the entire column. Much like the workload dataset, different groupings and aggregate levels are possible.

<u>Pick rate and malfunctions</u>

The final two acquired datasets relate to the pick rate and to the malfunctions at the largest order fulfillment department, DPS. The reason for this inclusion is that at DPS, the pick rate is greatly influenced by workload and workforce, whereas the output at other departments is more straightforward. The first dataset displays the performance of the workforce in 5-minute intervals with information such as the number of DOLs released, the number of DOLs completed, and the pick rate. The second dataset shows all the malfunctions that have been registered throughout the day, including the start time, duration, location, and identity. By merging the datasets, valuable insights can be obtained about the relationship between picking activity, malfunctions, and the overall impact on performance. While the available data only spanned two months, the logs were sufficient for valuable analysis, ultimately containing over 14,000 entries. With the preprocessing completed, all datasets mentioned are now ready for analysis.

3.2 Workload

In 2023, there were over 18.4 million DOLs, originating from 5.3 million purchasing orders. This demand was driven by approximately 55k unique customers. The demand of the review period has been plotted per day for a regular month and a summer period, as illustrated in Figure 10 and 11. In both graphs, there is an immediate weekday pattern noticeable, with minor fluctuations during working days and major decreases in the weekends. This lack of demand in the weekend is explained by the customer segment of the company which are exclusively businesses who are working on projects, hence on regular weekdays. Looking further at the graphs, demand appears to be relatively stable throughout the year. Only in the holidays season, especially in the summer, the demand is significantly reduced.

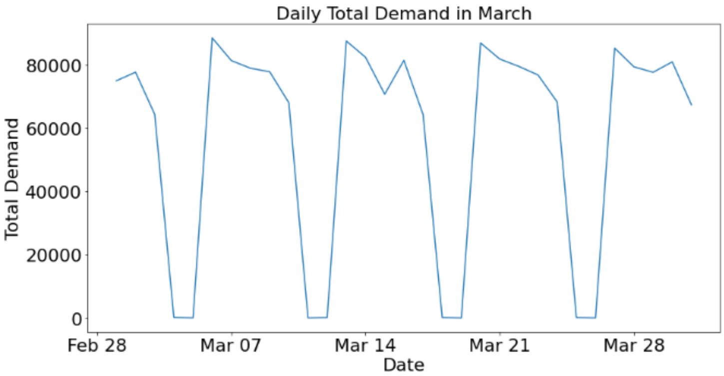

Figure 11: Daily demand pattern in a regular month

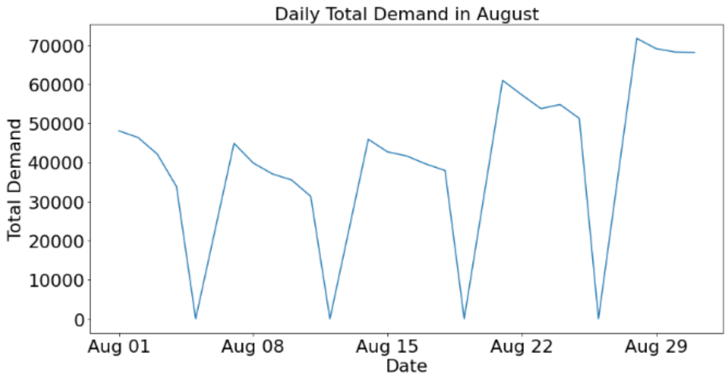

Figure 10: Daily demand pattern in the summer

The demand has been converted into workload for each working day of the week, taking into account the time windows as discussed during the preprocessing. The total workload per week has been plotted with and without outliers as shown in Figures 12 and 13. Figure 13 reveals a significant fluctuation in total weekly workload, with notable dips suggesting either periods of decreased demand, or business closures for holidays or other events. Figure 12 illustrates the same weekly workload when the extreme outliers have been removed. This results in a clearer view of a somewhat more stable workload, albeit fluctuations still being present. Further analysis has been conducted without the outliers to ensure most accurate representation of a regular demand week.

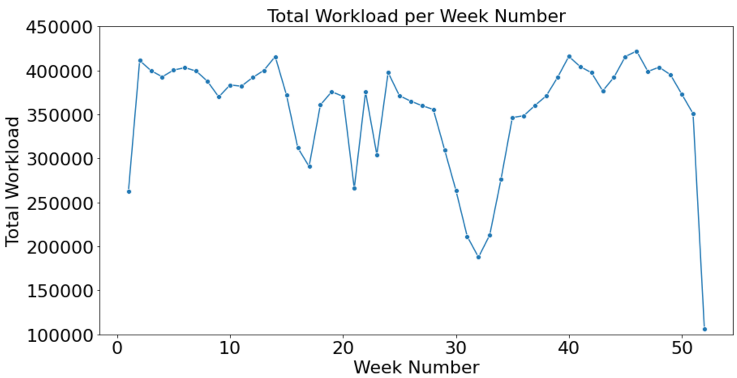

Figure 13: Total workload per week number

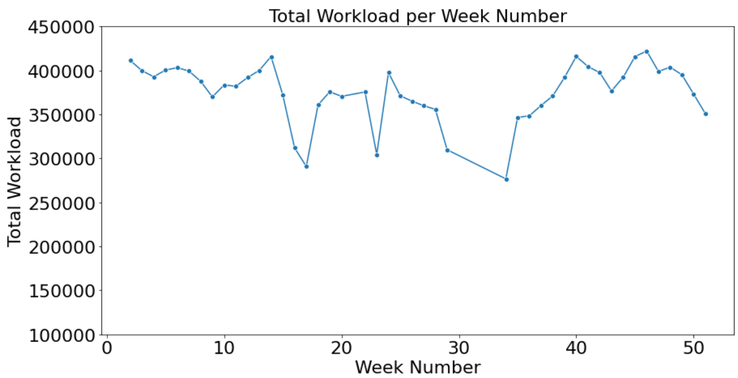

Figure 12: Total workload per week number, outliers removed

Upon examining the weekly data, as illustrated in Figures 14 and 15, a clear pattern emerges in the distribution of workload across weekdays. The majority of the workload occurs on Monday, which is expected due to the accumulation of orders placed over the weekend after Friday 8pm. Workloads on Tuesday, Wednesday, and Thursday are similar, showing a gradual decrease throughout the week. All three of these days have significantly lower workloads compared to Monday. Fridays experience the lowest workload of all weekdays, likely because customers do not feel the urgency to place orders before the weekend. Overall, the average daily workload decreases as the week progresses, with a relatively constant workload on each specific weekday over time.

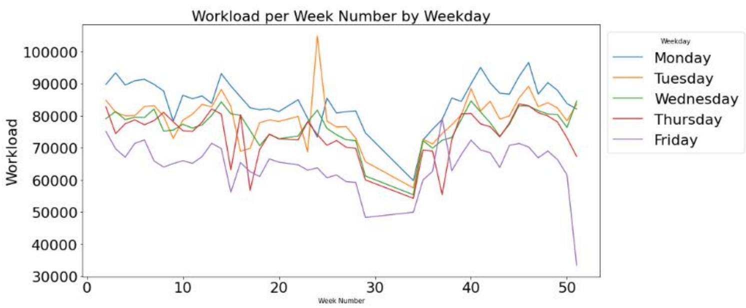

Figure 15: Workload per week number per weekday

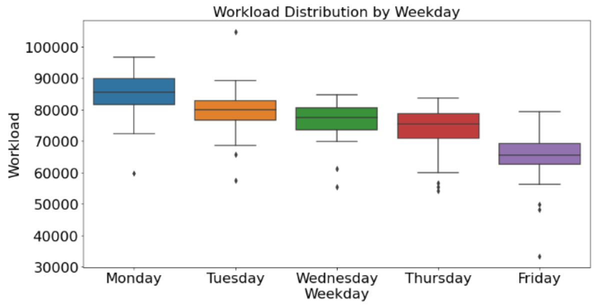

Figure 14: Workload distribution per weekday

<u>Workload per time interval</u>

Upon closer examination of the time intervals, a noticeable trend in workload emerges. Each time interval is defined as an hourly period starting from a specific time. For example, time interval 14 refers to the hour from 14:00 to 15:00. Figure 16 illustrates the average percentage workload per time interval, with the highest workload occurring at the first interval at 4am. This peak includes orders placed within this interval and all orders placed the previous day after 8pm, explaining the high workload. Due to the maximum limit of 5,000 DOLs per wave replenishment, MAO cannot send all DOLs to WMS at once, causing some DOLs to spill over to the next interval, resulting in a relatively high number at 5am. The 6am interval is the first 'regular' time interval, consisting solely of real-time DOLs.

Figure 16 shows the percentage of daily DOLs ready for picking per time interval, not the actual picking activity. For instance, approximately 14% of daily DOLs become available at 4am, but picking begins after 7am when employees start their shift. There is no rush to address the 4am workload immediately, as subsequent intervals are relatively low. Up until noon, the workload is manageable. However, in the afternoon, particularly around 4pm, there is a significant increase, with more than 50% of orders arriving after 3pm.

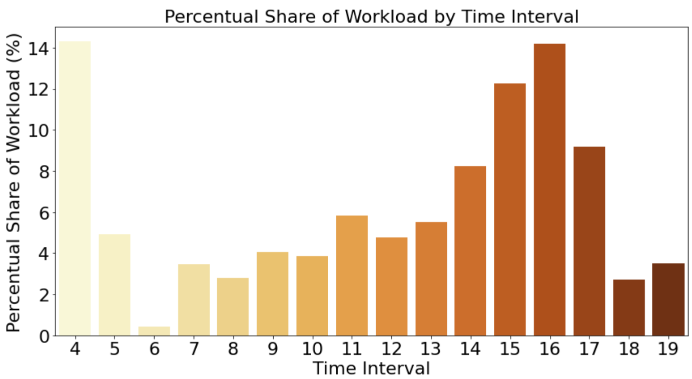

Figure 16: Percentual share of workload per time interval

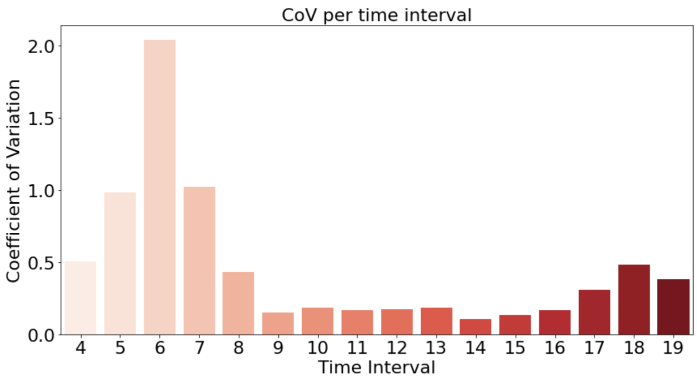

Figure 17: Coefficient of Variation per time interval

This leads to several important considerations. First, the manageable morning workload justifies a smaller workforce. However, if this workforce fails to complete all orders before the afternoon wave, the remaining workload must be tackled during already busy intervals, potentially requiring expensive overtime. Therefore, it is ideal for the afternoon shift to start with no carryover from the morning. Second, ensuring all morning orders are completed in the morning intervals requires a larger workforce for both shifts. This results in the morning shift completing their work early, leaving them with low workload towards the end of their shift. Similarly, the afternoon shift will have little to do initially, as most of the work has been completed and the early afternoon intervals are not very busy. Consequently, part of the workforce will be underutilized during certain intervals, requiring a balance between meeting productivity norms and efficiently managing existing workload.

When examining the average workloads per time interval, a notable observation is that the percentages from 8am onwards remain quite consistent. Each day, a similar workload can be expected during these intervals. Figure 17 illustrates the coefficient of variation (CoV) for each time interval, showing a low relationship between the averages and the standard deviation for most intervals, with only the final two intervals showing a low to moderate relationship. The morning intervals exhibit more variability, but this is expected and not a major concern since they fall outside working hours and are considered collectively.

Figure 18 depicts the spread of time intervals, highlighting their consistency throughout the year. Most intervals display a very small spread, with the majority of data points concentrated within a range of less than 1.5%. These differences are partially due to variations across weekdays, as the order pattern on Monday may differ from Thursday. When accounting for weekdays, the spread becomes even more concentrated.

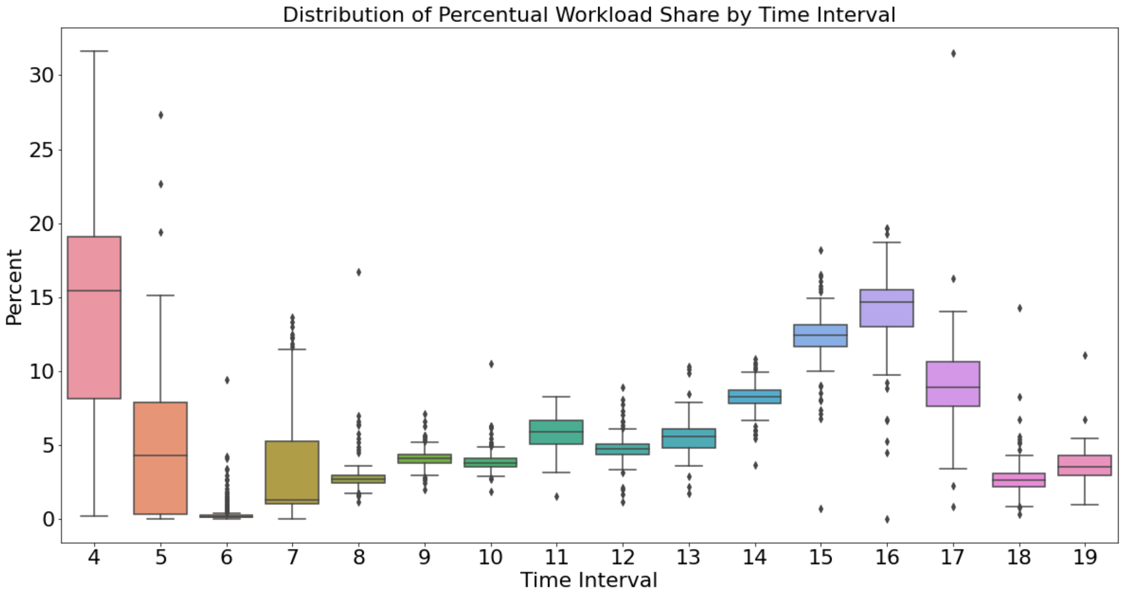

Figure 18: Boxplot of percentual workload distribution per time interval

<u>Departmental contribution</u>

Lastly, it is important to illustrate how the incoming workload is distributed across departments. Figure 19 displays the average workload per time interval, divided among the order picking departments (this figure is the absolute version of Figure 16). It is evident that the DPS department handles the majority of DOLs, accounting for nearly 90% of all daily orders. This is followed by Production, DCE, and Cable, with 5.5%, 2.2%, and 1.8% respectively. This sets the stage for the next section, which will analyze how the workload has been managed with varying workforce sizes. This section has shown how the workload arrives, highlighting the fluctuations by week, weekday, and time interval. Given that DPS is the largest department in terms of outbound orders, it logically has the largest workforce. However, many supportive processes are primarily handled by other departments to facilitate order picking. The next section will detail the workforce size distribution across departments and how each department manages the incoming workload.

