Abstract

In 2014 UNHCR introduced a centralized fleet management program named Global Fleet Management (GFM). Now, country offices request to lease vehicles from UNHCR. However, this vehicle supply process (VSP) has substantial room for improvement. This thesis focuses on improving the VSP of the GFM program. By means of a discrete event simulation model, we analyze two supply chain interventions. First, we consider inventory control policies that regulate the inventory position at each point in time. Second, we analyze a demand smoothing policy, which creates a more stable demand flow.

Results indicate that implementing a guaranteed service level in combination with time-varying base stocks yields substantial improvements with respect to lead times at a marginal increment in costs. This is an important finding for UNHCR, as this implies that realizing a much better service level does not need to be expensive. We also show that introducing a simple smoothing policy results in a lower level of workload variability at the hub.

Keywords: Fleet Management; Humanitarian Logistics; Inventory Control; Optimization; United Nations High Commissioner for Refugees

Supervisor Erasmus University:
Prof. Dr. Ir. Rommert Dekker

Supervisors INSEAD:
Prof. Dr. Luk N. van Wassenhove
Dr. Harwin de Vries
# Contents

1 Introduction 3

2 Background 4
   2.1 UNHCR 4
   2.2 Fleet Management 5
   2.3 Vehicle Supply Process 5
   2.4 Areas for Improvement 6

3 Literature Review 8

4 Methodology 11
   4.1 Simulation 11
   4.2 Performance Criteria 12
   4.3 Supply Chain Interventions 13

5 Results 20
   5.1 Data 20
   5.2 Parameters and Assumptions 21
   5.3 Interventions 23

6 Discussion 34

7 Conclusion 35

8 References 38

A Appendix 40
   A.1 Vehicle Model Coverage 40
   A.2 Confidence Intervals for Various Demand Distributions 40
   A.3 Vehicles Coming Off Lease Separated by Vehicle Model 41
   A.4 Sensitivity Analysis: Smoothened Demand versus The Benchmark 41
   A.5 Sensitivity Analysis: Demand Variation without Smoothing 42
## Acronyms

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AFMS</td>
<td>Asset and Fleet Management Section</td>
</tr>
<tr>
<td>CO</td>
<td>Country Office</td>
</tr>
<tr>
<td>CODP</td>
<td>Customer Order Decoupling Point</td>
</tr>
<tr>
<td>FM</td>
<td>Fleet Management</td>
</tr>
<tr>
<td>FMS</td>
<td>Fleet Management System</td>
</tr>
<tr>
<td>FO</td>
<td>Field Office</td>
</tr>
<tr>
<td>GFM</td>
<td>Global Fleet Management</td>
</tr>
<tr>
<td>IHO</td>
<td>International Humanitarian Organization</td>
</tr>
<tr>
<td>NGO</td>
<td>Non-Governmental Organization</td>
</tr>
<tr>
<td>SC</td>
<td>Supply Chain</td>
</tr>
<tr>
<td>SOP</td>
<td>Standard Operating Procedure</td>
</tr>
<tr>
<td>SS</td>
<td>Safety Stock</td>
</tr>
<tr>
<td>UN</td>
<td>United Nations</td>
</tr>
<tr>
<td>UNHCR</td>
<td>United Nations High Commissioner for Refugees</td>
</tr>
<tr>
<td>VSP</td>
<td>Vehicle Supply Process</td>
</tr>
</tbody>
</table>
1 Introduction

UN agencies and international humanitarian organizations (IHOs) work to alleviate suffering throughout the world. Aid is delivered almost regardless of security situation, operating context, and location. However, without vehicle fleets, aid workers cannot provide humanitarian assistance. Either by delivering relief items or transporting beneficiaries, the vehicle fleet of humanitarian organizations supports and is the backbone of program implementation. In many NGOs (Non-Governmental Organizations) and UN agencies, vehicle fleets represent the second largest operating cost after staff expenses (Disparte, 2007). Saving fleet costs can freeing substantial resources, which can be spent on additional help for beneficiaries. Additionally, fleet has a major direct impact on beneficiaries. If fleet is not properly managed, offering timely help becomes challenging if not impossible.

UNHCR (United Nations High Commissioner for Refugees) is the UN organization mandated to lead and coordinate international action for the worldwide protection of refugees and the resolution of refugee problems. The organization works in over 100 countries through country offices (COs) on oftentimes large-scale and expensive missions. UNHCR uses vehicles to support field operations. Prior to 2014, fleet management in the organization was decentralized. COs were in charge of all aspects by buying, maintaining, and disposing of vehicles locally. However, in 2014 UNHCR introduced a centralized fleet management program named Global Fleet Management (GFM), which was to be handled by the Asset and Fleet Management Section (AFMS) of UNHCR, headquartered in Budapest. Now, COs place orders for and request vehicles from AFMS, who buys vehicles directly from manufacturers and ships them to UNHCR’s fleet hub in Thailand. There, vehicles are prepared for use by adding extra, country-specific, equipment. Vehicles are then shipped to COs. The latter pay an annual rental fee to AFMS for the vehicle. Finally, after the rental period, the vehicles are auctioned.

Though the implementation of the centralized system is perceived to be successful in many aspects (Kunz et al., 2015), there is substantial room for improvement. This thesis focuses on the vehicle supply process (VSP) of the GFM program. The VSP involves all physical, administrative, and financial activities needed to fulfill a CO’s vehicle request. The process of delivering vehicles has grown haphazardly and become cumbersome. It is subject to a lack of formal policies and procedures, overload, many reiterations, and lack of clear communication and responsibility between the stakeholders. As a consequence, lead times are long and highly fluctuating. The issues put the process under severe tension and the risk of major trouble is considerable. One example is the process coordination and overview. Presently there is only one person with a holistic understanding and overview of the VSP. The knowledge and experience are tacit and will be lost if the person for some reason would leave. The lack of standardization makes it difficult to transfer knowledge and procedures by making the inherit experience explicit. AFMS does not have any formal documentation, training, or SOPs (Standard Operating Procedures) of the VSP to provide to newly hired staff.
In addition to standardization, optimizing the VSP is needed. AFMS places purchase orders to manufacturers mainly after the CO places a vehicle request. The result is long lead times, decreasing satisfaction in the COs who expect timely delivery of vehicles. Keeping stock at the hub can decrease lead times, thus improving the satisfaction in COs.

Further, the demand pattern includes big spikes and long periods of almost no demand. Fluctuating demand can result in inefficiencies for manufacturers and for the hub. Manufacturers have a capacity constraint on production, i.e., can only produce a set number of vehicles per time unit. The hub has a preparation constraint, i.e., can only prepare a set number of vehicles for shipment per time unit. A relatively stable workload will ease the planning and coordination for the hub and its employees. Establishing a smoothened and stabilized demand with formal policies and procedures would be beneficial.

For solving the identified problems, this thesis studies two supply chain (SC) interventions. The interventions were identified during in-depth interviews with AFMS staff. First, we consider inventory control policies that regulate the inventory position at each point in time. Second, we analyze a demand smoothing policy, which creates a more stable demand flow.

This thesis analyzes the two SC interventions by means of a discrete event simulation model. Based on real data from AFMS, our model simulates the entire VSP along with the impact of the SC interventions on the process. The performance criteria include costs, service\(^1\), and/or regularity of flows.

The remainder of the thesis is organized as follows. In Section 2 we provide background information on the problem, while in Section 3 we discuss the relevant literature and place the thesis research within this context. Next, in Section 4 we present the SC interventions we are analyzing and the simulation model we use. Finally, in Section 5 we present our results. We conclude by giving a brief summary and our ideas for improvement in Section 6 and 7.

2 Background

In Section 2.1 we explain the context in which UNHCR operates. In Section 2.2 we describe the fleet management system (FMS) UNHCR has used since the introduction of GFM. Then, in Section 2.3 we describe the VSP in detail. In Section 2.4 we state areas for improvement and describe the general situation to give an idea of the magnitude of the current problems. More details are provided in Section 4.

2.1 UNHCR

UNHCR has approximately 6000 4x4 vehicles to support field operations. The vehicles are mainly used for long term missions to provide basic goods and services to refugees,

\(^1\)Service is defined as reducing delay, i.e., the time between the CO submits a vehicle request and the vehicle starts being prepared for shipment.
for scheduled trips to camps, and administrative tasks. Vehicles are also used in times of emergency. Depending on security situation and operating context, UNHCR is using light and heavy duty vehicles, both armored and non-armored.