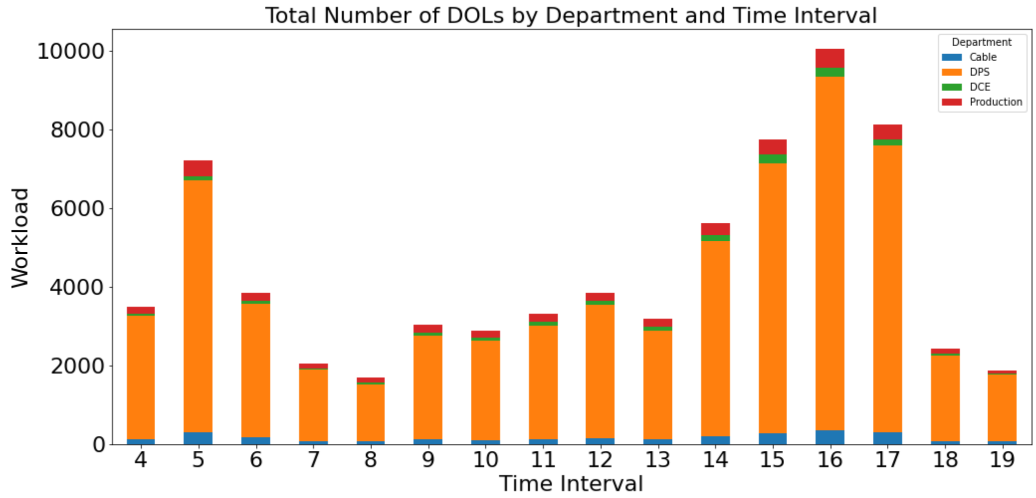

Figure 19: Total number of DOLs by department and time interval

3.3 Workforce

<u>Workforce per day and week</u>

The number of employees working per day hovers around 352 unique individuals. A similar weekday trend as with the workload is noticeable with the workforce, showing a gradual decrease from Monday to Friday. As the workload decreases, fewer employees are generally needed, as seen in Figure 20 and 21. The daily workforce also shows drops during holidays and the summer season, with significantly fewer employees working during Christmas or summer. In general, the workforce size is flexible, with the planned workforce mirroring the expected workload. The graphs also divide the employees by type. Most employees are regular staff, representing around 65% of the workforce. Although there are fewer subcontractors than regular employees, their number is still significant, reflecting the challenge of attracting human resources, as more than a third of the workforce consists of subcontractors. Interestingly, the fluctuation in workforce size for both employee types is very similar, with a standard deviation of 23 employees. This indicates that subcontractors are expectedly more flexible, but both workforces remain relatively stable over time.

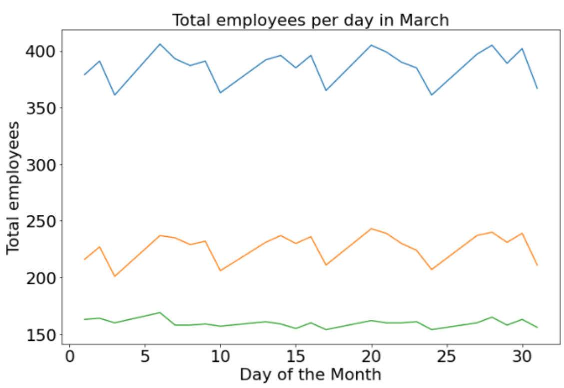

Figure 20: Daily workforce pattern in a regular month

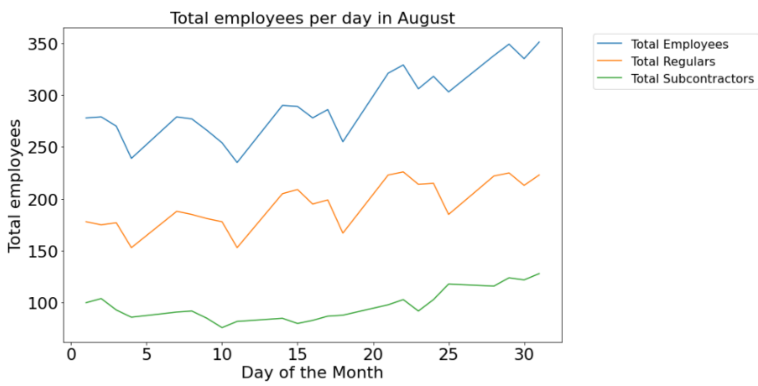

Figure 21: Daily workforce pattern in the summer

This pattern is also visible in Figure 22, which visualizes the total weekly employees for each department. The total number of unique employees allocated each week remains relatively stable, except during holiday periods. The company likely employs a mixed chase and level strategy, leaning more towards a level strategy. In this approach, they maintain a steady workforce and handle workload fluctuations primarily through overtime, with some adjustments in workforce size. Figure 22 also highlights the workforce size per department. Although DPS handles 90% of the order-picking workload, it only accounts for 50% of the workforce. This discrepancy is due to other processes, such as inbound activities, being primarily managed by other departments like Arrivals. These processes also require a significant workforce to support the order fulfillment operations.

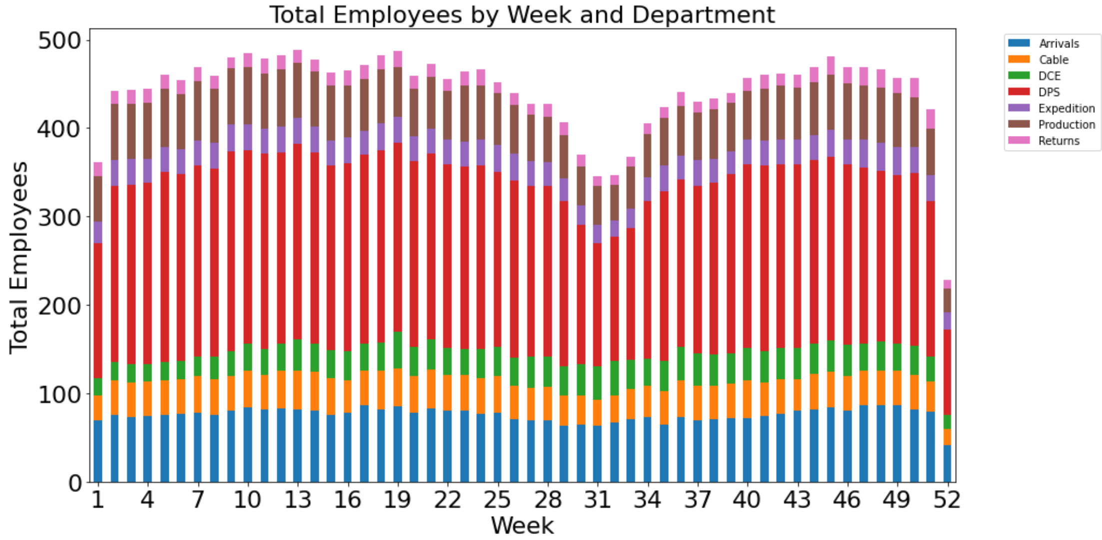

Figure 22: Total employees by week and department

<u>Workforce per time interval</u>

Nevertheless, the number of employees working on a given day or week does not provide insights into the hours worked or the efficiency of personnel utilization. For example, one employee might work a three-hour shift while another works an eight-hour shift, yet both count equally towards the daily workforce size. Therefore, it is more functional to examine the workforce size per time interval and calculate the total hours worked from there. For instance, if five people work from 10am to 11am, the workforce size for the 10am time interval is five. Each interval is an hour long. Combining all time intervals of the day provides a more detailed overview of labor input. This is demonstrated in Figure 23, which shows the average workforce size per time interval.

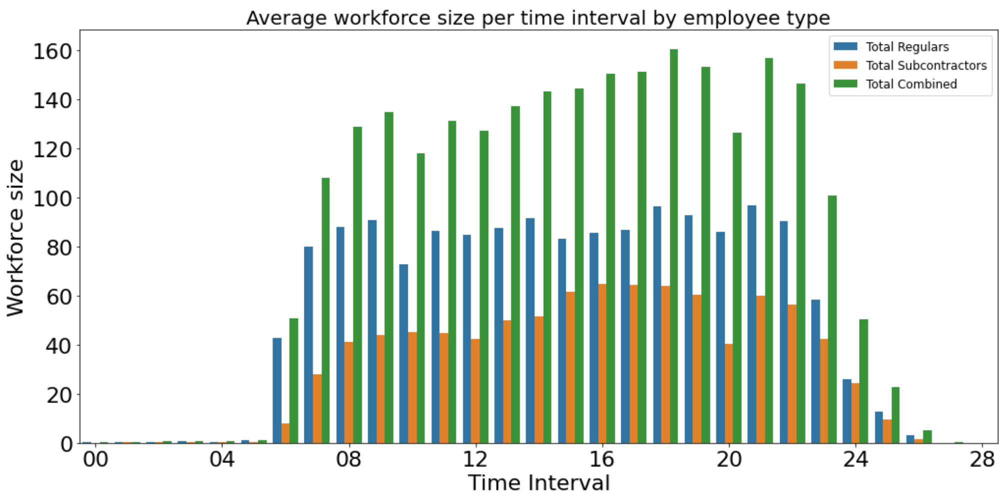

Figure 23: Average workforce size per time interval by employee type

The first employees arrive at 6am, with the number gradually increasing throughout the day. The busiest periods are in the afternoon and evening, where multiple starting and ending shifts overlap. There is a brief reduction in activity at 8pm, marking the middle of a shift when many employees take a half-hour break. At 11pm, activity drops as some evening shifts end. After this, employees are either working overtime to finish the latest order picks or are part of the Expedition department. The workforce size of regular employees remains relatively constant, roughly twice that of subcontractors. However, this ratio decreases as the day progresses, with the subcontractor workforce increasing and approaching the size of the regular workforce towards the end of the day. This suggests that subcontractors are more effectively utilized in the evening when the workload is at its peak, whereas in the earlier intervals, the need for an external workforce is lower due to less activity.

When accumulating the hours across all intervals, the average workforce size equals 2450 hours per day, equivalent to approximately 306 full-time shifts. In practice, this involves 352 employees, with a mix of 8-hour shifts and shorter 5.5-hour shifts, the latter being more common for order pickers, especially at DPS. DPS, the department with the largest daily workforce size, also translates this into the highest hourly workforce size per time interval. To better understand the allocation over departments, it is essential to recognize the possible roles within a department. Broadly, employees in a department may perform activities related to inbound or outbound processes, serve as key-users or team leaders, or be trained as new employees. Figure 24 visualizes the average workforce size in each department and the allocation of hours across these roles. This figure illustrates the average number of hours a department spends daily and how these hours are allocated among the different roles.

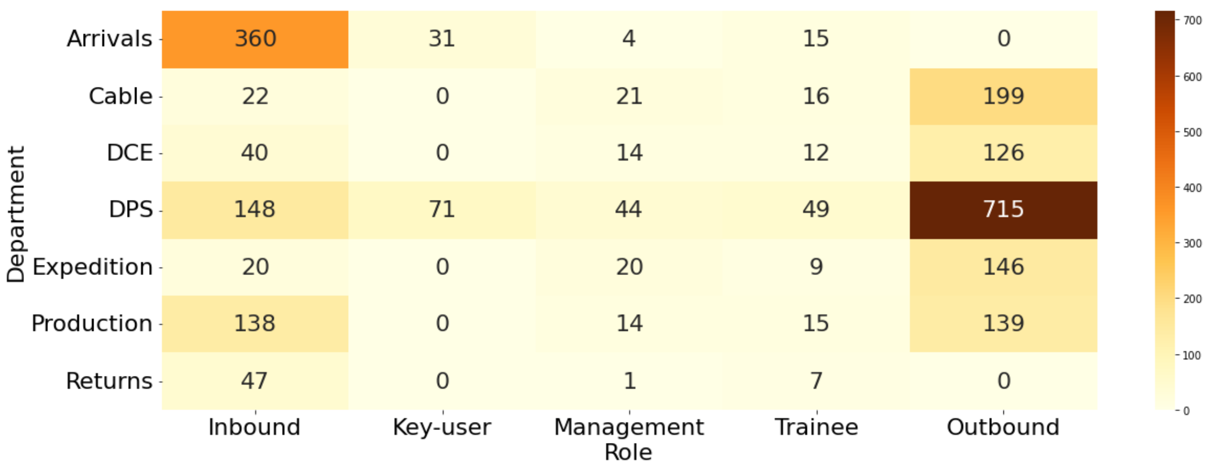

Figure 24: Distribution of workforce size per department and role

<u>Overtime</u>

As seen from the workload analysis, 40% of the workload arrives after 3pm, putting considerable pressure on the afternoon and evening workforce to complete tasks on time. Officially, most shifts cannot extend beyond midnight, with the Expedition shift ending at 1am when the last batch is loaded. However, overtime is more common than the exception. To better understand the drivers of overtime, Figure 26 illustrates the average workforce size per time interval for each role over the year. Three distinct patterns emerge from this data. First, Inbound work is highest in the morning, remains steady in the afternoon, and rarely extends past midnight. Second, the roles of Management, Trainee, and Key-Users maintain a consistent presence throughout the day. Trainee employees typically leave before midnight, but Management and Key-Users must stay until the end for oversight. Lastly, Outbound work starts at a moderate level and increases throughout the day, peaking in the evening.

This role is responsible for ensuring all orders are picked and fulfilled, and due to time pressure and potential malfunctions, it requires the most hours past midnight.

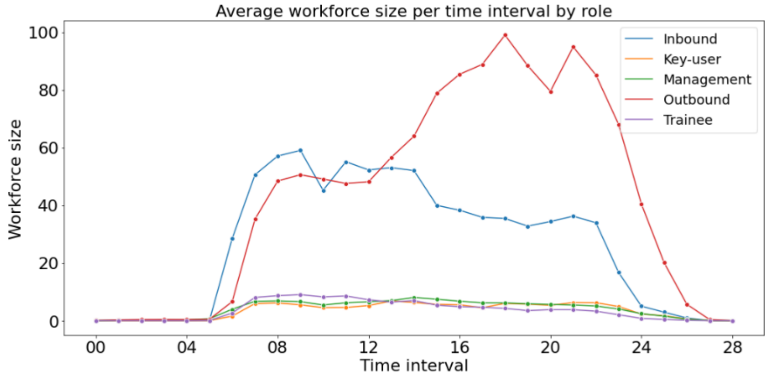

Figure 26: Average workforce size per time interval by role

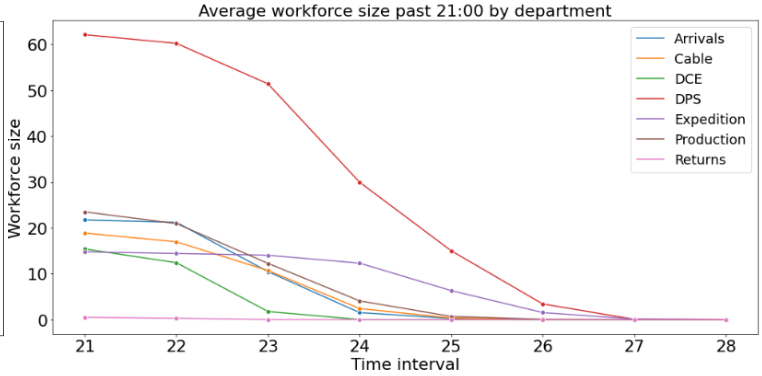

Figure 25: Average workforce size past 21:00 by department

To identify the source of these hours, the workforce active past 9pm has been plotted for each department in Figure 25. It is notable that DPS still has close to 30 people working past midnight, with most finishing between 2am and 3am on average. While the Expedition department has regular hours until 1am, they depend on other departments and, therefore, also often work until 3am as overtime. Other Outbound departments experience undesired overtime, but in smaller amounts. DCE is the only department that consistently finishes before midnight.

The above figures calculate the average, but in practice, the overtime problem is more pronounced. This is evident in Figure 27, which plots the total workforce size past midnight for each day of the review period. It can be inferred that on days with low workload, such as holidays and Fridays, there is almost no workforce past midnight. This implies that without these low-workload days, the average workforce size during those time intervals would be significantly higher, particularly on most non-Friday weekdays. It is not uncommon on these weekdays to have a workforce size equivalent to 100 combined hours spread over the midnight intervals. Considering that the Expedition's workforce until 1am equals a combined 11.7 hours, any amount above this figure can be attributed to substantial overtime.

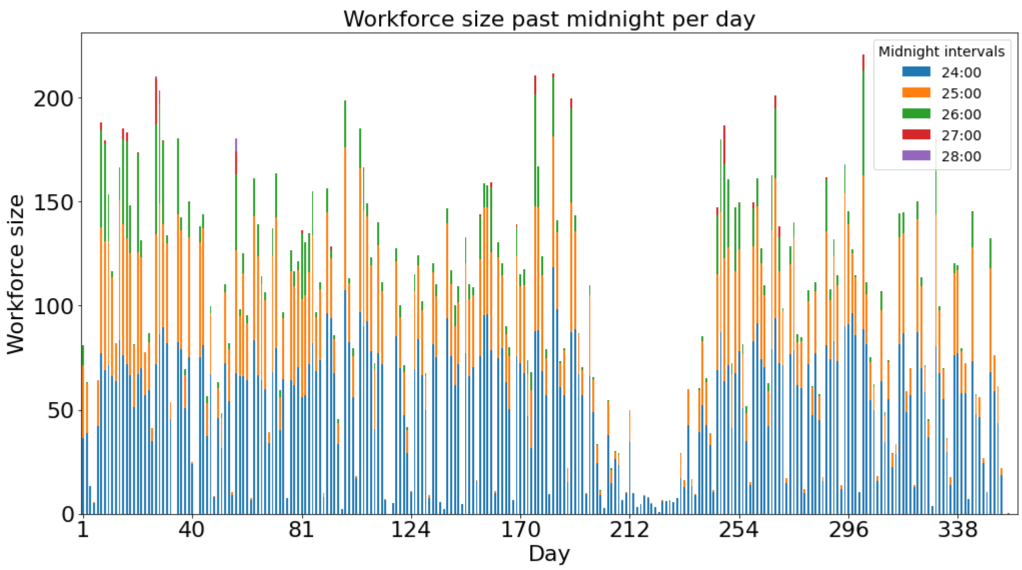

Figure 27: Workforce size past midnight per day

3.4 Pick rate and malfunctions

For the smaller order fulfillment departments, the pick rate is straightforward due to the structured nature of their order picking processes. However, for DPS, the pick rate can fluctuate significantly. To optimize workforce planning, it is integral to understand the effect of workforce size on productive output. The analysis for DPS first explored the relationship between available order lines and the pick rate, as visualized in Figure 28. This analysis indicates a clear relationship: the pick rate is low when the workload is minimal, increases after 5,000 order lines, rises again after 10,000 order lines, and then stabilizes. The lower pick rate below 10,000 available order lines is due to the AS/RS system, which has 16 aisles to distribute workload. When there is insufficient workload, certain aisles lack work, reducing the overall pick rate.

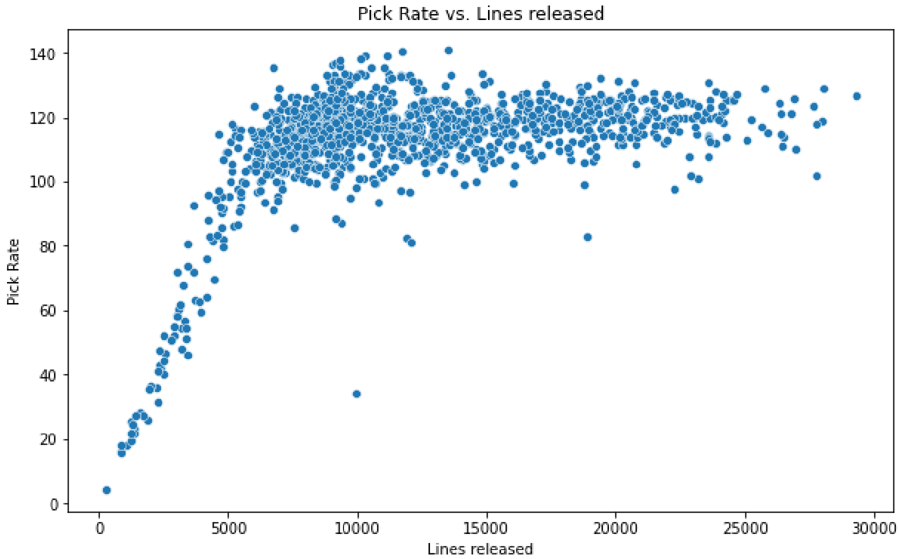

Figure 28: Pick rate vs. Order lines released

However, the pick rate for more than 10,000 order lines still shows considerable variance. To explain this difference, a regression analysis was conducted using multiple subsets of explanatory variables to predict the pick rate. The goal here is to more accurately predict how high the pick rate will be if certain variables in addition to the available order lines are known. Various regression methods were tested, and, ultimately, XGBoost was selected. XGBoost stands for Extreme Gradient Boosting and is a machine learning technique that uses an ensemble of decision trees to improve prediction accuracy. XGBoost is highly effective for regression analysis due to its ability to handle large datasets and model complex relationships between variables (Chen et al., 2015). It excels in providing high prediction accuracy by combining the predictions of multiple weak models (usually decision trees) to form a strong predictive model. Appendix A shows a ranking of all the tested methods.

To set up the XGBoost model, the dataset was first divided into a 70/30% training-test split to ensure robust model evaluation. The training set was used to fit the model, while the test set was reserved for evaluating its performance. Additionally, k-fold cross-validation (with k set to 5) was employed on the training data to mitigate overfitting and ensure the model's generalizability. Hyperparameter optimization was conducted using the Hyperopt library, which employs Bayesian optimization to find the optimal combination of hyperparameters. This process included tuning parameters such as the learning rate, maximum depth, and the number of estimators. The final model configuration was chosen based on the best cross-validated performance metrics, such as Mean Absolute Percentage Error (MAPE), R-squared (R2), and Root Mean Squared Error (RMSE)

The explanatory variables considered into the model ranged from the number of active pickers in a time interval to the type, duration, and number of errors. An excerpt of these variables is detailed in Appendix B. From this list, the variables with the highest correlations have been selected, and multiple permutations have been tested in the XGBoost model to predict the pick rate. The strongest results are summarized in Table 6, along with the MAPE, R2, and RMSE.

Table 6: Regression analyses results

	Variable(s)	**MAPE**	**R2**	**RMSE**
1	Available order lines	9.8%	0.64	13.64
2	Available order lines + active pickers	4.3%	0.83	5.12
3	Available order lines + active pickers + Malfunction identity	3.9%	0.87	4.98
4	Available order lines + active pickers + Malfunction identity + Malfunction duration	3.6%	0.88	4.93

The results show that adding an extra variable to the initial model with available order lines can explain a significantly higher portion of the variance in the pick rate, even up to 88%. For example, incorporating the number of active pickers more than halves the MAPE and RMSE, and the R2 increases from 0.64 to 0.83. This observation may be explained by the fact that with fewer pickers, employees must inefficiently traverse the entire aisle, while with more pickers, the aisles become too crowded, affecting the pick rate.

In terms of malfunctions, the explanatory power of the model is even higher. It is logical that an increased number of malfunctions would temporarily impede the systems, causing the pick rate to drop. However, the influence of malfunctions appears to be partially captured by other variables, as there may be a direct relationship between the number of pickers, the order lines available, and the number of malfunctions, indirectly affecting the pick rate.

For interpretation purposes, the model using available order lines in combination with active pickers has been chosen to aid in optimization in subsequent chapters. This combination offers a slightly below maximum predictive capability but enhances the model's interpretability. Including the malfunction variables would complicate the model, given there are over 1,000 types of malfunctions to distinguish from. The differences in MAPE and RMSE are also negligible. In the next chapter, the relationship between available order lines, active pickers, and pick rate will be used to help optimize workforce planning in DPS.

3.5 Conclusion

The first part of this chapter focused on workload patterns, highlighting weekly trends, weekday demand fluctuations, and the distribution of workload across time intervals, which proved to be of a fixed nature. It also revealed that the DPS department experiences the highest workload, greatly outweighing the other departments. The second part of the chapter examined workforce allocation. A primary issue uncovered was the prevalence of overtime hours, mainly caused by the requirements of order fulfillment in the Outbound role. DPS, again, was identified as the department with the highest workforce size and the greatest share in the overtime hours. Lastly, the final section analyzed the pick rate within the DPS department. It identified a clear relationship between workload, workforce size, and productivity rate. To optimize workforce planning, it is crucial to maintain high productivity rates.

In the subsequent chapter, the key insights from this data analysis concerning workload, workforce, and productivity rates will be used to develop an optimization model. This model will incorporate relevant parameters from the current chapter and the process analysis of Chapter 2. A holistic mathematical framework will be formulated to optimize workforce planning across all order fulfillment departments for all days of the week, aiming to enhance productivity, reduce inefficiencies, and prevent overtime.

Chapter 4 – Model design

4.1 Goals

The goal is to determine a workforce size for each time interval such that the total daily workload is met with the minimum necessary costs, subject to constraints. Figure 29 shows the hierarchical structure of the planning assignment. For each weekday, a plan must be created for four departments. Each department can be staffed by four types of employees: regular full-timers, regular part-timers, subcontracting full-timers, and subcontracting part-timers. Each employee type will be active in either the morning, afternoon, or evening shift.

Furthermore, each shift consists of a set of time intervals specific to the department and employee type, also denoted by the set $T_{d,e}^{s}$. This specific combination is called a *capacity*. For example, the regular full-timer in the morning shift at DPS may work a different set of hourly time intervals than the subcontracting part-timer in the evening shift at Production. Finally, for each time interval of each capacity, these individual employees can be assigned. The sum of these assignments determines the workforce size for a specific time interval. Since employees are unique, overlap must be prevented and working requirements must be respected, as will be detailed in the constraints part in section 4.6.

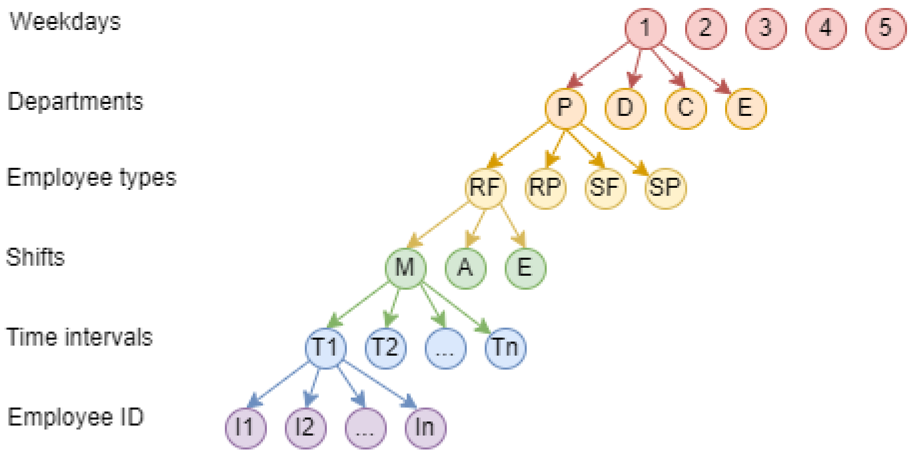

Figure 29: Hierarchical levels of the workforce planning

4.2 Sets and elements

The following section defines the various sets and elements used in the mathematical model. These sets encompass the weekdays, departments, employee types, shifts, and time intervals that form the basis of the workforce planning framework. Table 7 below lists these sets and their elements.

Table 7: Sets and elements

Name	Element	Set
Weekday	wd	$WD = \{Monday, \ldots, Friday\}$
Department	d	$\mathrm{D} = \{\mathrm{Production}, \mathrm{DPS}, \mathrm{Cable}, \mathrm{Expedition}\}$
Employee type	e	$\mathrm{E} = \{\mathrm{RegularF}, \mathrm{SubcontractorF}, \mathrm{RegularP}, \mathrm{SubcontractorP}\}$
Shift	s	$S = \{Morning, Afternoon, Evening, Overtime\}$
Time interval	t	$\mathrm{T} = \{4{:}00, 5{:}00, 6{:}00, \ldots, 25{:}00, 26{:}00, 27{:}00\}$
Time intervals per specific capacity	t_k, d, e, s	$T_{d,e}^{s} = \{t_{start}, t_{start+1}, \ldots, t_{end-1}, t_{end}\}$
Employee ID	i	$I = \{1, 2, 3, \ldots, 298, 299, 300\}$

Here, time interval 4:00 denotes the hour between 4 and 5, 5:00 denotes the hour between 5 and 6, etc. To ensure consistency within shifts and facilitate the application of constraints later on, t_k has been introduced as a specification of time interval t. Here, k denotes the index of the k-th time interval within the set $T_{d,e}^{s}$. For example, in a shift that runs from 7am to 3pm, t_{start} would correspond to 7am, while t_{end} would correspond to 3pm. This approach allows constraints to be applied generally across all sets $T_{d,e}^{s}$, rather than targeting specific time intervals for each unique set individually.