2.2 Fleet Management

Managing an enormous vehicle fleet can either be done centralized or decentralized (Pedraza-Martinez and Van Wassenhove, 2012). In a decentralized FMS the COs procure, operate, and dispose of their own vehicles. In a centralized FMS the headquarter procures the vehicles and give guidelines for their utilization, maintenance, and repair. Headquarters also manage the proper disposal of the vehicles at their end of life. Note, in both systems COs still manage the fleet. AFMS procures and disposes of the fleet, i.e., it manages the assets but not their use.

In January 2014, UNHCR switched from a decentralized to a centralized FMS. This change was followed by good results as shown by Kunz et al. (2015). After one year UNHCR’s fleet size decreased by 11 percent, and the average vehicle age fell 21 percent, dropping below five years. Procurement costs also decreased by 21 percent, which is expected to produce yearly savings of US $5 million. Another major benefit of centralization is increased fleet standardization, an important first step towards implementation of an organization-wide set of best practices. In the first year, the number of models represented in the fleet decreased by 34 percent, and the number of suppliers fell by 43 percent. The FMS has continued to show good results as the average fleet age reached 3.72 years in 2016, compared to 5.65 years in 2012. In the same period, the percentage of vehicles exceeding the five-year limit decreased from 47% to 25% (Delagarde et al., 2016).

2.3 Vehicle Supply Process

Figure 1 depicts the stakeholders involved in the VSP, as well as the vehicle and order flows between them. AFMS issues a PO to the manufacturer based on incoming requests, the upstream pipeline\(^2\), and the available budget.

\(^2\)AFMS defines the upstream pipeline as all the unallocated vehicles in 1) the hub, 2) 1st-leg transport, 3) production.
All vehicles are purchased and ordered centrally by AFMS, which has its headquarters in Budapest. Then, vehicles are shipped by sea from manufacturer to one of the UNHCR vehicle hubs, either in Laem Chabang, Thailand; Brussels, Belgium; or Dubai, United Arab Emirates. All shipments are handled by freight forwarders. Light vehicles are shipped to Thailand while armored vehicles are sent to Brussels. A security stock of vehicles for emergency situations is kept in Dubai. At the hubs, the vehicles are prepared for use by adding extra equipment, e.g., vehicle tracking system, bull-bar, windshield, winch, and jerry-cans. Next, the vehicles are shipped to the requesting CO. The latter pays an annual rental fee to AFMS, which, in exchange, provides the vehicle and additional services such as training, insurance, and disposal. AFMS currently starts the lease period when the vehicle is ready for shipment from the hub, meaning that the vehicles start aging from this moment. After five years of use, vehicles are auctioned by AFMS locally. Figure 2 shows the journey of a vehicle from Japan through Thailand to Uganda.

Figure 2: The Journey of a Vehicle from Japan through Thailand to Uganda.

2.4 Areas for Improvement

Though the present FMS improved upon the previous in many aspects, substantial improvement can still be achieved. First, formal procurement policies are presently absent. As a result, the current FMS is largely based on intuition. In one part, it resembles a pull system. Each vehicle request has to 1) wait until the vehicles are manufactured and shipped to the hub, 2) risk that the requested vehicles are assigned to different (more urgent) operations, 3) wait until the vehicle arrives. In another part, the FMS resembles a push system. To enable a swift response to emergencies, UNHCR keeps a safety stock (SS) of vehicles and tops up the upstream pipeline of vehicles ordered to the manufacturer (using intuition and information about potential crises). The latter process is unsystem-
atic and depends on the skill set and experience of one or two persons, i.e., it would be
difficult to duplicate for persons trying to re-engineer the process. The well-known cus-
tomer order decoupling point (CODP) is situated at the hub. The CODP separates the
part of the value stream that is order-driven (downstream of the CODP), from the part
that is forecast-driven (upstream of the CODP) (Romme and Hoekstra, 1992).

Second, demand forecasting is presently lacking. Vehicles are leased for a 5-year period.
Hence, it can be anticipated when current vehicles should be replaced.\(^3\) This especially
holds if the hub preparation time for a vehicle and the shipping time to a CO can be
reasonably predicted constant. Demand uncertainty is mainly due to new or expanding
missions and ending or shrinking programs. The first induce additional, unforeseen orders.
The latter results in some or all of the vehicles not being replaced after 5 years.

Third, UNHCR is experiencing high demand volatility. Figure 3 shows the sudden and
large quantity orders that increase the complexity of inventory control and forecasting.
This can be avoided by smoothing the demand caused by existing vehicles coming off
lease. For example, vehicles can come off lease one month earlier or later. Such demand
smoothing does currently not take place.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure3.png}
\caption{Number of Vehicles\(^4\) Coming Off Lease per Month from 2019 until 2023 Aggre-
gated from Weeks to Months based on Data from AFMS.}
\end{figure}

Complaints about lead times and uncertainty are rising, and has been evidenced in
a recent survey conducted by INSEAD Humanitarian Research Group (Van Rijn et al.,
2017). Proposing and analyzing system changes is therefore key.

\(^3\)Given that field operations comply with the 5-year replacement policy.
\(^4\)Toyota Land Cruiser 76 and Toyota Prado both left and right hand drive.
3 Literature Review

In this section we discuss the relevant literature. First, we present the challenges of implementing Operations Research in a humanitarian versus a commercial world. Next, we discuss the obstacles in implementing proper FMSs. Last, we discuss the relevant literature on inventory control.

**Humanitarian versus Commercial**

SC optimization is an essential topic in both the commercial and humanitarian sector. However, the areas are generally not one to one interchangeable. Especially in times of emergency, effective and efficient operations management may decrease human suffering and economic damage. Time plays a critical role and directly impacts the survival rate in affected areas. Time, or the lack of it, also makes logistics planning and SC management in a humanitarian setting more complex than conventional distribution problems (Nahleh et al., 2013).

Many IHOs keep stock of vehicles in order to increase their levels of preparedness against sudden disasters. However, similar to commercial SCs, high inventory holding costs are a financial burden. Especially for IHOs because of their limited funds and operating resources. Designing effective inventory systems for a humanitarian organization is therefore of great importance (Nikbaksh et al., 2011).

A vast majority of humanitarian research is dedicated to global disaster relief efforts. According to Manopiniwes and Irohara (2014), research related to disaster relief logistics can be divided into three groups: facility locations, distribution models, and inventory models. In their review paper, they give a broad overview of this challenging research area with emphasis on the corresponding optimization problems.

This thesis focuses on another core task for IHOs. Next to emergencies, IHOs use their vehicles for long term missions, providing basic goods and services to vulnerable groups. These programs consist of scheduled missions and corresponding administrative tasks. The focus during emergencies versus development is more on speed of vehicle delivery in emergencies and more on costs effectiveness for development. COs requesting vehicles are often not served from inventory as stock keeping is expensive.

Additional challenges are created by the peculiar decision-making process of humanitarian organizations. Qualitative heuristics are often used to establish fleet sizes and compositions. Eftekhari et al. (2014) emphasize the benefits that can be gained from simple validated tools that aggregate demand information to guide procurement and allocation decisions. While these problems have been studied extensively, De Vries and Van Wassenhove (2017) explain why the adoption of optimization approaches in humanitarian field operations are absent.

**Fleet Management**

Logistics is the second-largest operating expense of an IHO and much of the logistics-
related costs are associated with fleet management (FM) (Eftekhari et al., 2014). Therefore, cost savings and performance improvements expected from optimized FM processes are brought to the attention of IHOs. The IHOs that operate a centrally leased fleet for their development missions need a functional FMS. Multiple factors increase the complexity of FM in humanitarian versus commercial settings. According to Pedraza Martinez et al. (2010), it is challenging to calculate the optimal fleet size due to the unusual environment in which IHOs operate. Kovács and Spens (2007) illustrate the negative effects that lack of infrastructure, reliable roads, and security constraints might have on vehicle usage for IHOs. These and other factors increase the complexity of proper FM in humanitarian versus commercial settings.