4.3 Parameters

The following section details the parameters used in the mathematical model, which are essential for defining the productivity, demand, workforce constraints, and costs associated with the workforce planning process. Most importantly, the workload distribution is defined by $F_d^{wd,t}$, which denotes the percentage of the daily workload that arrives during a specific time interval in a department on a weekday, as has been reported in Section 3.2. These fractions are treated as fixed parameters. Later, in the impact analysis, the effects of altering these fractions will be explored. All the parameters are detailed in Table 8.

Table 8: Parameters

Name	**Notation**	**Description**
Norm	N_d	Productivity norm in units per hour per employee for department $d \in D$.
Workload distribution	$F_d^{wd,t}$	Fraction of demand on weekday $wd \in WD$ and department $d \in D$ in time interval $t \in T$.
Minimal and maximal workforce	$Min_{d,e,}^{s}\ Max_{d,e}^{s}$	The minimal and maximal workforce size in each capacity.
Costs	$C_{d,e}^{t}$	Costs for utilizing employee type $e \in E$ for department $d \in D$ in time interval $t \in T$.

4.4 Workforce, workload & productivity rate

In this section, the decision variable and auxiliary variables are introduced, both for the workforce and workload side of the optimization. Then, the two parts are reconciled by formulating the effects on the productivity rate.

<u>Workforce</u>

If an employee has been assigned to work in a specific capacity, this activity leads to requirements throughout the week. An active regular employee, for example, must meet a certain number of days in that specific capacity, whereas a subcontractor has similar requirements but no shift restrictions. To capture the activity status of an employee, two supplementary variables are added, which will be utilized later in the constraints. The workforce variables are summarized in Table 9.

Table 9: Workforce variables

Name	**Notation**	**Description**
Time interval assignment	$x_{d,e,i}^{wd,s,t}$	Binary <u>decision</u> variable stating whether employee $i \in I$ is assigned to a specific capacity in time interval $t \in T$ on weekday $wd \in WD$.
Week activity status	$y_{d,e,i}^{s}$	Binary variable stating whether employee $i \in I$ is assigned to a specific capacity during the week.
Week activity status without shift specification	$z_{d,e,i}$	Binary variable stating whether employee $i \in I$ is assigned to a specific capacity during the week, with liberal shift placement.

<u>Workload</u>

The total number of employees assigned to a time interval must be balanced against the workload $W_d^{wd,t}$. The workload depends on the demand A^{wd}, for which numerous input methods are possible. This input will be discussed in Chapter 5, where the case study design is explicated. Furthermore, the workload in a time interval also depends on how much of the workload in the previous time interval(s) has been completed. It is not immediately evident how high the leftover workload $L_d^{wd,t}$ at the end of each time interval must be, considering the relationship between workload, active pickers, and the pick rate. The variables related to the productivity rate and how it culminates into leftover workload are elaborated below Table 10, which displays a summary of the workload variables.

Table 10: Workload variables

Name	**Notation**	**Computation**	**Description**
Total demand	A^{wd}	Demand input	Total demand on weekday $wd \in WD$.
Time and department specific demand	$A_d^{wd,t}$	$A^{wd} * F_d^{wd,t}$	Demand on weekday $wd \in WD$ in department $d \in D$ and time interval $t \in T$.
Time and department specific workload	$W_d^{wd,t}$	$A_d^{wd,t} + L_d^{wd,t-1}$	Workload in department $d \in D$ and time interval $t \in T$, which is equal to the demand in the current time interval plus the leftover workload of the previous time interval.

<u>Productivity rate</u>

Based on the demand input, the optimization will focus on how many employees to allocate, taking into account the relationship between the available order lines, the active pickers, and the pick rate for the different departments. Due to overlapping shifts and multiple employee types being able to work simultaneously, the number of active pickers in a department during a specific time interval is defined as the workforce size $E_d^{wd,t}$ and is computed as follows:

$$E_d^{wd,t} = \sum_{e}^{E} \sum_{s}^{S} \sum_{i}^{I} x_{d,e,i}^{wd,s,t}, \forall d \in D, wd \in WD, t \in T$$

With the variables for available workload $W_d^{wd,t}$ and active pickers $E_d^{wd,t}$ established, the variables related to the productivity rate can now be defined. This mainly encompasses the theoretical, adjusted, and real productivity rates, which will be explained step by step below.

First of all, the Theoretical Productivity Rate (TPR) is the maximum rate achievable given the currently available workload. The pick rate analysis in section 3.4 showed that for the DPS department, with over 10,000 order lines available the pickers can attain the maximum pick rate. When order lines fall below this threshold, productivity begins to decline, dropping even more when the workload falls below 5,000. This relationship has led to the introduction of a hindrance factor H_d^t. This relationship has been modeled as a piecewise linear function such that the problem can be solved optimally. For DPS, the hindrance factor is as follows:

$$H_{DPS}^t = \begin{cases} 0.95, if\ W_{DPS}^{wd,t} < 5000 \\ 0.98, if\ 5000 \leq W_{DPS}^{wd,t} < 10000 \\ 1, if\ W_{DPS}^{wd,t} \geq 10000 \end{cases}$$

For Production and Cable, the pick rate heavily relies on the percentage of total workload available; picking activities commence only when 90% of the workload has been reached. If this threshold is met, a pick assignment is released; otherwise, it is deemed too inefficient. Consequently, if less than 90% of the daily workload is available, the hindrance factor is set to a deliberately low value. This ensures that the optimization model assigns employees only to those time intervals where the workload threshold has been sufficiently met:

$$H_d^t = \begin{cases} 0.1, if \dfrac{W_d^{wd,t}}{A_d^{wd}} < 0.9 \\ 1, if\ \dfrac{W_d^{wd,t}}{A_d^{wd}} \geq 0.9 \end{cases} \quad for\ d \in \{Production, Cable\}$$

For Expedition, the hindrance factor is always equal to 1, considering this department has a standardized work procedure that is not influenced by the workload or workforce size. Now, the TPR and the corresponding theoretical output TO can be defined as:

$$TPR_d^{wd,t} = N_d * H_d^t$$

$$TO_d^{wd,t} = TPR_d^{wd,t} * E_d^{wd,t}$$

Here, N_d represents the norm, or standard productivity level, set for the department. This formulation implies that when a certain workload is available, the TPR and TO denote the upper limit of picking performance in a specific time interval given the conditions.

The second main variable related to productivity is the Adjusted Productivity Rate (APR), which represents the maximum achievable rate given a specific workforce size. The APR can be regarded as a practical limit. If there are too many employees, the available workload might not be sufficient to maintain high pick rates. For instance, if the workload is 1,000 picks and the workforce size is 100, the pick rate can be at most 10 picks per employee per hour, even if the theoretical productivity rate is higher. This practical limit leads to the following definition of APR and the corresponding Adjusted Output (AO):

$$APR_d^{wd,t} = \min\left(\frac{W_d^{wd,t}}{E_d^{wd,t}}, TPR_d^{wd,t}\right)$$

$$AO_d^{wd,t} = APR_d^{wd,t} * E_d^{wd,t}$$

Ideally, the TPR and APR are equal as that would indicate sufficient workload and no productivity loss. However, if the TPR is higher than the APR, it signifies that the workload is insufficient for the allocated workforce, resulting in a shortage loss. This difference between the TPR and APR is a key performance indicator (KPI) used to measure productivity. Therefore, ensuring adequate workload for the allocated workforce is crucial to maintaining high productivity rates.

The final main variable is the Real Productivity Rate (RPR). The RPR builds on the TPR by further considering the relationship between the number of active pickers and the pick rate, again modeled with a piecewise linear relationship. While the APR focuses on the practical limits due to workforce size, the RPR incorporates inefficiencies caused by said workforce size. For DPS, inefficiencies arise when the workforce is too low, leading to employees covering larger distances, or too high, increasing the probability of malfunctions. To account for this, the inefficiency factor M has been introduced:

$$M_{DPS}^{wd,t} = \begin{cases} 0.95, if\ E_{DPS}^{wd,t} \leq 32 \\ 1, if\ 33 \leq E_{DPS}^{wd,t} \leq 48 \\ 0.90, if\ E_{DPS}^{wd,t} \geq 49 \end{cases}$$

The TPR is multiplied with this factor. For the other departments, the workforce size is already constrained within a small range, so the inefficiency factor plays a negligible role and has therefore been set to 1 for those departments.

Ideally, the APR and RPR are equal as that would indicate efficient workforce allocation and no productivity loss. Note that due to its practical limit, the RPR can never exceed the APR. Contrarily, if the APR is higher than the RPR, it signifies the workforce size has not been allocated efficiently, resulting in an inefficiency loss. This difference between the APR and RPR is the second KPI used to measure productivity. Therefore, ensuring an adequate workforce size is crucial to maintaining high productivity rates. The full list of KPIs used to measure performance will be discussed in Chapter 5. The RPR and the corresponding real output RO can be defined as:

$$RPR_d^{wd,t} = \min\left(TPR_d^{wd,t} * M_d^{wd,t}, APR_d^{wd,t}\right)$$

$$RO_d^{wd,t} = RPR_d^{wd,t} * E_d^{wd,t}$$

The RO represents the actual output produced given the workforce size and available workload. At the start of each time interval, incoming demand determines the workload. The workload minus the RO is the remaining work at the end of the time interval, which must be carried over to the next interval and added to the new workload. Thus, finally, the leftover workload can be defined as:

$$L_d^{wd,t} = W_d^{wd,t} - RO_d^{wd,t}$$

To illustrate the differences between the theoretical productivity rate, adjusted productivity rate, and real productivity rate, consider three scenarios, as summarized in Table 11. For clarity, a productivity norm of 100 picks per employee per hour, a hindrance factor of 1, and simplified values for the inefficiency factor are used.

In the first scenario, with a workload of 500 picks and 10 employees, the TPR is 100 picks per hour, resulting in a theoretical output of 1000 picks. However, due to a workload shortage, the APR is practically limited at 50 picks per hour. The inefficiency factor at the current employee level is 0.8, so the RPR is constrained by the APR limit of 50 picks per hour, leading to a real output of 500 picks. The loss of output in this case can be attributed solely to shortages. Then, in the second scenario, with a workload of 1000 picks and 10 employees, the TPR remains 100 picks per hour, and since there is no workload shortage, the APR is also 100 picks per hour. However, the inefficiency factor of 0.8 sets the RPR to 80 picks per hour and the RO to 800 picks. This loss can be attributed solely to inefficiency.

In the third scenario, with a workload of 1000 picks and 12 employees, the TPR is 100 picks per hour, yielding a TO of 1200 picks. In this scenario the workforce size has been slightly increased. The APR is then 83 picks per hour due to a workload shortage, resulting in an adjusted output of 1000 picks. The increased workforce size further reduces the inefficiency factor to 0.6, decreasing the RPR to 60 picks per hour and the RO to 720 picks. In this final scenario, productivity losses occur both due to shortages in workload and inefficiency of the workforce size. As the leftover workload in the final column demonstrates, it is not always as straightforward as simply increasing the workforce size to achieve higher output. In some cases, it may lead to inefficient use of personnel, whereas in other situations the available workload may not be sufficient to maintain satisfactory productivity rates. Finding this balance in relation to costs across all time intervals of the planning period becomes a crucial element of the optimization framework.

Table 11: Example scenarios of TPR, APR, and RPR

Scenario	Workload	Workforce	TPR	TO	APR	AO	M	RPR	RO	L
1	500	10	100	1000	50	500	0.8	50	500	0
2	1000	10	100	1000	100	1000	0.8	80	800	200
3	1000	12	100	1200	83⅓	1000	0.6	60	720	280

4.5 Objective function

The objective of the workforce planning model is to minimize the total costs across all order fulfillment departments for every weekday in the week, as shown in Table 12. The function is subject to several constraints, which will be discussed in the subsequent section.

Table 12: Objective function

Component	Minimize	Description
Workforce	$\sum_{wd}^{WD}\sum_{d}^{D}\sum_{e}^{E}\sum_{s}^{S}\sum_{t}^{T}\sum_{i}^{I}(C_{d,e}^{t} * x_{d,e,i}^{wd,s,t})$	Minimize the total costs of utilizing workforce size x during the entire week.

4.6 Constraints

The optimization model includes a variety of constraints to ensure the integrity of the workforce planning process. These constraints ensure employee consistency both during the day and throughout the week. Specifically, there are constraints for maintaining consistency for regular employees and subcontractors, with the latter having more liberal shift placements, meaning they are not constrained to the same type of shift on different days. Additionally, the model respects departmental personnel constraints to ensure that each department has the necessary workforce, but does not exceed the limits placed on them by the availability of, for example, specific machines or work stations. The model also ensures that the daily workload is reduced to zero by the end of each day, or else no workforce would be allocated at all. Furthermore, there is a constraint that addresses the dependency between departments, specifically the dependency of Expedition on the other three order fulfillment departments. The constraints are summarized in Table 13.

Table 13: Constraints

#	Shift constraints	Description
1a	$\sum_{d}^{D}\sum_{e}^{E}\sum_{s}^{S} x_{d,e,i}^{wd,s,t_{start}} \leq 1, \forall\, i \in I, for\ t_{start} \in T_{d,e}^{s}, for\ s \in S \backslash Overtime$	If an employee is assigned to the first time interval of a shift (t_{start}), the same employee cannot be assigned to the first time interval of another shift. This constraint ensures that an employee works a maximum of one shift each day, with the exception of the overtime shift.
1b	$x_{d,e,i}^{wd,s,t_k} = x_{d,e,i}^{wds,t_{k+1}}, \forall\, t_k \in T_{d,e}^{s} \backslash t_{end}, for\ s \in S \backslash Overtime$	If an employee is assigned to a certain time interval, they must also be assigned to the subsequent time interval of the shift. This and the previous constraint ensure that an active employee is assigned to all time intervals within a shift and exactly one shift.
1c	$x_{d,e,i}^{wd,s,t} = 0, \forall\, t \notin T_{d,e}^{s}$	An employee cannot be assigned to a time interval that falls outside a specific capacity.
1d	$x_{d,e,i}^{wd,Overtime,t} \leq x_{d,e,i}^{wd,Evening,t}, \forall i \in I, \forall t \in T_{d,e}^{Overtime}$	An employee can only be assigned to the overtime shift if assigned to the evening shift.
1e	$x_{d,e,i}^{wd,Overtime,t_k} = x_{d,e,i}^{wd,Overtime,t_{k-1}}, \forall t_k \in T_{d,e}^{Overtime} \backslash t_{start}$	If assigned to the k-th time interval $t_k$in the overtime shift, it must also be assigned to the preceding time interval within that shift, except for the first time interval in that shift, t_{start}.
2a	$y_{d,e,i}^{s} \geq x_{d,e,i}^{wd,s,t}, for\ e = RegularF, RegularP$	If a regular employee is assigned to any time interval at a specific capacity, it means this employee will have an active status in this week for that specific capacity.
2b	$\sum_{d}^{D}\sum_{e}^{E}\sum_{s}^{S} y_{d,r,e,i}^{s}, \leq 1, \forall\, i \in I$	If a regular employee has an active status during a week in a specific capacity, they can only be active in that specific capacity during the week. E.g., a regular employee being assigned to DPS/Morning on Monday must operate in the same capacity if they get an assignment on a different weekday.
2c	$\sum_{wd}^{WD} x_{d,e,i}^{wd,s,t_{start}} \geq 4 * y_{d,e,i}^{s}, for\ e = RegularF$ $\sum_{wd}^{WD} x_{d,e,i}^{wd,s,t_{start}} \geq 2 * y_{d,e,i}^{s}, for\ e = RegularP$	If a regular employee has an active status in the week and they are a full-timer, they must be assigned to a minimum of 4 weekdays. The part-timers must be assigned to a minimum of 2 weekdays.
3a	$z_{d,e,i} \geq x_{d,e,i}^{wd,s,t}, for\ e = SubcontractorF, SubcontractorP$	If a subcontractor employee is assigned to any time interval at a specific capacity, not taking into account the shift, it means this employee will have an active status in this week in that specific capacity.
3b	$\sum_{d}^{D}\sum_{e}^{E} z_{d,r,e,i}, \leq 1, \forall\, i \in I$	If a subcontractor employee has an active status during a week in a specific capacity, not taking into account the shift, they can only be active in that specific capacity during the week.