Martinez et al. (2011) present a theoretical framework showing the potential factors affecting field vehicle fleet management objectives in humanitarian programs. Their fleet management objectives include speed of vehicle delivery, fleet availability, and cost effectiveness. In their paper they consider external factors such as donors, demand uncertainty and operating conditions. They describe how external factors can affect FM objectives directly but mostly seem to affect IHOs at the organization and fleet level. Fleet level and fleet management objectives are simultaneously affected by internal factors such as mandate, organizational structure, and funding structures.

In literature, most decisions regarding FM are based on stochastic vehicle demand. In this thesis we base vehicle demand on vehicles coming off lease. A demand that can be reasonably predicted constant. This part of the research within FM differs from our research and is not used.

Inventory Control
Forecasting demand in humanitarian operations has been found to be challenging due to many types of uncertainties. Uncertainty in demand is mainly due to new or expanding missions and due to ending or shrinking missions. Additionally, requests can also be canceled last-minute. These uncertainties can be mitigated to a large extent by mechanisms such as SS (Nahleh et al., 2013). The objective of a SS is to absorb fluctuations in demand and supply, e.g., unexpected demands and shortages in supplies, and to buffer errors in the stock records that may occur during manufacturing. There are various studies on determining the desired level of SS. Many SS levels are based on \((R, Q)\) or \((s, S)\) inventory control policies. In this thesis we look at a simple order-up-to policy. However, the initial demand in our situation shows an erratic pattern, therefore we use a dynamic SS policy instead of the extensively covered static SS policy.

Optimal inventory control decisions depend on stock level, demand information (taking into account forecasted quantity), lead time variation, lead time, inventory holding costs, shortage cost, and cost of ordering (cf. Onyango (2016)). In our case, the number of vehicles coming off lease form the basis of new demand and is known. However, the desired number of new vehicles often deviates from the number of vehicles coming off lease as previously discussed. Manufacturer lead times can be reasonably predicted constant
and cause little disturbance from the supply side. Shortage costs are not immediately quantifiable. Stock-outs result in longer lead times and leads to dissatisfaction in field operations, which is not directly measurable. This shows that not everything is applicable to our case.

Our trade-off is all about keeping the COs satisfied and costs low at the same time. The needs of the COs must thus be balanced with the goal of AFMS. The balance should be found in an inventory management model where at the ideal time the right assets are at the ideal place. The inventory level investment should be in line with the expected customer service level. Inventory is a critical asset in any organization (Barnes, 2008). To maintain optimal inventory levels, a robust system is needed to accurately track levels of inventory control. Further, there is a need for internal processes which monitors demand patterns, maintains inventory counts accurately, tracks inventory performance, and ensures adherence to supplier commitments.

A general approach of including customer service level was first suggested by Kimball in 1955. This guaranteed-service approach (GSA) has been well developed over the years for single- and multi-echelon inventory system optimization. The assumptions with regard to this approach are most often different than ours. The GSA in literature is mostly based on a continuous review serial inventory system. Sobel (2004) show that a periodic review system is more difficult to characterize and compute, however we can not work with this as our replenishment goods are only shipped periodically. Based on a comparison study of Li and Wu (2018) the GSA model is most efficient in solving complicated inventory problem in terms of computation time, and costs. It is shown that GSA is especially effective when warehouse processing time is long and retailer service level high. This is consistent with our case. In our research we follow a key assumption of the GSA, which assumes a maximum reasonable lead time demand level for lead time demand of the customer. As a consequence, the guaranteed service can be accomplished with a finite stock of inventory. Graves and Willems (2003) state that this is reasonable if the replenishment time is predictable and deterministic which is the case in our research. Graves and Willems (2000) provide various generalizations of the GSA and show how to minimize total cost for safety stock, but finding the optimal service level can be time expensive and complex which is exactly the opposite of what we aim for in this research.

Managing inventory with non-stationary, stochastic demand is found to be challenging. The problem is often separated into strategic and tactical components. If and where to locate inventories, what levels of service to provide, and whether these choices should change over time are the strategic decisions. Tactical decisions mainly consist on how to calculate inventory targets that change with demand over time. Neale and Willems (2009) present a guaranteed service modeling framework that help companies make these strategical and tactical decisions. Their model is based on the work of Graves and Willems (2000) where they extend it to the case of non-stationary demand. In their work they focus on one time period review periods which is different from our work. They also place much emphasis on multistage systems.
Kalchschmidt et al. (2003) explain that in many industries demand is becoming more and more variable and uncertain. As previously discussed, our demand shows an erratic pattern. Literature has devoted major attention to forecasting erratic demand and the development of stock policies. There are some well-known forecasting techniques, but in all studies, it is shown that demand smoothing tends to be substantially more beneficial than forecasting.

Our research is positioned relative to previous work on stochastic, non-stationary inventory problems in a humanitarian setting. Relative to stationary demand inventory models, much less work on non-stationary demand exists. Non-stationarity typically complicates the analysis and limits the results that can be obtained. The vast majority of humanitarian research is dedicated to global disaster relief efforts while this research focuses on development.

4 Methodology

To investigate the effects of our interventions on our VSP problem, we develop a discrete-event simulation model. In Section 4.1 we describe the details and structure of the simulation model. Then, in Section 4.2 we describe how we evaluate the performance of the proposed interventions. In Section 4.3 we explain the different inventory control policies together with the technique for smoothing.

4.1 Simulation

We chose to use a simulation model because the performance of (a combination of) our supply chain interventions cannot (trivially) be expressed analytically. In this research we use a finite horizon together with non-stationary demand which also suggest to use a simulation model. Our model simulates the whole process from ordering a vehicle to the delivery of that vehicle and every step in between. With our model, the effects of the different inventory control policies are evaluated. The same is done for the smoothing policy.

The simulation model mimics the journey of a vehicle that goes through different phases. We describe the following phases: the vehicle is being manufactured, the vehicle is being shipped from the manufacturer to the hub, the vehicle is at the hub, the vehicle is being prepared for shipment, the vehicle is being shipped to the designated country, the vehicle is in service, and finally the vehicle is disposed of.

The CO request for a certain vehicle can occur at three different moments within this journey: 1) before the PO for this vehicle is submitted to the manufacturer, 2) while the vehicle is in the pipeline (i.e., the PO is placed but the vehicle has not yet arrived at the hub yet), 3) while the vehicle is on stock at the hub. Figure 4 illustrates these three options.
To simulate the vehicle journey we use a discrete event simulation model. All blocks in Figure 4 are events. The end-of-lease period that triggers new requests is another event. These events will be simulated for two cycles of five years, using time steps of one week, resulting in 520 time periods. The entities in our model are the vehicles and the capacity constrained is found at the hub for the preparation of the vehicles for shipment. The number of vehicles that is requested is determined by drawing from our derived demand distribution and is based on the number of vehicles coming off lease. We assume that there are no failures in the field that trigger new requests. Combinations of the different policies will be used as input for our simulation model, whereas their performance with respect to the evaluation criteria described in the next section will be the output. As our starting situation we use the current situation of UNHCR, where the number of vehicles in each country and their age is known. The number of vehicles in the upstream pipeline is not known and can thus not be used in our starting position. To account for this we start recording performance after 60 weeks so that vehicles have accumulated at the hub.

4.2 Performance Criteria

While simulating the combination of the different interventions we look at three performance measures. **First,** we look at the average inventory level at the hub to indicate holding costs. We calculate the average number of vehicles on stock per week. If holding costs are known we can easily calculate the actual costs of keeping stock. **Second,** we look at delay. In our research, delay is defined as the time between the CO submits a vehicle request and the vehicle starts being prepared for shipment. We leave all processes after the preparation phase out of the equation as this is outside the area of influence of AFMS. This measure is used to show the possible benefits of having vehicles in stock waiting to be requested, compared to make-to-order, where the vehicles still need to be procured, manufactured, and shipped to the hub. To show the difference we use the average delay and to indicate if the supply chain is stable enough to give a more or less guaranteed delay indication to the CO we show the standard deviation of the delay. **Third,** we evaluate the capacity at the hub based on the number of accessorized vehicles per month. Manufactur-
ing and 1st-leg transport times are independent of the order size. However, accessorizing vehicles at the hub is not. It is realistic to assume that the hub has a constant processing speed, high variability thus leads to bigger delays. If this rate is known, the effect of smoothing can easily be expressed in terms of lead times. However, as we can not derive this rate from the data we simplify the analysis by assuming that the hub can prepare any number of vehicles for shipment independent of previous months. We evaluate the workload variability at the hub based on the standard deviation of the number of vehicles being accessorized per month.