		E.g., a subcontractor being assigned to DPS/Outbound/Morning working on Monday must operate in the same capacity if they get an assignment on a different weekday, but it does not have to be the morning shift.
3c	$\sum_{wd}^{WD}\sum_{s}^{S} x_{d,e,i}^{wd,s,t_{start}} \geq 4 * z_{d,e,i}\text{ , } for\ e = SubcontractorF$ $\sum_{wd}^{WD}\sum_{s}^{S} x_{d,e,i}^{wd,s,t_{start}} \geq 2 * z_{d,e,i}\text{ , } for\ e = SubcontractorP$	If a subcontractor employee has an active status in the week and they are a full-timer, they must be assigned to a minimum of 4 weekdays. The part-timers must be assigned to a minimum of 2 weekdays.
4	$y_{d,e,i}^{s} + z_{d,e,i} \leq 1, \forall i \in I$	An employee can only be active as either a regular or a subcontractor, but not both, during the week.
5	$Min_{d,e}^{s} \leq \sum_{i}^{I} x_{d,e,i}^{wd,s,t} \leq Max_{d,e}^{s}$	The workforce size in a specific capacity must respect the lower and upper bounds of the personnel in that capacity.
6	$0.2 \leq \frac{\sum_d^D \sum_e^E \sum_i^I z_{d,e,i}}{\sum_d^D \sum_e^E \sum_s^S \sum_i^I (y_{d,e,i}^{s} + z_{d,e,i})} \leq 0.5$	The workforce during the week must consist for a minimum of 20% and a maximum of 50% of subcontractors.
7	$L_{d}^{wd,t} = 0, for\ t = 27{:}00$	At the end of the day, the leftover workload must be zero.
8	$RO_{DPS}^{wd,t} + RO_{Production}^{wd,t} + RO_{Cable,}^{wd,t} = W_{Expedition}^{wd,t+1}$	The sum of the actual output of DPS, Production and Cable in time interval $t \in T$ is the workload of Expedition an hour later.
9	$x_{d,e,i}^{wd,s,t} \in \{0,1\}, \forall\ wd, s, t, d, e, i$ $y_{d,e,i}^{s} \in \{0,1\}, \forall\ wd, s, d, e, i$ $z_{d,e,i} \in \{0,1\}, \forall\ wd, d, e, i$	Binary constraints.

Chapter 5 – Case study design

The mathematical model developed in the previous chapter is now ready to be applied to the case study environment. In this chapter, the design of the case study will be discussed. One crucial decision for the application of the model is the input for the demand. The performance of the model, like any model, is greatly influenced by how demand is incorporated. Overestimating or underestimating demand could lead to suboptimal outcomes. Therefore, four approaches for incorporating demand are proposed.

The first approach uses the current situation, employing a company-designed forecast method for demand without optimizing the workforce. The second approach uses the same forecast method but applies the mathematical model from Chapter 4 for workforce optimization. The third approach introduces a stochastic element to the demand, while the fourth approach regards demand as deterministic to assess the value of perfect information. These approaches will be discussed step by step in section 5.1. Subsequently, section 5.2 will discuss the setup for the impact analysis regarding workload distribution. Then, section 5.3 delves into how the case study will be evaluated. Finally, section 5.4 provides an overview of the entire setup.

5.1 Approaches to incorporating demand

<u>Approach 1: Current situation</u>

In Approach 1, the performance of the company during the review period of 2023 is examined. The company has developed a forecast method that it consistently applies to its planning. This method involves taking the forecast from the same day of the previous year and adjusting it slightly. Thus, forecasts for the entire year are made at the beginning of the review period, allowing the company to predict approximate demand for specific days, with some exceptions like holidays.

To create a workforce plan using this forecast method, the company follows these steps: First, the forecasted demand is split by department. Then, the demand is converted into a workforce requirement in hours, based on the productivity norm per employee per hour. For example, if the forecast is 1000 and the norm is 100, 10 workforce hours are required. Finally, the required hours are allocated to the morning, afternoon, and evening shifts, with allocations of 22%, 40%, and 38% respectively. Planners manually assign available employees to the shifts until the required hours for each shift are filled.

This approach has been the company's customary method. However, the fixed percentage allocation per shift may not be optimal, especially considering interdepartmental differences and the variation in workload distribution across weekdays. Approach 1 serves as a baseline to quantify the company's performance with its current method. The other approaches will compare this baseline to see if optimizing workforce planning improves performance.

<u>Approach 2: Current situation, optimized</u>

In the second approach, the company's demand forecasts are used as input, but the fixed percentage allocation is replaced with the mathematical model. This model aims to provide a more precise allocation based on the forecasted demand, producing a workforce plan for each time interval. This optimized approach seeks to improve efficiency by aligning workforce allocation more closely with actual demand patterns.

<u>Approach 3: Stochastic situation</u>

In the previous scenario, workforce planning is based on a single demand forecast for a given day, assuming this forecast will be accurate. However, in reality, demand is unpredictable and inherently uncertain, often fluctuating above or below the initial forecast. While the forecast aims to closely approximate actual demand, it does not account for the asymmetrical costs associated with inaccurate predictions. For example, overtime costs are twice as high as the costs of inefficient workforce planning. Therefore, it would be more cost-effective to have forecasts that tend to be higher than actual demand rather than lower. The current approach does not address this issue.

To account for demand uncertainty and minimize the long-term costs of workforce imbalances, the Newsvendor model will be applied to the forecast error. This model helps determine whether it is better to anticipate higher or lower demand than forecasted, and by what margin. The Newsvendor model, a classical inventory management tool, aims to find the optimal order quantity—or in this case, the optimal units that should be added or subtracted from the forecast—that minimizes total relevant costs (Nahmias & Olsen, 2015). The Newsvendor formulation is as follows:

$$Q^* = F^{-1}(\frac{C_o}{C_o + C_u})$$

Where:

- Q^* is the optimal quantity to order.
- F^{-1} is the inverse of the cumulative distribution function of demand.
- C_o is the cost of overestimating demand; C_u is the cost of underestimating demand.

In this context, Q^* represents the optimal workforce adjustment based on the forecast error distribution.

To apply the Newsvendor problem to the forecast errors of the current forecast method and acquire an optimally adjusted forecast, it is necessary to ensure that the forecast method is suitable. This requires verifying that the residuals of the forecasts are normally distributed and that the method is an unbiased estimator of the actual demand. If these conditions are met, the mean forecast and the standard deviation of the forecast error (FE) can be used as estimators of the probability distribution. Given that the demand pattern and corresponding forecast errors vary by weekday, this method has been checked for each weekday separately.

Generally, the prediction errors are assumed to be normally distributed, except when their CoV is larger than one, making this assumption less probable (Montgomery & Runger, 2010). If this is not the case, the mean forecast and standard deviation of the error can be used to estimate the properties of the distribution. For each weekday, the CoV is lower than one, as shown in Table 14. To confirm this, a Chi-square goodness-of-fit test has also been conducted. The p-value for each weekday is significantly higher than 0.05, indicating no reason to reject the null hypothesis, which means the prediction errors can be assumed to be normally distributed. All relevant values are summarized in Table 14. As can be seen, the bias is quite large, and must therefore be incorporated into the forecast calculation.

Table 14: Relevant values per weekday for the Newsvendor application

Weekday	Mean Forecast	Bias/Mean Forecast Error (FE)	Standard deviation of FE	Coefficient of Variation	Chi^2 GoF	P-value
Monday	84180	3413	5024	0.059	2.96	0.57

Tuesday	79565	3240	5217	0.065	3.42	0.63
Wednesday	74988	2127	4392	0.058	4.28	0.50
Thursday	75723	3411	5340	0.070	2.57	0.76
Friday	65169	3440	3501	0.053	4.39	0.49

Using the above information, the optimal forecast level to be used as input for the optimization can be determined as follows:

1. Calculate the critical ratio (CR) using the underage and overage costs.
2. Take the inverse of the cumulative distribution function (CDF) that equals the CR.
3. Compute the unbiased forecast using the formula: Unbiased forecast = Forecast – (Forecast/Mean forecast * Bias).
4. Determine the adjusted forecast using the formula: Adjusted forecast = Unbiased forecast + CDF^{-1}(CR) * Standard deviation of the forecast error (Std. dev. FE).

For each weekday of the review period, the optimal forecast level derived through this method will be used as input for the mathematical model.

Approach 4: Deterministic situation (benchmark)

In the fourth optimization approach, demand is considered to be known ahead of time and thus completely deterministic. In this scenario, the input for the mathematical model is the actual demand, eliminating any difference between the forecast and the demand that has actually arrived. This hypothetical situation aims to measure the performance of having a perfect forecast, providing an upper limit for the optimization model. The goal is to evaluate the potential performance of the optimization model if demand were perfectly known in advance. Since the deterministic situation is hypothetical, the outcomes can also be regarded as a benchmark to which other approaches are compared.

The four optimization approaches discussed serve to provide insights into how workforce optimization relates to the current situation, and the impact of forecast quality on performance. By comparing the different approaches, the importance of accurate forecasting in workforce planning can be better understood. The role of forecasting can then be juxtaposed against the role of the workload distribution, which is the topic of the subsequent impact analysis.

5.2 Impact analysis

An impact analysis will be conducted to assess the effect of changing the workload distribution throughout the day. The previously discussed approaches are designed to show the effect of optimization and the degree to which forecasting affects performance. Here, the models have been subjected to the fractions of the workload distribution that were found during the data analysis in section 3.2. In the impact analysis, the impact of altering these fractions will be studied.

A significant percentage of the daily demand arrives in the afternoon time intervals, which puts a strain on the end time of the evening shift. The company is considering two policies that could potentially alleviate this strain. In policy 1, the key is to shift a percentage of the demand that arrives in the time intervals starting from 17:00. A portion of the demand arriving during these intervals will be moved to the first time interval of the next working day. The impact analysis will then study the effects of increasing or decreasing this percentage. The second policy is a variant of policy 1, which states that 100% of the demand from all time intervals will be shifted to the next day. In practice, this equates to a 48-hour delivery policy, where pickers will be picking all the orders that arrived yesterday for delivery the following day. The added benefit is that the total workload will already be known at the beginning of the workday. The company aims to see the impact of delivering orders placed on a certain day not the next day, but the day after that. This approach would allow for accurate and efficient workforce allocation. This hypothetical situation combines the advantages of deterministic forecasts, as all demand from the previous day is known, with complete workload availability at the start of the day. The results of this analysis will serve as an upper limit, as achieving better performance is practically impossible.

5.3 Evaluation

During the process analysis, a set of KPIs has been identified to evaluate workforce planning. These KPIs have been crucial for developing a mathematical model aimed at minimizing total daily costs, while also accounting for various secondary effects of workforce planning that are of interest to the company. The most significant KPIs include the total daily costs, the total hours allocated, and the overtime hours required.

Total daily costs C_d^{wd} reflect the overall expense incurred in a day, while the total workforce size E_d^{wd} and overtime hours used OT_d^{wd} provide insights into workforce utilization. The real productivity rate RPR_d^{wd} and real output RO_d^{wd} measure the average productivity and output before midnight, respectively, taking into account the number of pickers. The percentage of workload completed by specific times (p_{17}, p_{20}, p_{23}) indicates critical milestones in the workflow: p_{17} denotes progress during the peak workload at the end of 17:00, p_{20} represents completion before the cutoff for order acceptance at 20:00, and p_{23} signifies the ideal end of all workload at midnight. Additionally, productivity loss due to workload shortages α_d and inefficiency β_d are measured to understand losses in pick rate caused by scheduling mismatches or operational issues. The utilization of the DPS system γ shows how effectively the system is used as a percentage of its maximum capacity. Table 15 summarizes the KPIs discussed.

Table 15: Evaluation KPIs

Name	Computation	Description
E_d^{wd}	$\Sigma_{t=4}^{27} E_d^{wd,t}$	The combined hours allocated
OT_d^{wd}	$\Sigma_{t=24}^{27} E_d^{wd,t}$	The combined overtime hours needed
C_d^{wd}	$\Sigma_{t=7}^{27} C_d^{wd,t}$	Total daily costs in €
RPR_d^{wd}	$\frac{\Sigma_{t=7}^{23} RPR_d^{wd,t}}{16}$	Real productivity rate in picks per hour per employee (p/he)
RO_d^{wd}	$\frac{\Sigma_{t=7}^{23} RO_d^{wd,t}}{16}$	Real output in picks (p)
p_{17}	$\frac{\Sigma_{t=7}^{17} RO_d^{wd,t}}{A_d}$	Percentage of workload finished at the end of time interval 17:00
p_{20}	$\frac{\Sigma_{t=7}^{20} RO_d^{wd,t}}{A_d}$	Percentage of workload finished at the end of time interval 20:00
p_{23}	$\frac{\Sigma_{t=7}^{23} RO_d^{wd,t}}{A_d}$	Percentage of workload finished at the end of time interval 23:00
α_d	$TPR_d^{wd,t} - APR_d^{wd,t}$	Productivity loss due to workload shortage in p/he
β_d	$APR_d^{wd,t} - RPR_d^{wd,t}$	Productivity loss due to inefficiency in p/he
γ	$\frac{RO_{DPS}^{wd,t}}{Max}$	Utilization DPS as a percentage of the maximum output

5.4 Optimization methodology

The optimization process consists of three stages. In the first stage, the collected demand datasets are used. These datasets contain the incoming demand on a given weekday along with the company's forecasts of the review period. To optimize workforce planning for the days of a given week, it is imperative to have sufficient weeks for comparison. Therefore, the demand dataset for all weeks of 2023 will be used. After removing outliers, irregular weeks, and holiday periods, a total of 12 data points were excluded, resulting in 40 full weeks for review. For each week, the workforce will be optimized for all weekdays, and this process will be repeated for each subsequent week. The average of the daily costs and corresponding KPIs will be presented based on performance over these 40 weeks. From the acquired demand dataset, for approach 1 and 2, the forecasted demand will be used as input. For approach 3, the adjusted forecast will be used after application of the Newsvendor model. For the benchmark approach, the actual demand will be used.

In the second stage, the mathematical model is implemented and initialized with parameter settings derived from process and data analyses. The model is then optimized using a solver, producing an optimal workforce plan for all days of a given week, specifying the ideal workforce size for each time interval. Thus, for a given week, the entire optimization framework is applied holistically, considering all days together. The third stage involves evaluating how the provided workforce plan for a weekday would have performed against the actual demand on that day. This evaluation uses data from the actual demand datasets per time interval. The corresponding costs and KPIs described in the previous section are then derived from this performance. Finally, in stage four, the impact analysis is conducted by altering the workload distribution. Figure 30 provides an overview of how the four stages and four approaches are tied together.

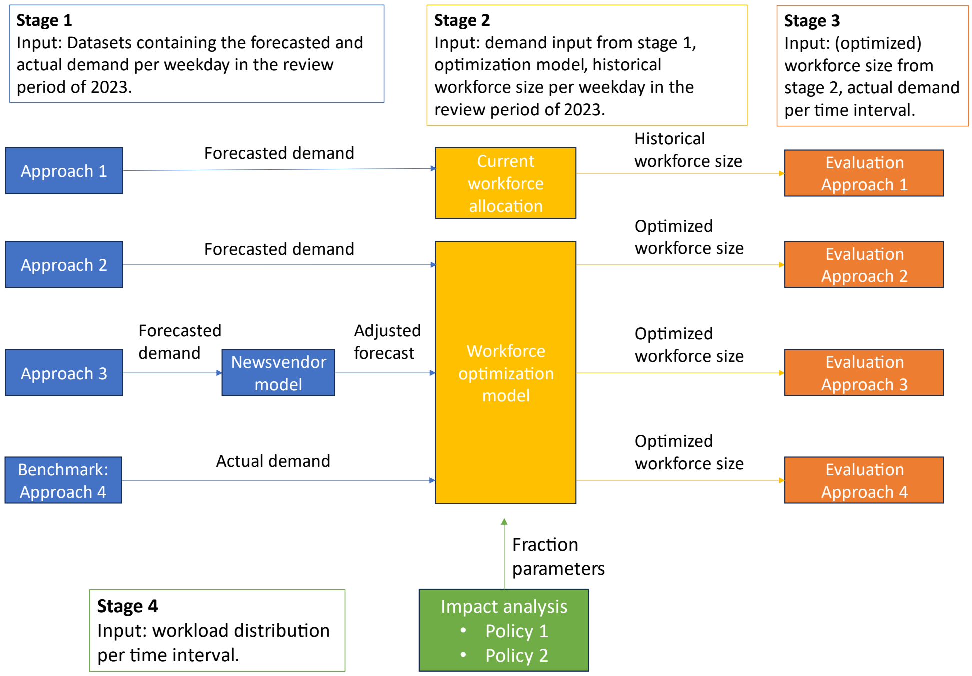

Figure 30: Stages of the optimization framework

All four stages of the optimization have been implemented in Python within the Spyder environment. The mathematical model has been optimized with the commercial solver Gurobi 11.0.1, using an academic license. Due to the nonlinearities within the model, including the hindrance and inefficiency factor, quadratic relationships ensued. The model is therefore solved using the mixed-integer quadratically constrained programming (MIQCP) package. The recorded results discussed later were achieved with a CPU model 12th Gen Intel(R) Core(TM) i7-1265U, taking approximately 4.5 minutes to optimize the workforce for a single week. This solution method is a convex solver, allowing for finding the optimal solution.

Chapter 6 – Results

6.1 Approaches to incorporating demand

Approach 1: Current situation

Many insights have already been obtained from the data analysis in Chapter 3. It was shown how the workforce varies by department throughout the day and highlighted overtime as a significant issue. Starting with the DPS department, Table 16 quantifies the effects of the original workforce planning, presenting the most important KPIs for the 40 weeks of the review period in 2023, with mean values shown per weekday and per week.

What is immediately evident is that the weekday pattern translates almost proportionally to the KPIs. On Monday, the highest demand must be handled, while Friday has the least, with a gradual decrease throughout the week. This trend is reflected in the workforce size E_d^{wd}, overtime hours OT_d^{wd}, and costs C_d^{wd}, which are highest on Monday and decrease as the week progresses. This pattern persists across most KPIs. A higher percentage of the workload is shifted on days with lower demand, which holds true for the 17:00, 20:00, and 23:00 time intervals, as denoted by p_{17}, p_{20}, and p_{23}. Additionally, the DPS system has higher utilization (γ) before midnight since more workload needs to be completed.

Another point of interest is the productivity rate RPR_d^{wd} and the corresponding shortage (α_d) and inefficiency (β_d) loss, in picks per hour per employee (p/he). On days with high workload, there appears to be sufficient workload available at all times for the workforce to maintain a high pick rate. However, there is also a need for a high workforce in the later time intervals, considering the relatively low percentage of workload finished by 17:00 and 20:00. This higher workforce increases inefficiency, leading to a substantial loss in the pick rate. These trends reverse as the week advances. Generally, the less workload on a given day, the higher the probability of achieving a more positive performance.

Table 16: Performance of Approach 1 (baseline situation)

	Monday	**Tuesday**	**Wednesday**	**Thursday**	**Friday**	**Mean**
$\boldsymbol{E_d^{wd}}$	820 hrs	783 hrs	733 hrs	735 hrs	615 hrs	737 hrs
$\boldsymbol{OT_d^{wd}}$	81 hrs	65 hrs	46 hrs	38 hrs	9 hrs	47.8 hrs
$\boldsymbol{C_d^{wd}}$	€23,892	€22,638	€21,050	€21,017	€17,239	€21,167
$\boldsymbol{RPR_d^{wd}}$	108.63	109.01	109.79	109.71	109.92	109.41
$\boldsymbol{RO_d^{wd}}$	4308	4215	4069	4101	3572	4053
$\boldsymbol{p_{17}}$	56.43%	57.07%	57.70%	59.58%	62.63%	58.68%
$\boldsymbol{p_{20}}$	73.51%	74.89%	75.93%	78.20%	82.40%	76.98%
$\boldsymbol{p_{23}}$	91.50%	93.48%	94.75%	96.25%	97.47%	94.69%
$\boldsymbol{\alpha_d}$	0.007	0.016	0. 0271	0.038	0.057	0.029
$\boldsymbol{\beta_d}$	3.82	3.44	2.57	2.59	2.46	2.98
$\boldsymbol{\gamma}$	86.17%	84.30%	81.39%	82.02%	71.45%	81.06%

Extensive results are provided for the DPS department due to its significant impact on overall operations, its complex workflows, and its critical role in the supply chain process. Conversely, the Production, Cable, and Expedition departments are much smaller in size and less complex. For these departments, a concise summary of key metrics such as workforce size, overtime hours, and costs is sufficient to capture their operational performance. The summarized results for these departments are as follows in Table 17.

Table 17: Performance of other order fulfillment departments

Department	Workforce size	Overtime hours	Costs
Production	155 hrs	3.6 hrs	€4380
Cable	216 hrs	3.9 hrs	€6092
Expedition	152 hrs	22.8 hrs	€4513

One notable observation is the overtime hours in the Expedition department. This is partly explainable because the evening shift extends until 1am to complete the last truck departures. However, even considering this extra hour, the overtime hours are still comparatively high. This indicates a need for further investigation into the factors driving this overtime and potential optimization to reduce these hours.

<u>Approach 2: Current situation, optimized</u>

Approach 2 is crucial for several reasons. Firstly, it serves as a validation of the model. Ideally, the model should allocate a similar number of hours as the actual workforce planning. If the model allocates significantly fewer hours, it might indicate that some modeling decisions are too generous compared to reality. Conversely, if the hours are too high, the model may be overly strict, leading to suboptimal results. As shown in Table 18, the model closely matches the actual allocation level. On high-demand days, it uses slightly fewer hours, while on low-demand days, it uses slightly more. Overall, the model aligns well with the actual workforce required to meet the workload.

Table 18: Performance Approach 2

	Monday	Tuesday	Wednesday	Thursday	Friday	Mean	+/- %
$\boldsymbol{E_d^{wd}}$	818 hrs	781 hrs	740 hrs	734 hrs	628 hrs	740 hrs	0.41
$\boldsymbol{OT_d^{wd}}$	24 hrs	15 hrs	14 hrs	12 hrs	7 hrs	14.4 hrs	-69.87
$\boldsymbol{C_d^{wd}}$	€23,169	€22,035	€20,873	€20,697	€17,646	€20,884	-1.34
$\boldsymbol{RPR_d^{wd}}$	110.85	110.69	110.93	110.69	111.38	110.91	0.90
$\boldsymbol{RO_d^{wd}}$	4609	4440	4232	4207	3642	4226	4.26
$\boldsymbol{p_{17}}$	61.31%	61.42%	60.62%	62.28%	68.94%	62.91%	10.99
$\boldsymbol{p_{20}}$	80.04%	80.94%	80.97%	82.57%	86.67%	82.24%	6.83
$\boldsymbol{p_{23}}$	97.72%	98.30%	98.38%	98.59%	99.04%	98.41%	3.92
$\boldsymbol{\alpha_d}$	0.247	0.519	0.381	0.359	0.210	0.343	1082.76
$\boldsymbol{\beta_d}$	0.511	0.389	0.587	0.745	0.445	0.535	-82.04
$\boldsymbol{\gamma}$	92.18%	88.81%	84.65%	84.15%	72.85%	84.53%	4.28

However, the manner in which the workforce is utilized differs considerably, which emphasizes the second crucial aspect of Approach 2: demonstrating the potential for enhancement through improved planning while keeping all other variables constant. This potential emerges in several ways, as captured in the final column of Table 18, showing the percentual difference with Approach 1. Firstly, overtime hours have been reduced by nearly 70%, as more employees are allocated before the afternoon peak. Additionally, the percentage of workload completed earlier in the day has increased, as reflected by p_{17}, p_{20}, and p_{23}. This improvement is evident in Figure 31, which shows the completion percentages for all time intervals past 17:00 for the optimized approach next to the current situation. A stark increase is noticeable for each weekday.

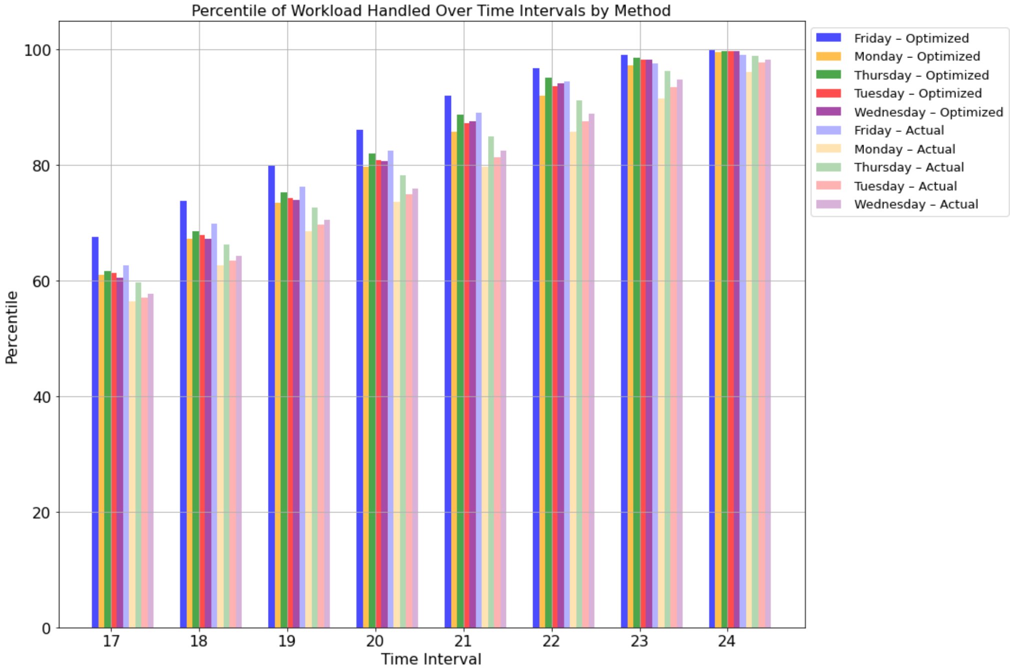

Figure 31: Percentile of workload handled over time intervals: Baseline (Approach 1) vs Optimized (Approach 2)

The trade-off made by the model is between achieving the productivity norm and managing the workload. The model clearly favors handling more workload, even if it means reducing the productivity rate in certain time intervals. This is evident in the increased shortage loss, indicating that more productivity is wasted due to workload being unavailable, compared to Approach 1. However, this occurs because a greater amount of workload is being completed overall. Despite the increase in shortage loss, the results demonstrate that this waste is acceptable. The reduction in inefficiency loss significantly outweighs the rise in shortage loss. Additionally, the real productivity rate has been higher on average across all days of the week, as illustrated in Figure 32. This suggests that the model's approach to balancing workload and productivity is effective, leading to overall improved performance.

While the improvements are positive, the translation to costs are somewhat modest, with a little bit over 1% reduction in the costs. This can still lead to substantial reductions due to the sheer volume in a year, but the cost reductions are constrained. This can be explained by the two drives of cost reductions, namely an increase in efficient order picking, or a decrease in overtime hours. As seen in Table 18, the overtime hours have been significantly reduced. However, the output that was originally created during overtime must still be completed. If this output is handled during regular hours instead of premium hours, the cost reduction will reflect the difference between regular and premium wages, rather than eliminating the total premium hours entirely. Ultimately, the improvements across the board are very positive, as demonstrated by the final column in Table 18.