4.3 Supply Chain Interventions

The results of our simulation model are evaluated for a combination of different policies. First, we will describe the intervention of introducing systematic inventory control on the basis of forecasted demand. Second, we introduce the intervention of smoothing demand.

Inventory Control

Until a request is made, actual demand is unknown. We can thus not start documentation finalization, vehicle preparation, and or shipment. However, we can anticipate future demand by keeping stock of vehicles in use. We have to take the lead times for all phases before the hub into account (i.e. time to manufacture and 1st-leg transport time).

In our research we evaluate three inventory control strategies based on Axsäter (2015). Reflecting current practice, the first strategy entails placing monthly orders at the supplier so as to bring the upstream pipeline up to a constant base stock level $S$. However, as we work with different vehicle models, all varying in quantity, we use different levels of $S$ per vehicle model. To be able to visualize and analyze this strategy we define our vehicle model base stock level as $S_m$. We base this level on a percentage of the total number of all vehicles in use of that specific model, $V_m$. We define the percentage as $x$ and to evaluate the strategy we test multiple levels of $x$. We calculate our model specific base stock level as follows:

$$S_m = x \cdot V_m$$  \hspace{1cm} (1)

The second strategy we evaluate is make-to-order, where we keep no stock and only procure after COs request. The benefits will be no stock and limited storage holding costs, but the disadvantage will be the longest delay of the strategies evaluated.

Third, we test a guaranteed service level inventory control policy with time-varying base stocks. We use the fill rate ($FR$), defined as the fraction of demand that can be satisfied immediately from stock on hand, as our service level criterion. In the next paragraph we discuss the policy in detail.

**Guaranteed Service Level Inventory Control Policy**

We refer to the literature, specifically Axsäter (2015), for a theoretic approach to optimize
our inventory control. We use a periodic review \((S-1, S)\) policy where \(S\) is dynamic. The expected demand during one unit of time is equal to the number of vehicles coming off lease and is defined as \(\mu_t\) for every period \(t\). The analysis below is built on the work of Axsäter (2015) but adds value to this work due to the introduction of \(\mu_t\) as the expected demand per period. Due to time-varying demand we choose \(S\) to be time-dependent and therefore use the \(S_t\) notation. To determine our time-varying base stock level, we introduce some additional notations. The time it takes from placing a PO at the manufacturer until the vehicles physically arrive at the hub is our lead time, defined as \(L\). We assume a constant lead time. We define the inventory position at time \(t\) as \(IP_t\). The inventory position is defined as the stock on hand + outstanding orders - backorders. The physical inventory at the hub at time \(t\) is called the inventory level and is defined as \(IL_t\). The inventory level is defined as the stock on hand - backorders. \(D(t, t+\tau)\) is now defined as the stochastic demand in the interval \((t, t+\tau]\), while \(D(t)\) is defined as the stochastic demand in period \(t\). Finally, we define \(T\) as our review period, i.e., the time interval between reviews. In our research we place a PO at Toyota once every month, which is defined as our review period.

Consider Figure 5 and let \(t\) be an arbitrary review time. The orders can take place at times \(..., t, t+T, ...\). A possible order at time \(t\) is delivered at time \(t+L\). The next possible delivery time is evidently \(t+L+T\). This research shows the trade-off between costs and lead time. We use a target \(FR\) as a tool to evaluate this trade-off. Setting a high target \(FR\) will lead to inventory control policies with high stock levels and thus high costs, but also shorter lead times. In our experiments we will vary the target \(FR\) to investigate the trade-off between costs and lead times. When determining the \(FR\) we consider the interval between these two times \((t+L, t+L+T]\). It is easy to determine the expected demand in the considered interval once demand distributions are known. Furthermore, it is clear that the demand that can be delivered from stock on hand in this interval depends only on the inventory level after the possible delivery at time \(t+L\) and the total stochastic demand during the interval.

![Figure 5: Considered Time Epochs.](image)

We first derive the distribution of the inventory level at the two times \(t+L\) and \(t+L+T\), i.e., the start and the end of the considered interval. At time \(t+L\) we consider the inventory level after a possible delivery, but right before time \(t+L+T\), a possible delivery has not taken place. Orders that have been triggered in the interval \((t, t+L]\) have not yet reached the inventory due to the lead-time. Therefore, the inventory level \(IL'\) at time \(t+L\) after a possible delivery can be obtained as
\( IL' = IL(t + L) = IP(t) - D(t, t + L) \) (2)

And the inventory level \( IL'' \) at time \( t + L + T \) can be obtained as

\[ IL'' = IL(t + L + T) = IP(t) - D(t, t + L + T) = IP(t + T) - D(t + T, t + L + T) \] (3)

Consider now the interval between the two times \( t + L \) and \( t + L + T \). The expected demand between two review periods is obviously \( \sum_{i=t+L}^{t+L+T} \mu_i \).

Consider then the part of the demand in the considered interval between \( t + L \) and \( t + L + T \) that cannot be met from stock on hand. This demand is backordered so it can be determined as \( E(IL'') - E(IL') \), we define \( E(IL) \) as the expected negative inventory level. This means that we can determine the \( FR \) as

\[ FR = 1 - \frac{E(IL(t + T + L)) - E(IL(t + L))}{E(\text{review period demand})} = 1 - \frac{E(IL'') - E(IL')}{\sum_{i=t+L}^{t+L+T} \mu_i} \] (4)

The first vehicles will come off lease early 2019. Hence, we presently lack empirical data to base the demand distribution on. Based on expert interviews we assume that demand follows the following discrete uniform distribution \( \left[ \left( 1 - \alpha \right) \cdot \mu_t, \left( 1 + \alpha \right) \cdot \mu_t \right] \), where \( \alpha \) is determined by experts and the rounding is done for discretization purposes. These demand distributions are independent.

We express the probability mass function (PMF) of the discrete uniform demand distribution, for any \( j \in \left[ \left( 1 - \alpha \right) \cdot \mu_t, \left( 1 + \alpha \right) \cdot \mu_t \right] \), as

\[ P(D(t) = j) = \frac{1}{\left( 1 + \alpha \right) \cdot \mu_t - \left( 1 - \alpha \right) \cdot \mu_t + 1} \] (5)

We calculate our service level based on Equation (4). The base stock policy approach is visualized in Figure 6. \( S_t \) is the base stock level we want to optimize based on our \((S_t - 1, S_t)\) inventory control policy. In our numerical experiments, we use a bisection algorithm to find the smallest value of \( S_t \) that makes the \( FR \) in period \((t + L, t + L + T)\) exceed the target \( FR \). We know that demand during \((t, t + L + T)\) lies in the interval \( \left[ \sum_{i=t}^{t+L+T} \left( 1 - \alpha \right) \cdot \mu_i, \sum_{i=t}^{t+L+T} \left( 1 + \alpha \right) \cdot \mu_i \right] \), therefore the appropriate value of \( S_t \) lies in this interval. We use these bounds for our bisection algorithm.
Figure 6: Inventory Level at Different Points in Time.