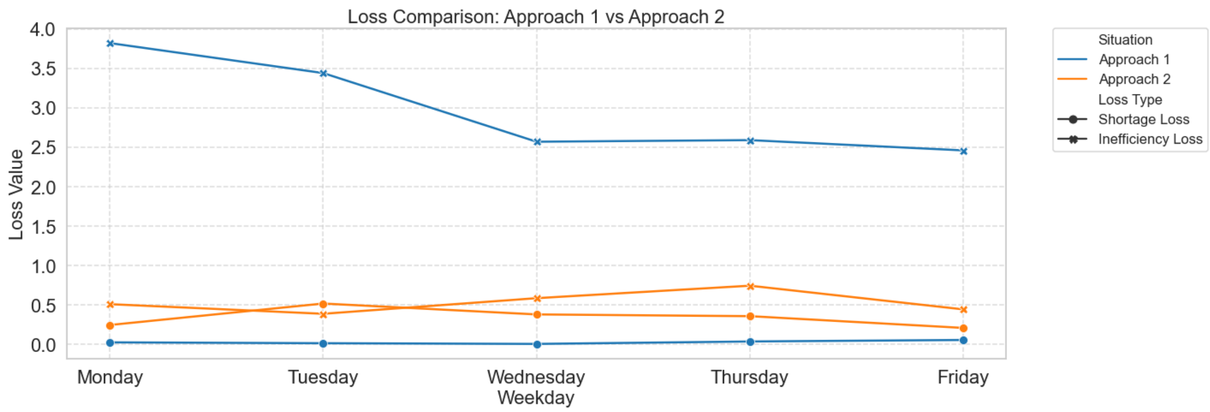

Figure 32: Productivity rate loss comparison: Approach 1 vs Approach 2

When examining the other departments, the results are virtually the same as in Approach 1. There are several reasons for this. Firstly, Production and Cable have a unique way of picking orders. Unlike DPS, which picks orders in real time, these departments wait until a certain threshold is met before releasing orders. This approach allows them to pick orders efficiently from the pallet bays. The efficiency difference is substantial. If orders were picked in real time, the number of trips to the pallet bays would increase dramatically, making the process inefficient. Therefore, these departments start work in the later time intervals when more workload has accumulated. This method underscores the systemic effects of workload distribution, which will be explored further in the impact analysis.

Secondly, workforce constraints contribute to the marginal changes. These departments often operate with a limited supply of employees, preventing them from utilizing a larger workforce. Additionally, the minimal required workforce is around 10 employees, resulting in the workforce size in regular intervals typically ranging between 10 and 15 employees. This limited range restricts the potential for flexibility in workforce planning, unlike DPS, where the gap between minimum and maximum workforce size is significantly larger.

Thirdly, forecasting plays a less critical role in these departments. Because they maintain a stable workforce size and begin picking only after a high percentage of the workload has arrived, the quality of the forecast is less impactful. In contrast to DPS, any discrepancy between the forecast and actual demand is much less pronounced in Production and Cable due to their lower demand size. This reduces the potential for optimization, as work must begin at a later stage regardless of forecast accuracy, given the nature of workload arrival. The results of the subsequent analyses for these departments are stored in Appendix C.

Approach 3: Stochastic situation

For Approach 3, no significant differences were found in the results compared to Approach 2. The results have been summarized in Table 19. The key difference in this approach was the use of the Newsvendor model to alter the forecast used in Approaches 1 and 2. While the new forecasts have been improved to account for the asymmetry in costs between regular time and overtime, this adjustment has resulted in only marginal changes: fewer hours are used on average, but more overtime hours are incurred. This may be because the altered forecasts are lower than the original forecasts. Consequently, during periods of higher demand than anticipated, Approach 3 relies more on overtime than Approach 2. Conversely, if demand is lower than anticipated, Approach 3 plans for a smaller workforce size. The differences between Approaches 2 and 3 are minimal, which can be attributed to the stable nature of demand. Throughout the review period of 2023, the demand exhibited little erratic fluctuation for all weekdays, with a maximum CoV of 0.21. This stability in demand means that the benefits of applying the Newsvendor model to adjust forecasts are less pronounced.

Table 19: Performance Approach 3

	Monday	Tuesday	Wednesday	Thursday	Friday	Mean	+/- %
E_d^{wd}	817 hrs	779 hrs	740 hrs	733 hrs	627 hrs	739 hrs	0.27
OT_d^{wd}	25 hrs	18 hrs	14 hrs	14 hrs	13 hrs	16.8 hrs	-64.85
C_d^{wd}	€23,159	€22,000	€20,873	€20,676	€17,702	€20,882	-1.35
RPR_d^{wd}	111.00	110.82	111.01	110.54	111.81	111.04	1.02
RO_d^{wd}	4588	4430	4228	4194	3606	4209	3.84
p_{17}	60.83%	61.14%	60.40%	61.25%	66.87%	62.10%	9.56
p_{20}	79.54%	80.60%	80.75%	81.55%	84.18%	81.32%	5.63
p_{23}	97.29%	98.08%	98.31%	98.30%	98.10%	98.01%	3.50
α_d	0.364	0.420	0.348	0.753	0.040	0.385	1227.59
β_d	0.313	0.434	0.550	0.617	0.393	0.461	-84.53
γ	91.76%	88.60%	84.57%	83.89%	72.13%	84.19%	3.86

Approach 4: Deterministic situation (benchmark)

The final approach assumes demand is fully known, allowing the daily workload to be optimally matched with the workforce. This approach theoretically yields the best possible performance, hence it is portrayed as the benchmark. As shown in Table 20, this method indeed yields improvements across the board, including lower hours utilized, reduced overtime hours, decreased daily costs, fewer losses due to shortages and inefficiencies, and a higher percentage of workload completed earlier in the day with greater utilization of DPS.

Remarkably, despite these optimizations, overtime hours still average 12.6 hours per day. One would expect that if forecasts perfectly aligned with actual demand, the workforce could be precisely tailored to the workload, thereby minimizing overtime. However, the persistence of 12.6 overtime hours indicates that this is not entirely the case. The primary reason is the distribution of demand throughout the day. While perfect forecasting can significantly minimize downsides, reduce costs, and improve other key metrics, the inherent nature of workload distribution means that overtime will remain a factor. The impact of altering the distribution of the workload, and how it could further optimize the system, will be the focus of the next section.

Table 20: Performance Approach 4 (benchmark)

	Monday	Tuesday	Wednesday	Thursday	Friday	Mean	+/- %
E_d^{wd}	814 hrs	773 hrs	735 hrs	729 hrs	625 hrs	735 hrs	-0.27
OT_d^{wd}	23 hrs	14 hrs	7 hrs	9 hrs	6 hrs	11.8 hrs	-75.31
C_d^{wd}	€23,040	€21,807	€20,649	€20,516	€17,567	€20,715	-2.17
RPR_d^{wd}	111.34	111.63	111.80	111.82	112.24	111.77	1.68
RO_d^{wd}	4600	4457	4271	4227	3627	4236	4.51
p_{17}	60.65%	60.05%	59.19%	59.42%	64.60%	60.78%	7.23
p_{20}	79.31%	79.53%	79.42%	79.62%	81.87%	79.95%	3.85
p_{23}	97.52%	98.64%	99.22%	99.04%	98.62%	98.61%	4.13
α_d	0.224	0.183	0.076	0.144	0	0.125	331.03
β_d	0.202	0.225	0.396	0.366	0.306	0.299	-89.96
γ	92.00%	89.14%	85.43%	84.54%	72.55%	84.73%	4.52

6.2 Impact analysis

While it has become clear that optimizing the workforce planning enhances a lot of the areas that the operation is currently lacking in, it also proved that these enhancements are constrained because of systemic limitations. For periods with high demand, the workforce is still subjected to overtime, despite optimized planning and despite improving the quality of the forecast. However, in other areas, the improvements are very solid.

Having established the potential of optimized planning with the current forecast method, it is worthwhile to explore the potential of changing the workload distribution. Under the first policy, the company aims to shift demand that arrives after 17:00 to the next day. This approach would mean that orders arriving after this time are handled the next day and delivered the day after, thereby alleviating the total workload during peak hours and before the midnight deadline. To determine the optimal share of orders to shift after 17:00, an impact analysis has been conducted, ranging from 0% (no workload shifted) to 100% (all workload shifted). The impact on workforce size, total costs, and the percentage of picks completed before midnight has been plotted in Figure 33.

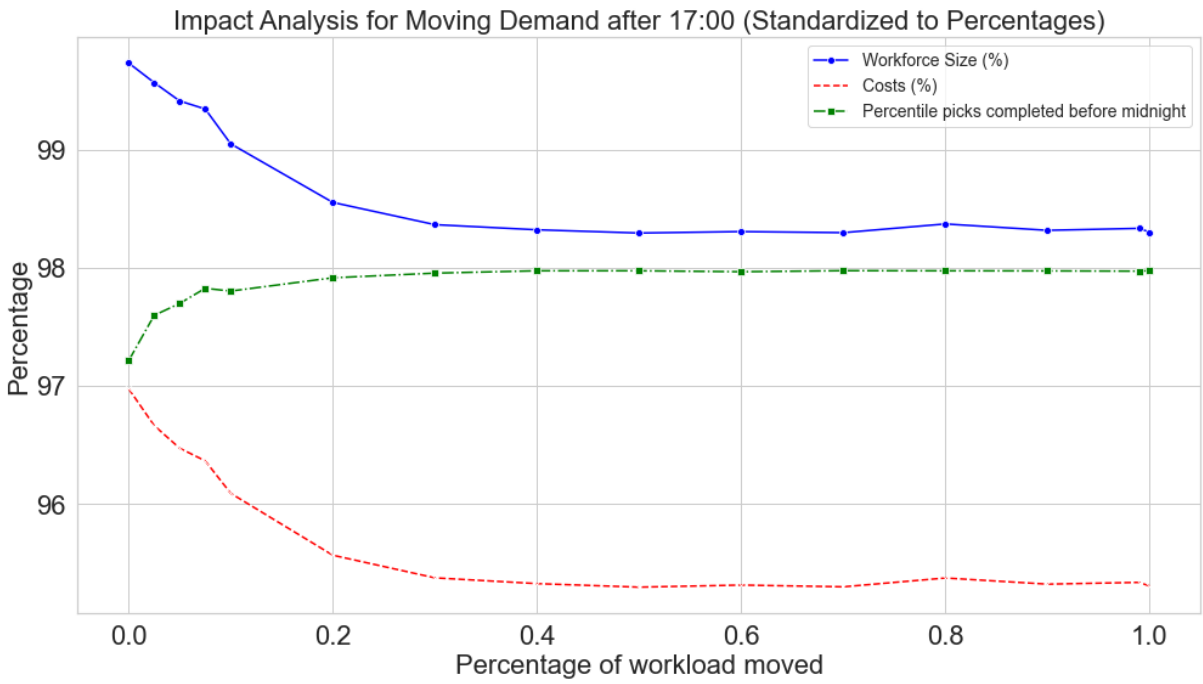

Figure 33: Impact analysis of policy 1: moving demand after 17:00

In the graph, the results of the current situation from Approach 1 are used as the baseline, representing 100%. Initially, with no workload shifted, the values reflect the improvements from Approach 2, which only optimizes the planning. As more workload is shifted after 17:00, the subsequent improvements are shown. The analysis reveals that all key measures—workforce size, daily costs, and percentage of workload completed before midnight—improve as more workload is shifted to the following day. More interestingly, the outcomes imply that peak improvements are achieved relatively quickly. Beyond shifting 30% of the workload, additional performance gains become marginal, indicating diminishing returns. This suggests that while shifting workload can significantly enhance efficiency and reduce costs, the majority of benefits are realized with moderate adjustments, and further changes yield progressively smaller improvements.

In the second policy, the entire day's workload is shifted to the next day, effectively implementing a 48-hour delivery policy. Customers placing orders today will have their orders picked tomorrow and delivered the day after. This policy offers maximal advantages because the exact workload is known in advance, and all the workload is already present at the beginning of the workday, eliminating systemic issues and enabling optimal performance. The results are summarized in Table 21.

Table 21: Impact analysis of policy 2: 48-hours delivery

	Monday	Tuesday	Wednesday	Thursday	Friday	Mean	+/- %
E_d^{wd}	807 hrs	761 hrs	726 hrs	720 hrs	616 hrs	726 hrs	-1.49
OT_d^{wd}	0 hrs	0 hrs	0 hrs	0 hrs	0 hrs	0 hrs	-100
C_d^{wd}	€22,596	€21,294	€20,314	€20,168	€17,236	€20,321	-3.99
RPR_d^{wd}	112.27	112.46	112.54	112.49	112.19	112.39	2.25
RO_d^{wd}	4730	4488	4285	4253	3631	4277	5.52
p_{17}	65.21%	65.99%	67.60%	68.04%	67.07%	66.78%	17.81
p_{20}	82.51%	83.18%	83.99%	84.23%	83.77%	83.54%	8.52
p_{23}	99.96%	99.97%	99.99%	99.99%	100%	99.99%	5.59
α_d	0.021	0.048	0.049	0.066	0.095	0.056	93.10
β_d	0.463	0.088	0	0.014	0.255	0.164	-94.49
γ	94.61%	89.79%	85.71%	85.07%	72.62%	85.55%	5.53

What is evident is that the overtime hours have been completely eliminated. Each weekday, zero hours are needed after midnight to complete the picks, which translates into a nearly 4% reduction in costs—a substantial improvement and triple the reduction possible with Approach 2, which only optimized workforce planning.

Another advantage of postponing workload is that the workforce can be allocated efficiently at each time interval, preventing losses from inefficient use or malfunctions. Moreover, since there is always plenty of workload available, shortage losses are minimized. However, these losses are not entirely eliminated because during the final time interval of the day, the workforce may not be active for the entire period, leading to a reduced pick rate. Despite this, the pick rate remains significantly higher than in the current situation and other approaches due to the ideal circumstances. Ultimately, this translates to a smaller required workforce size. Not only are overtime hours decreased, but the overall workforce size is also reduced due to more efficient usage.

Chapter 7 – Discussion

7.1 Conclusion

In the conclusion of this thesis, the main research question will be answered: *How can an aggregate planning approach effectively optimize workforce size and workload distribution in response to stochastic demand in order fulfillment operations?* This is done by revisiting the research questions.

Firstly, the process analysis mapped the order fulfillment workflow. It highlighted how different departments within DCA operated, detailing the functions each department performed and the resources used. The analysis revealed which departments contributed to order fulfillment and illustrated the differences between real-time order picking in DPS, order collection and releasing after the order deadline by Production and Cable, and the dependency on all three by Expedition. Additionally, key parameters such as minimum and maximum workforce size, production rate norms, and shift times were identified, providing a comprehensive overview for optimizing workforce size.

Secondly, the data analysis provided insights into workload, workforce, and their interplay. Clear patterns emerged regarding workload distribution across weekdays and time intervals, with a significant portion of the workload arriving in the afternoon, creating pressure on the midnight end time. This pattern was reflected in workforce allocation, where overtime was prevalent. The severity of this issue was quantified, highlighting the importance of DPS, which handles 90% of the workload with 50% of the workforce. The analysis also examined the relationship between workload levels and pick rates, providing valuable information for workforce optimization.

Thirdly, insights from the process and data analyses were integrated into a mathematical model, considering weekdays, departments, employee types, and corresponding constraints. The current situation served as a baseline, while the MIQCP model demonstrated the ability to achieve better results, reducing costs and improving predefined KPIs. The quality of demand input, including stochastic elements, was also incorporated. Although stochasticity had an insignificant impact, the quality determining demand forecasts could influence the optimization of the aggregate plan.