To calculate the FR, we must first derive the probability distribution of $IL'$. We can then derive the probability distribution of $IL''$ by conditioning on $IL'$. It remains to determine $E(IL'')^-$ and $E(IL')^-$. 

$$E\left(IL(t+T+L)\right)^- = \sum_{k=-\infty}^{0} P\left(IL(t + L + T) = k\right) \cdot k$$  \hspace{1cm} (6)

$$E\left(IL(t+L)\right)^- = \sum_{j=-\infty}^{0} P\left(IL(t + L) = j\right) \cdot j$$  \hspace{1cm} (7)

As seen before, maximum demand during $(t, t+L+T]$ is $\sum_{i=t}^{t+L+T} [(1+\alpha) \cdot \mu_i]$, therefore the range of $k$ can start at the negative equivalent of this number. With the same reasoning we let $j$ start at $-\sum_{i=t}^{t+L} [(1+\alpha) \cdot \mu_i]$. To calculate the expected values we have to introduce one last piece of notation

$$D(t) = \text{stochastic demand at time } t$$
$$D(t, t + \tau) = \text{stochastic demand in the time interval } (t, t + \tau]$$
$$P\left(D(t) = j\right) = \text{probability of demand size } j \text{ at time } t$$
$$P\left(D(t, t + \tau) = j\right) = \text{probability that period } (t, t + \tau] \text{ gives the total demand } j$$

Because our demand distribution is non-stationary we now derive the distribution of $D(t, t + \tau)$ by obtaining the $j$-fold convolution of $P\left(D(t, t + \tau) = j\right)$, recursively as
\[ P(D(t, t + \tau) = j) = \sum_{i=0}^{j} P(D(t, t + \tau - 1) = i) \cdot P(D(t + \tau) = j - i) \]  

(8)

In more detail, Equation (8) can be written as

\[ P(D(t, t + 1) = j) = \sum_{i=0}^{j} P(D(t) = i) \cdot P(D(t + 1) = j - i) \]

\[ P(D(t, t + 2) = j) = \sum_{i=0}^{j} P(D(t, t + 1) = i) \cdot P(D(t + 2) = j - i) \]

\[ \vdots \]

\[ P(D(t, t + \tau) = j) = \sum_{i=0}^{j} P(D(t, t + \tau - 1) = i) \cdot P(D(t + \tau) = j - i) \]

We derive the distribution of \( IL' \) as follows

\[ P(IL' = j) = P(IL(t + L) = j) = P(D(t, t + L) = S_t - j) \]

\[ = \sum_{i=0}^{S_t - j} P(D(t, t + L - 1) = i) \cdot P(D(t + L) = S_t - j - i) \]  

(9)

Using (9) we derive the distribution of \( IL'' \) as

\[ P(IL'' = k) = P(IL(t + L + T) = k) \]

\[ = \sum_{j=-\infty}^{S_t} P(IL(t + L) = j) \cdot P(IL(t + L + T) = k | IL(t + L) = j) \]

\[ = \sum_{j=k}^{S_t} P(IL(t + L) = j) \cdot P(D(t + L, t + L + T) = j - k) \]  

(10)

Note that we can calculate the final term in Equation (10) in a similar way as done in Equation (9). Now, we have all the ingredients to calculate the \( FR \) from Equation (4).

**Normal Approximation**

The approach described above works but could be extensive. If the number of vehicles coming off lease is large, and the lead time \( L \) is long, the use of an approximation with the normal distribution is a faster alternative.

We know from the **central limit theorem** that, under general conditions, a sum of many independent random variables will have a distribution that is approximately normal.

We now consider a continuous normally distributed demand. We denote the mean and standard deviation of the demand per unit of time by \( \mu \) and \( \sigma \). Furthermore, the mean and standard deviation of the demand during the lead-time \( L \) are denoted \( \mu' = L \cdot \mu \) and \( \sigma' = \sqrt{L} \cdot \sigma \). We also consider the demand during the time \( L + T \) and the corresponding
mean and standard deviation \( \mu'' = (T + L) \cdot \mu \) and \( \sigma'' = \sqrt{(T + L) \cdot \sigma^2} \). The inventory position \( IP \) just after the arbitrary review time \( t \) as depicted in Figure 5 is now uniform on the continuous interval \( (S - 1, S] \). We again apply Equation (2), and obtain the inventory level distribution at time \( t + L \) after a possible delivery as

\[
F(x) = P(II' \leq x) = \sigma' \left[ G\left( \frac{S - 1 - x - \mu'}{\sigma'} \right) - G\left( \frac{S - x - \mu'}{\sigma'} \right) \right]
\]

With the loss function \( G(x) \) denoted as

\[
G(x) = \phi(x) - x(1 - \Phi(x))
\]

We obtain the \( FR \) by using the following formula

\[
E(II') = \int_{-\infty}^{0} F(x)dx = \int_{-\infty}^{0} \sigma' \left[ G\left( \frac{S - 1 - x - \mu'}{\sigma'} \right) - G\left( \frac{S - x - \mu'}{\sigma'} \right) \right] dx = (\sigma')^2 \left[ H\left( \frac{S - 1 - \mu'}{\sigma'} \right) - H\left( \frac{S - \mu'}{\sigma'} \right) \right]
\]

With \( H(x) \) defined as

\[
H(x) = \int_{x}^{\infty} G(v)dv = \frac{1}{2} \left[ (x^2 + 1)(1 - \Phi(x)) - x\phi(x) \right]
\]

\( E(II'') \) can be obtained in the same manner, replacing \( \mu' \) and \( \sigma' \) by \( \mu'' \) and \( \sigma'' \) in Equation (13).

**Smoothing**

We design a policy on how to smooth future demand. Spikes in demand should be smoothed to reduce risk and to make proper use of the above-mentioned policies. Demand spikes can gradually be lowered by allowing certain vehicles to come off lease in an earlier or later stage.

Vehicles are thus not replaced after a lease period of exactly five years. A policy with variable lease periods obviously has its disadvantages. First, vehicles coming off lease later will force the allocated vehicles to stay at the hub for an additional time period, which will increase inventory holding costs. Second, there is much administrative work involved in having variable lease periods.

Consider as example 50 vehicles that will come off lease in period \( n \) and 30 vehicles that will come off lease in period \( n + 1 \). We can decide to take the average of these two and let 10 vehicles come off lease one period later and thus replace 40 vehicles in both periods. Then the allocated vehicles will be on stock an additional period, but demand variability decreases. We can have an even bigger impact and let the smoothing go faster.
by averaging over more periods, but this would increase the complexity.

Figure 7 shows the effect of demand smoothing for another example in which we re-apply the heuristic. This example originally has two demand peaks in one life cycle, visualized by cycle 0. A life cycle is defined as the life time of an entity. In this example the cycles are 6 periods long. Cycle 0 represents the original demand and in this example the initial peaks are in period 1 and 4, both consisting of demand of 50 vehicles. The first cycle is the first iteration of our heuristic. If we continually re-apply the heuristic as discussed, we see the immediate effect. After only 3 cycle iterations there is already a decrease in standard deviation from over 23 to less than 0.6, showing the reduction that can be achieved by demand smoothing. We smooth demand over period \( t \) and \( t + 1 \) after which we move forward one period and re-apply the heuristic. \( \tilde{d}_t = \frac{d_t + d_{t+1}}{2} \) and \( \tilde{d}_{t+1} = \frac{d_{t+1} + d_{t+2}}{2} \). As a result the entities can thus be smoothened over multiple periods. One can also see that, in this example, there will not be much gain from more iterations.

Figure 7: The Effect of Demand Smoothing with the Number of Vehicles Coming Off Lease per Period per Cycle.

To capture the demand smoothing process we use the heuristic described below. We execute Equation (15) until Equation (20) for period \( t \) after which we move to period \( t + 1 \) and repeat this sequence. The sequence consists of five steps. First, we calculate the average number of vehicles coming off lease (\( VoL^*_t \)) over period \( t \) til period \( t + p \) with Equation (15). We define \( p \) as our smoothing horizon. A smoothing horizon of 1 is equal to our current situation of non-smoothing.

\[
VoL^*_t = \sum_{i=t}^{t+p-1} \frac{VoL^*_{old_i}}{p}
\]  

(15)

Second, after calculating the average number of vehicles coming off lease over period \( t \) til period \( t + p \) we have to check whether \( VoL^*_t \) is integer. If this is not the case we round
the number of vehicles down in some periods while we round the number of vehicles up in
the other periods. To save administrative work we round up the periods with originally the
largest number of vehicles. To decide in which periods to round up we first use Equation
(16) to calculate the number of periods that need to be rounded up.

\[ n \equiv \left( \sum_{i=t}^{t+p-1} \text{Vol}_i \right) \mod (p) \] (16)

Third, we define two sets of periods. We define a set of periods with the \( n \) highest
number of vehicles coming off lease, and we define a set of periods with the other periods.
Let us define set \( A_0 = \{ \text{Vol}_i \mid t \leq i < t + p \} \), the set with all periods and their
corresponding number of vehicles coming off lease. Next, we define set \( B_n \) as a subset of
\( A_0 \) with the \( n \) largest numbers. We write

\[ B_1 := \{ b \in A_0 : b \geq a \ \forall a \in A_0 \} \] (17)

and define \( A_1 := A_0 \setminus B_1 \). Then

\[ B_{i+1} := \{ b \in A_i : b \geq a \ \forall a \in A_i \} \] (18)

and define \( A_{i+1} := A_i \setminus B_{i+1} \). We do this inductively until we reach \( B_n \).

In our fourth step (Equation (19)) we adjust the new number of vehicles coming off
lease with the rounding we propose.

\[ \text{Vol}_i^{\text{new}} = \begin{cases} \lceil \text{Vol}_i^* \rceil, & \text{if } \text{Vol}_i^{\text{old}} \in B_n \\ \lfloor \text{Vol}_i^* \rfloor, & \text{if } \text{Vol}_i^{\text{old}} \in A_n \end{cases} \ \forall i = t, \ldots, t + p - 1 \] (19)

Finally, we have our new number of vehicles coming off lease at period \( t \) til period
\( t + p \). We set our new numbers now as our old numbers, as shown in Equation (20), and
we repeat this process as we iterate over time.

\[ \text{Vol}_i^{\text{old}} = \text{Vol}_i^{\text{new}} \ \forall i = t, \ldots, t + p - 1 \] (20)

5 Results

In Section 5.1 we describe how the data is gathered and what the obstacles are. Then,
in Section 5.2 we state the parameters that we use in this research and clearly state
the assumptions behind our numerical experiments. We give the results of the different
interventions in Section 5.3.

5.1 Data

The data for our research was retrieved through interviews with personnel from AFMS
and through analysis of procurement and shipping data. UNHCR keeps track of vehicle
orders from COs, orders from HQ to the manufacturer, the arrival of vehicles at the hub, and the shipment to COs. These data are used to identify the required parameters and distributions.

The retrieved data leaves out the emergencies and their corresponding increase in demand for vehicles. During a declared emergency, money plays almost no role and the focus is on delivering vehicles as soon as possible. Emergency requests have priority and, independent of our policies, delay is thus minimized for these requests. Hence, emergencies have an impact on the delay of other requests. Excluding emergency requests is a simplified assumption which might give a small bias in the calculations. However, vehicles used for emergencies have different funding and processing and only form a small portion of the total requests. Therefore, we expect only a small bias. So all demand for vehicles used in this study is assumed to be used for development missions. Next, we only focus on light vehicles, the hub in Thailand, and the Toyota manufacturer in Japan.

Some vehicle flows unnecessarily skew the data for various reasons. For example, countries are landlocked, customs are highly unpredictable, or procurement is decentralized. These special cases require extra personal attention and are thus not captured in our model. This was already identified by AFMS.

5.2 Parameters and Assumptions

Simulation

The stochastic part of our simulation is the uncertainty around the number of vehicles coming off lease. We define a discrete uniform distribution to indicate the uncertainty and to simulate the actual number of vehicles coming off lease. We are interested in our performance metrics stated in Section 4.2 and how they are influenced by the policies. Inherent to a simulation model, these outcomes differ per simulation run. We need to run the simulation model “enough” times to get an accurate estimate of the expected performance of the policies. We determine the required number of simulation runs to get a 95% confidence interval (CI)(Petty, 2012). CIs are written $[\text{LB}, \text{UB}]$, where $\text{LB}$ is defined as the lower bound and $\text{UB}$ is defined as the upper bound of the interval. We set the width of our target CI equal to $1/7$, which corresponds to a standard deviation of average delay, $\sigma^*$, of one day (= 1/7 week). We use 50 simulation runs to determine our sample standard deviation, $\hat{\sigma}_s$. We choose the maximum average delay over all our inventory control policies but without smoothing as our sample standard deviation. Together with the confidence interval of the normal approximation, where $z_c$ is the critical value for the normal distribution for confidence level $c$, the bounds look as follows:

$$\left[\mu - \frac{\hat{\sigma}_s}{\sqrt{n}} \cdot z_c , \mu + \frac{\hat{\sigma}_s}{\sqrt{n}} \cdot z_c \right]$$  \hspace{1cm} (21)

To make sure the width between the upper and lower bound is equal to our target standard deviation we rewrite Equation (21) and end up with our required number of
replications:
\[
2 \cdot \frac{\hat{\sigma}}{\sqrt{n}} \cdot z_c = \sigma^* \\
\left[2 \cdot \frac{\hat{\sigma}}{\sigma^*} \cdot z_c \right]^2 = n
\]  

(22)

In our experiment we find a sample standard deviation of the average delay equal to 0.43 weeks. If we use Equation (22) and a 95% confidence interval we end up with a required number of replications equal to \( \approx 150 \). Executed in Java this takes approximately 20 seconds per policy.

**Costs**

Most cost elements are constant. Both 1st and 2nd Leg Transport costs are constant per vehicle unit shipped and per destination. Contracts with the freight forwarders (FF) do not include a fixed cost per shipment. Thus, economies of scale cannot be gained, as the costs are independent of the size of the order. The same holds for procurement costs. The contract with Toyota in Japan does not include a fixed cost per production batch. The procurement costs are constant per type of vehicle.

The costs that really motivate this research are capital related holding costs. As it is often difficult to exactly define holding costs per vehicle per day we decided to evaluate the average inventory level per week such that it is easy to calculate total holding costs once holding costs per vehicle per day are known.

**Lead Times**

The other aspect that drives this research is the lead time. The processing times associated with each stage in the flow process are assumed to be deterministic. AFMS orders the 15\textsuperscript{th} each month and manufacturing and 1st Leg Transport together takes 13 weeks from the moment it starts the manufacturing process in Japan until it arrives at the hub in Thailand. At the hub, the vehicle is stocked until it is allocated to a field operation. When the vehicle is allocated the preparation phase starts in which the hub puts the accessories on the vehicles. The vehicle accessorization process takes 3 weeks independent of the number of vehicles being accessorized. Next, the vehicle is ready for shipment to the requesting country and is officially in service. The vehicle is assumed to be in service for 5 years. These assumptions of deterministic times are reasonable, they are often true within AFMS.

**Vehicle Models**

Van Wassenhove and Martinez (2012) state that standardization is basic for supply chain excellence. Field programs benefit from standardization in terms of increasing speed of response. AFMS is moving towards a more standardized vehicle catalog which is of huge benefit for inventory control policies as discussed by Martinez et al. (2011). Over 9 different vehicle models are currently in use. Each model has a left (LHD) and a right (RHD) hand drive variant. Figure 8 shows the relative quantity of the vehicle models requested from
April 2017 until October 2018. The top 2 models cover almost 80% of all vehicles. In this research, only these 2 models are used for analysis. The exact model types can be found in Appendix A.1. In total we are left with 1587 vehicles of these four models, of which 728 vehicles are Toyota Land Cruiser 76 LHD (TLC76-LHD), 406 Toyota Land Cruiser 76 RHD (TLC76-RHD), 266 Toyota Prado LHD (PRADO-LHD), and 187 Toyota Prado RHD (PRADO-RHD). We could run the simulation model for each type separate as the inventory policies are model specific. However, smoothing is based on the aggregated number of vehicles coming off lease. The hub does not distinguish between the models when accessorizing the vehicles. Figure 3 shows the 1587 vehicles coming off lease.

5.3 Interventions

In this section we give the results of our simulation model. We analyze three cases. Case 0 is the theoretical case. Case 1 is the benchmark case. Smoothing is implemented in case 2. The results are visualized as a Pareto curve where a point on the curve represents a specific inventory control policy. The result is given in terms of average delay, standard deviation of delay, and average inventory level. For the smoothing intervention we also analyze the effect on the standard deviation of the workload at the hub. Lastly, we check for robustness with respect to the chosen demand distribution.