Finally, the mathematical framework was used to test the impact of workload distribution. The analysis showed that inherent overtime issues remaining after optimization could be eliminated by redistributing the workload throughout the day. Notably, the required change was not extensive: shifting a moderate percentage of the post-17:00 workload to the next day nearly achieved maximum results.

7.2 Recommendations

The results of this thesis have demonstrated that improved planning can enhance performance in the current operational setting, even without altering other factors. The company is advised to consider implementing a mathematical model to optimize workforce aggregate planning, replacing the existing planning methods. By differentiating between weekdays and accounting for the effects of workforce size within specific time intervals, the company can achieve more aligned planning. This approach can lead to substantial cost reductions and secondary benefits, such as earlier completion of workloads and earlier end times.

In addition to better planning, accurate demand forecasting is crucial for effective workforce management. While the current situation has shown stable demand, the deterministic approach used in this research indicates that there is still room for improvement. By investing in a more advanced forecasting tool compared to the current method, the company can achieve large cost reductions on a yearly basis. Enhanced forecasting capabilities would enable more precise planning, allowing the company to better match workforce levels with actual demand, thus optimizing overall performance.

Finally, the analysis also highlighted the critical role of workload distribution. While improvements in planning and forecasting yield notable advantages, their effectiveness is limited by the current distribution of workload throughout the day, with a big portion arriving late. By shifting a part of the peak workload to the next day, the company can achieve further performance enhancements. The analysis indicates that even a moderate percentage shift can have a significant impact. Additionally, the company could experiment with shifting a percentage of the entire day's workload instead of just the peak intervals. Given that the total daily volume is larger, an even smaller percentage shift of the whole day could be more attainable than a larger percentage shift during peak times. Mechanisms with which desired shifts can be effectuated are discussed in section 7.4. The company is encouraged to explore the practical feasibility and associated costs of such adjustments.

7.3 Academic contributions

This thesis addresses a notable gap in the field of APP identified in the literature study. Traditional APP methods, which typically assume the possibility of inventory building and often overlook the dynamic nature of demand within operative periods, are not well-suited for make order fulfillment environments where production cannot commence ahead of incoming demand. These environments require a more granular analysis of demand variations within individual time intervals of a day.

To fill this gap, this research has developed a planning framework on a weekly basis, explicitly accounting for the workload distribution over the time intervals within specific weekdays. This framework operates under the constraint that building up inventory or allowing backorders is not possible. By focusing on the non-linear nature of production output and the irregular arrival of orders throughout the day, this thesis introduces a novel approach to workforce and production planning in order fulfillment environments. This approach optimizes workforce size and provides a more accurate method for managing resources in contexts where traditional APP methods may fall short.

Furthermore, the thesis includes an impact analysis of altering the workload distribution, building on the work of Jamalnia et al. (2017), who considered different policies for various shifts within a day to influence production levels. This research extends it by analyzing how changes in workload distribution within a day affect overall productivity.

7.4 Limitations & future research

Despite the findings and contributions of this thesis, several limitations must be acknowledged. First, the scope of this research was confined to an order fulfillment environment with a stable demand pattern. While this provided a controlled setting for analyzing workforce and production planning, it does not account for scenarios with highly volatile demand. Future research could explore the applicability of the proposed models in environments with greater demand variability to ascertain generalizability.

Secondly, the analysis primarily considered operational efficiency metrics, such as cost reduction and workload completion times. However, other factors, such as employee satisfaction, sickness leave, and long-term sustainability of workforce practices, were not within the scope of this study. While the assumption that advantages in performance also result in psychological advantages may seem valid—for example, earlier end times would mean employees are less overworked and thus mentally healthier—future research could incorporate these human resource dimensions to provide a more holistic view of workforce planning and its implications.

Lastly, a major assumption of this study is that it is possible to modify workload distribution by influencing customer behavior. Various research has demonstrated its feasibility through mechanisms such as differentiated pricing, dynamic pricing, discounts, or other pricing strategies (e.g., Berger & Nasr, 1998; Familmaleki, Aghighi & Hamidi, 2015; Den Boer, 2015). These methods can encourage customers to place orders at alternative times, resulting in a shift in the pattern of order arrivals at DCs and consequently altering the workload distribution. To fully understand the impact of a given pricing policy on customer behavior, a detailed customer analysis would be necessary. Due to the scope of this project's timeline, conducting this analysis was not feasible. Hence, the findings of this study are predicated on the assumption that the desired workload distribution can be achieved. This limitation presents an opportunity for future studies to build upon this foundation, possibly by expanding the objective function to include the costs associated with implementing the aforementioned pricing policies.

References

Accenture. (2022, May 10). *Impact of war in Ukraine on oil & gas industry*. https://www.accenture.com/us-en/insights/energy/ukraine-oil-gas

Benchmark International. (2023). 2023 Global Distribution Industry Report. *Benchmark International*. https://blog.benchmarkcorporate.com/2023-global-distribution-industry-report

Berger, P. D., & Nasr, N. I. (1998). Customer lifetime value: Marketing models and applications. *Journal of interactive marketing*, *12*(1), 17-30.

Bowman, E. H. (1956). Production scheduling by the transportation method of linear programming. *Operations Research*, *4*(1), 100-103.

Charnes, A., Cooper, W. W., & Farr, D. (1953). Linear programming and profit preference scheduling for a manufacturing firm. *Journal of the Operations Research Society of America*, *1*(3), 114-129.

Chen, T., He, T., Benesty, M., Khotilovich, V., Tang, Y., Cho, H., ... & Zhou, T. (2015). Xgboost: extreme gradient boosting. *R package version 0.4-2*, *1*(4), 1-4.

Deloitte. (2016, September 12). *Industry 4.0 and distribution centers*. Deloitte Insights. https://www2.deloitte.com/us/en/insights/focus/industry-4-0/warehousing-distributed-center-operations.html

Demirel, E., Özelkan, E. C., & Lim, C. (2018). Aggregate planning with flexibility requirements profile. *International Journal of Production Economics*, *202*, 45-58.

Den Boer, A. V. (2015). Dynamic pricing and learning: historical origins, current research, and new directions. *Surveys in operations research and management science*, *20*(1), 1-18.

EAM Mosca. (2022, May 24). *Strategies to deal with warehouse labor Shortages*. EAM Mosca. https://www.eammosca.com/blog/solving-warehouse-labor-shortages/

Familmaleki, M., Aghighi, A., & Hamidi, K. (2015). Analyzing the influence of sales promotion on customer purchasing behavior. *International Journal of Economics & management sciences*, *4*(4), 1-6.

Flowers, A. D., & Preston, S. E. (1977). Work force scheduling with the search decision rule. *Omega*, *5*(4), 473-479.

Hanssmann, F., & Hess, S. W. (1960). A linear programming approach to production and employment scheduling. *Management science*, (1), 46-51.

Holt, C. C., Modigliani, F., & Muth, J. F. (1956). Derivation of a linear decision rule for production and employment. *Management Science*, *2*(2), 159-177.

Holt, C. C., Modigliani, F., & Simon, H. A. (1955). A linear decision rule for production and employment scheduling. *Management Science*, *2*(1), 1-30.

Jamalnia, A., Yang, J. B., Feili, A., Xu, D. L., & Jamali, G. (2019). Aggregate production planning under uncertainty: a comprehensive literature survey and future research directions. *The International Journal of Advanced Manufacturing Technology*, *102*, 159-181.

Jamalnia, A., Yang, J. B., Xu, D. L., & Feili, A. (2017). Novel decision model based on mixed chase and level strategy for aggregate production planning under uncertainty: Case study in beverage industry. *Computers & Industrial Engineering*, *114*, 54-68.

Jones, C. H. (1967). Parametric production planning. *Management Science*, *13*(11), 843-866.

Leung, S. C., Wu, Y., & Lai, K. K. (2006). A stochastic programming approach for multi-site aggregate production planning. *Journal of the Operational Research Society*, *57*(2), 123-132.

McRary, J. (2023, October 11). *Bridging the labor gap: Warehouse automation's role in addressing labor shortages*. BastianSolutions. https://www.bastiansolutions.com/blog/bridging-the-labor-gap-warehouse-automations-role-in-addressing-labor-shortages/

Mirzapour Al-E-Hashem, S. M. J., Baboli, A., & Sazvar, Z. (2013). A stochastic aggregate production planning model in a green supply chain: Considering flexible lead times, nonlinear purchase and shortage cost functions. *European journal of operational research*, *230*(1), 26-41.

Montgomery, D. C., & Runger, G. C. (2010). Applied statistics and probability for engineers. John wiley & sons.

Myers, J. (2021, May 14). *Workforce Management in Warehousing and Distribution.* Zebra Technologies. https://www.zebra.com/us/en/blog/posts/2021/how-warehouse-operators-can-better-manage-labor.html

Nahmias, S., & Olsen, T. L. (2015). Production and operations analysis. Waveland Press.

Nam, S. J., & Logendran, R. (1992). Aggregate production planning—a survey of models and methodologies. *European Journal of Operational Research*, *61*(3), 255-272.

Posner, M. E., & Szwarc, W. (1983). A transportation type aggregate production model with backordering. *Management Science*, *29*(2), 188-199.

Romaine, E. (2021, July 27). *Warehouse labor trends in 2023: Seismic changes, shortages, retention challenges & more*. Conveyco. https://www.conveyco.com/blog/warehouse-labor-challenges-trends/

Schild, A. (1959). On inventory, production and employment scheduling. *Management Science*, *5*(2), 157-168.

Singhal, K., & Adlakha, V. (1989). Cost and shortage trade-offs in aggregate production planning. *Decision Sciences*, *20*(1), 158-165.

Sunol, H. (2023, September 22). *Distribution Center: 6 Best practices to unlock efficiency*. https://articles.cyzerg.com/distribution-center-best-practices-to-unlock-efficiency-and-cost-savings

Taubert, W. H. (1968). A search decision rule for the aggregate scheduling problem. *Management Science*, *14*(6), B-343.

Thompson, S. D., Watanabe, D. T., & Davis, W. J. (1993). A comparative study of aggregate production planning strategies under conditions of uncertainty and cyclic product demands. *The International Journal Of Production Research*, *31*(8), 1957-1979.

Törn, J. (2022, February 15). *Four common challenges in warehouse order picking (and how to address them)*. TAWI. https://www.tawi.com/insights/four-common-challenges-in-warehouse-order-picking-and-how-to-address-them/

Van de Panne, C., & Bosje, P. (1962). Sensitivity analysis of cost coefficient estimates: The case of linear decision rules for employment and production. *Management Science*, *9*(1), 82-107.

Vergin, R. C. (1966). Scheduling maintenance and determining crew size for stochastically failing equipment. *Management Science*, *13*(2), B-52.

World Trade Organization. (2023). World Trade Statistical Review 2023. https://www.wto.org/english/res_e/publications_e/wtsr_2023_e.htm

Appendices

Appendix A – Regression methods

Rank	Model	Rank	
1	XGBRegressor	**21**	LassoCV
2	LGBMRegressor	**22**	Ridge
3	GradientBoostingRegressor	**23**	BayesianRidge
4	BaggingRegressor	**24**	RidgeCV
5	RandomForestRegressor	**25**	SGDRegressor
6	KNeighborsRegressor	**26**	ElasticNetCV
7	ExtraTreesRegressor	**27**	Lasso
8	HistGradientBoostingRegressor	**28**	LassoLars
9	AdaBoostRegressor	**29**	OrthogonalMatchingPursuit
10	DecisionTreeRegressor	**30**	ElasticNet
11	ExtraTreeRegressor	**31**	PassiveAggressiveRegressor
12	NuSVR	**32**	TweedieRegressor
13	SVR	**33**	PoissonRegressor
14	LassoLarsCV	**34**	MLPRegressor
15	TransformedTargetRegressor	**35**	LinearSVR
16	OrthogonalMatchingPursuitCV	**36**	HuberRegressor
17	LinearRegression	**37**	DummyRegressor
18	LassoLarsIC	**38**	RANSACRegressor
19	LarsCV	**39**	KernelRidge
20	Lars	**40**	GaussianProcessRegressor

Appendix B – Explanatory variables

#	Variable name	#	
1	SNIS_NEW	**16**	WEEKDAY_Friday
2	SNIS_TOTAL	**17**	Ident_W-CRASH_ERROR_ACTIVITY
3	active_pickers	**18**	Ident_W-GAPC-R_ERROR_SIM
4	PICK_TOTAL	**19**	Ident_W-CRASH_ERROR_SIM
5	Ident_HEIGHT-ERROR_total_errors	**20**	Ident_W-UNDEF-E_ERROR_SIM
6	Ident_W-CRASH_total_errors	**21**	TN
7	Ident_W-RUEK-P_ERROR_SIM	**22**	Ident_HEIGHT-ERROR_ACTIVITY
8	Ident_scanner_reading_total_errors	**23**	Ident_HEIGHT-ERROR_SIM
9	Ident_W-SC-Q_total_errors	**24**	Ident_LOC-EMPTY_ACTIVITIY
10	Ident_W-FUSE_SIM	**25**	Ident_BVOL-processing_SIM
11	Ident_W-GAPC_total_errors	**26**	Ident_W-SE-DOOR_ERROR_ACTIVITY'
12	WEEKDAY_Monday	**27**	Ident_W-CRASH_duration
13	WEEKDAY_Wednesday	**28**	Ident_CRANE-NOT-AVAIL_total_errors
14	WEEKDAY_Tuesday	**29**	Ident_W-TO-RUN_ERROR_ACTIVITY
15	WEEKDAY_Thursday	**30**	Ident_W-SE-DOOR_total_errors

Appendix C – Performance of other departments

Approach 2

Department	Workforce size	Overtime hours	Costs
Production	153 hrs	3.4 hrs	€4322
Cable	210 hrs	3.7 hrs	€5921
Expedition	147 hrs	17.9 hrs	€4318

Approach 3

Department	Workforce size	Overtime hours	Costs
Production	154 hrs	3.2 hrs	€4348
Cable	211 hrs	3.4 hrs	€5946
Expedition	147 hrs	17.8 hrs	€4317

Approach 4

Department	Workforce size	Overtime hours	Costs
Production	151 hrs	3.1 hrs	€4263
Cable	208 hrs	3.4 hrs	€5862
Expedition	146 hrs	17.6 hrs	€4286

Policy 1 (30%)

Department	Workforce size	Overtime hours	Costs
Production	152 hrs	0.6 hrs	€4262
Cable	211 hrs	0.9 hrs	€5918
Expedition	147 hrs	17.1 hrs	€4309

Policy 2

Department	Workforce size	Overtime hours	Costs
Production	148 hrs	0 hrs	€4144
Cable	206 hrs	0 hrs	€5768
Expedition	146 hrs	15.5 hrs	€4263