Case 0: Theoretical

To see if we can predict the effects of our inventory control policies on our data we first apply our policies to theoretically generated demand. We generate 10 different demand patterns for 1 distinct vehicle model. The coefficient of variation for the various generated patterns can be found in Table 1. For this analysis we make sure that the average number of vehicles coming off lease is equal to 10 vehicles for all generated data. Our first distribution concerns a deterministic demand pattern of 10 vehicles coming off lease every week. This
pattern is visualized in Figure 9a. Next, we generate a Poisson process visualized in Figure 9b. We also generate three data sets based on sinusoidal functions which differ in variation. We generate patterns with an amplitude of respectively 3, 6, and 9. These patterns are visualized in Figure 9c, Figure 9d, and Figure 9e respectively. To generate a lumpy demand pattern we use a random number generator to generate a pattern that on average 75% of the time contains zero demand, and the other weeks a Poisson process is used to generate an average demand of 40 vehicles. We generate two other lumpy demand patterns with respectively 66.66% and 50% of the time no demand, and for the other weeks we respectively use a Poisson process with an average of 30 and 20 vehicles coming off lease. The lumpy demand patterns are visualized in Figure 9f, Figure 9g, and Figure 9h. Last, we generate two erratic demand patterns. Again, we use a random number generator to generate a pattern that uses on average 50% of the time a Poisson process with a high mean and a Poisson process with a low mean in the other periods while we make sure that on average 10 vehicles are coming off lease every week. For this process we use a low mean of 5 and 8, and respectively a high mean of 15 and 12. We visualize these patterns in Figure 9i and Figure 9j.

Table 1: The Coefficient of Variation for the Various Demand Distributions with their Parameters between Brackets.

<table>
<thead>
<tr>
<th>Distribution</th>
<th>CV</th>
<th>Distribution</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deterministic(10)</td>
<td>0.00</td>
<td>Lumpy(40)</td>
<td>1.76</td>
</tr>
<tr>
<td>Poisson(10)</td>
<td>0.31</td>
<td>Lumpy(30)</td>
<td>1.28</td>
</tr>
<tr>
<td>Sinus(3)</td>
<td>0.22</td>
<td>Lumpy(20)</td>
<td>1.12</td>
</tr>
<tr>
<td>Sinus(6)</td>
<td>0.43</td>
<td>Erratic(8,12)</td>
<td>0.36</td>
</tr>
<tr>
<td>Sinus(9)</td>
<td>0.63</td>
<td>Erratic(5,15)</td>
<td>0.59</td>
</tr>
</tbody>
</table>
(a) Deterministic Generated Demand of 10 Vehicles.

(b) Poisson Process Generated Demand with Mean 10.

(c) Sinusoidal Generated Demand with an Amplitude of 3.

(d) Sinusoidal Generated Demand with an Amplitude of 6.

(e) Sinusoidal Generated Demand with an Amplitude of 9.

(f) Lumpy Generated Demand with a Poisson Process with Mean 40.

(g) Lumpy Generated Demand with a Poisson Process with Mean 30.

(h) Lumpy Generated Demand with a Poisson Process with Mean 20.

(i) Erratic Generated Demand with a Poisson Process with Mean 5 and Mean 15.

(j) Erratic Generated Demand with a Poisson Process with Mean 8 and Mean 12.

Figure 9: Theoretical Generated Demand Patterns.
For comparison of our different inventory control policies, we show the trade-off for the various theoretical generated demand patterns. Reflecting current practice, we evaluate the policy which entails placing monthly orders so as to bring the upstream pipeline up to a constant level $S$ (i.e. a base stock policy). In practice AFMS often deviates from this policy. So for several values of $x$ we calculate the effect of strictly adhering to such a policy, where we use $x$ as a percentage of the total number of vehicles per type in service as shown in Equation (1). We use for all models a constant upstream pipeline with $x \in \{1.0\%, 2.5\%, 5.0\%, 7.5\%, 10.0\%\}$, meaning that we base the number of vehicles on Equation (1), where we procure until our inventory position reaches $S$, including backorders, which is $S = x \cdot V$, with $V$ being the total number of vehicles in use of a certain type.

The percentage values in Figure 10 refer to the target $FR$ used in our guaranteed service level approach from Equation (4) to determine our time-varying base stock level $S_t$. We do not make use of our Normal Approximation as the exact approach is computationally not extensive, executed in Java this takes approximately 20 seconds per policy. We base our value of $\mu_t$, the number of vehicles coming off lease, in Equation (5) on the generated data patterns in Figure 9 where we use a range factor of demand $\alpha = 0.3$, which means that we can get a deviation of at most 30% from $\mu_t$. To see to which extent our results depend on the choice of $\alpha$ we perform a sensitivity analysis with different values of $\alpha$ at the end of this section. Finally, $MTO$ refers to our make-to-order policy. We indicate the base stock level with ◆, the time varying fill rate policy with △, and the make-to-order policy with □. So for every policy we have several values corresponding to the 10 different demand patterns. We do this analysis based on 150 replications, the same number as determined in Section 5.2, our biggest width in the confidence interval is approximately 2.5 days. The other confidence intervals are listed in Appendix A.2.
Figure 10: The Pareto Curve Shift Showing the Trade-off Between Average Delay Versus Average Inventory Level for the Theoretical Case.

From Figure 10 we can see that all Pareto curves have the same shape. It is the shape that we expect to see based on inventory control theory. Our make-to-order policy results in the highest average delay possible but no stock. Using a base stock policy is not beneficial as average delay shows almost no decrease. Note, using a value of $x$ equal to 10.0% is not even included in the optimal curve for any of the demand patterns as its values are outside the scale. The policies that mainly define the shape of the curve are the guaranteed service level policies. Introducing these policies immediately decrease average delay while the average inventory level is barely increased as can be seen when moving from the 10.0% target fill rates to the 30.0% target fill rates. The trade-off is clearly visible when moving towards target fill rates of 50.0%, 70.0%, and 90.0%. For these target fill rates average delay decreases at the cost of increasing stock levels. It can be seen that setting higher target fill rates is not beneficial as average delay is barely decreased while stock levels increase significantly. From Figure 10 we see that the lumpy generated demand curves with Poisson process with mean 40 and 30 are situated a little bit above the other generated demand curves which makes sense as their coefficient of variations are also higher, which results in more uncertainty and thus higher average delays. However, in general, one can note that the specific inventory control policies are clustered for the different distributions. This gives us a good indication of the shape of the curve we can expect when we visualize the trade-off for the data of AFMS.

Case 1: Benchmark
Because there is a difference in quantity and variability requested between the vehicle
models, we investigate if we can apply one strategy to all models or if the models should be treated differently. We evaluate the impact of the inventory control policies on the separate models. Results of the distinct vehicle models (TLC76 LHD, etc) can be found in Figure 11. We base our value of $\mu_t$ in Equation (5) on the data from AFMS visualized in Appendix A.3 where we use $\alpha = 0.3$, which means that we can get a deviation of at most 30% from $\mu_t$.

![TLC76 LHD - Average Delay](image1)
![TLC76 RHD - Average Delay](image2)
![PRADO LHD - Average Delay](image3)
![PRADO RHD - Average Delay](image4)

(a) The Average Delay Versus Average Inventory Level.
(b) The Average Delay Versus Average Inventory Level.
(c) The Average Delay Versus Average Inventory Level.
(d) The Average Delay Versus Average Inventory Level.

Figure 11: The Pareto Curve Showing the Trade-off Between Delay Versus Inventory Level for the Separate Vehicle Models without Smoothing.

As can be seen in Figure 11 all Pareto curves show more or less the same shape. This is in line with our findings from Figure 10. The distribution of the demand pattern does not affect the trade-off between average delay and average inventory level much.

As we want to introduce our smoothing intervention in a later stadium, which has impact on the aggregated number of vehicles, we aggregate the results of the 4 vehicle models in a weighted way and show the result in one curve. We will call this the benchmark curve. As previously discussed, we also look at the standard deviation of delay as one of
our performance metrics. These results are visualized in Figure 12.

As can be seen in Figure 12a following UNHCR’s current policy of keeping a constant upstream pipeline a decrease in average delay goes at a substantial increase in average inventory level. We can see that our current situation\(^5\) (depicted with a red circle) is situated at approximately the same average inventory level of the base stock policy with value \(x = 10.0\%\), but at a significant higher level of average delay. These constant upstream pipeline policies together represent an imaginable curve that is above our Pareto curve in all instances.

Our Pareto curve shows a clear downward convex shape. The upper left part of the curve is defined by our make-to-order policy resulting in the highest average delay possible but no stock. The policies that define the shape of the curve are the guaranteed service level policies.

We make a clear transition on the curve when we move from our make-to-order policy to a target \(FR\) of 10.0%. We experience an almost vertical drop in average delay from over 16 weeks to approximately 11 weeks while barely increasing our average inventory level. The following transition to a target \(FR\) of 30.0% again barely increase our average inventory level while reducing our average delay to approximately 6.5 weeks. The average inventory level starts to increase when moving to a target \(FR\) of 50.0%. The average inventory level is now increased up to almost 6 vehicles, while again decreasing average delay to almost 3 weeks. Moving towards a target \(FR\) of 70% accounts for another 2 weeks decrease in average delay at the cost of an increase in the average inventory level of 8 vehicles. Our next target \(FR\) of 90% only marginally decreases average delay at a substantial increase in average inventory level.

The last part of the curve is characterized by inventory control policies that result in

\(^5\)Based on data from April 2017 to October 2018.
low levels of average delay. Moving from one policy to another in this region does not improve our performance much.

The results for the standard deviation of delay can be seen in Figure 12b. Standard deviations are lowest for the inventory policies with high guaranteed service levels. Following UNHCR’s current constant upstream pipeline policy results in higher standard deviations and thus more uncertainty in guaranteeing a certain lead time. The make-to-order inventory control policy and the guaranteed service level inventory control policy with the highest target fill rate define the shape of the curve. This is as expected as we are most confident about our lead times if we always choose the longest lead time, or if we almost always have sufficient vehicles in stock by guaranteeing a target fill rate of over 90.0%.

If we combine Figure 12a and Figure 12b we can show that implementing a guaranteed service level in combination with time-varying base stocks yields substantial improvements with respect to delay at a marginal increment in the average inventory level. This is an important finding for UNHCR, as this implies that realizing a much better service level does not need to be expensive. At the same time the uncertainty in guaranteeing lead times is also lowered.

Based on our analysis we can already reduce average delay from 11.6 weeks to only 0.2 weeks, standard deviation of delay from 7.4 weeks to almost 1.2 weeks, and average inventory level from 62.4 vehicles to 27.6 vehicles. This can be done by simply introducing our guaranteed service level policy with a target $FR$ of 90.0%. Respectively, improvements are made of approximately 98%, 84%, and 56%. A recent evaluation of UNHCR’s GFM programme showed an average lead time of almost 6 months (Universalia, 2018). If the included delay would thus be reduced we can expect a realization of average lead time of approximately 3.5 months, resulting in a 40% gain.

If we introduce our guaranteed service level policy with a more conservative target $FR$ of 70.0%. We reduce our average delay to 1.1 weeks, standard deviation of delay to 2.6, and average inventory level to 13.7 vehicles. Approximate improvements are now made of respectively 90%, 65%, and 78%.

**Case 2: Demand Smoothing**

We introduce a policy that gradually lowers the variability in the number of vehicles coming off lease based on Figure 3. In Figure 13 the results are visualized where smoothing over two periods is indicated as $p = 2$ and the original situation of non-smoothing is indicated as $p = 1$ (the top red curve). The points that are not on the curve represent the specific inventory control policies belonging to the smoothing intervention. We use the Pareto curves shown in Figure 12 as a benchmark for our new intervention. As can be seen in Figure 13 the shape of the curves is almost identical to the benchmark case. The savings are small in absolute terms, a delay reduction of about 1 week (e.g. from 3 to 2 weeks) at the same inventory level.
Sensitivity Analysis: Smoothing Horizon

To minimize complexity lower levels of smoothing are preferred as there is less administrative work involved in having smaller smoothing horizons. We evaluate the use of a smoothing horizon \( p \in \{1, 2, 5, 10\} \) in Equation (15) to show the benefits of increased complexity. \( p = 1 \) is the original situation of non-smoothing. Figure 14 shows that the impact of higher levels of smoothing is rather similar in terms of delay. Increasing our smoothing horizon makes us able to move to slightly better outcomes.

(a) The Average Delay Versus Average Inventory Level.

(b) The Standard Deviation of Delay Versus Average Inventory Level.

Figure 13: The Pareto Curve Shift Showing the Trade-off Between Delay Versus Inventory Level when Introducing Smoothing.

(a) The Average Delay Versus Average Inventory Level.

(b) The Standard Deviation of Delay Versus Average Inventory Level.

Figure 14: The Pareto Curve Shift Showing the Trade-off Between Delay Versus Inventory Level for Various Smoothing Levels.

The most important impact of smoothing can be found at the hub. Figure 15 shows the decrease in standard deviation of the workload at the hub for the different levels of smoothing. Let us look at the most interesting policies in our research, the guaranteed
service level policies with target fill rates of 70.0% and 90.0%. For these cases a smoothing horizon of two periods only marginally decrease workload variability. Smoothing has more impact if we increase our smoothing horizon from two to five or even ten periods. Table 2 shows the decrease in coefficient of variation for these cases as a result. Standard deviation decreases when we increase the smoothing horizon. The average workload at the hub stays the same. Thus, the coefficient of variation decreases when increasing the smoothing horizon. Implementing smoothing seems to substantially reduce variability in the hub. However, smoothing horizons of five or even ten periods increase complexity.

Table 2: The Coefficient of Variation for the Different Smoothing Horizons for the Guaranteed Service Level Policies with Target Fill Rates of 70.0% and 90.0%.

<table>
<thead>
<tr>
<th>FR = 70.0%</th>
<th>FR = 90.0%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smoothing Horizon</td>
<td>CV</td>
</tr>
<tr>
<td>p=1</td>
<td>0.760</td>
</tr>
<tr>
<td>p=2</td>
<td>0.733</td>
</tr>
<tr>
<td>p=5</td>
<td>0.628</td>
</tr>
<tr>
<td>p=10</td>
<td>0.526</td>
</tr>
</tbody>
</table>

To illustrate the effect of the spikes in demand, we show the difference in our trade-off when AFMS can make sure that the vehicles coming off lease from Figure 3 are completely smoothened versus the situation as it is now. Notice that we still have a demand variation, as it can be 30% more or less than the mean. We show these results in Figure 16. The

Figure 15: The Effect of Different Smoothing Horizons on the Variability of the Workload at the Hub.
result for the standard deviation of delay is shown in Appendix A.4. The base stock policies now perform much better, but still worse compared to the guaranteed service level inventory control policies.

Figure 16: The Pareto Curves Showing the Trade-off Between Delay Versus Inventory Level for the Benchmark Case Versus the Case in which Vehicles Coming Off Lease are Completely Smoothened.

It is clear that Figure 16a shows a more desired curve than Figure 16b. If vehicles coming off lease are already smoothened, spikes are not present anymore. Our proposed inventory control policies perform better when implemented on more stable demand flows. It is interesting to note that it is not worth increasing our target fill rate anymore after 70.0% as the average delay is already minimal and this would thus only increase our average inventory level of vehicles.

Sensitivity Analysis: Demand Variation
AFMS will start disposal of GFM rental vehicles as of 2019, therefore it is impossible to determine the exact distribution that captures the variation. Upon replacing the vehicles it can be decided to increase or decrease the fleet size. For the whole analysis we used a uniform demand distribution that allows for a range of a 30% increase or decrease in demand based on the vehicles coming off lease as shown in Equation (5)(depicted in red). To check robustness we show the Pareto curves for two other uniform demand distributions in Figure 17. The first represents a more optimistic distribution in which demand deviates up to 10% from the number of vehicles coming off lease (depicted in blue). The other represents a more pessimistic distribution in which demand deviates up to 50% (depicted in green). Figure 17a shows the three Pareto curves for the average delay for the different distributions. Figure 17b shows the curves for the standard deviation of delay. The resulting curves when we do not introduce smoothing have similar results and can be found in Appendix A.5.
(a) The Average Delay Versus Average Inventory Level.  (b) The Standard Deviation of Delay Versus Average Inventory Level.

Figure 17: The Pareto Curve Shift Showing the Trade-off Between Delay Versus Inventory Level for Various Demand Variation Distributions with Smoothing.

Figure 17a shows that the Pareto curves do not differ substantially. For higher target fill rates the inventory levels differ while the average delay stays the same. This seems logical as guaranteeing high service levels under higher demand uncertainty generally comes with high costs. This not only shows the importance of minimizing uncertainties through adequate communication and coordination with COs on vehicle needs. It also implies that adequate demand forecasting is important to meet the targeted service levels.

6 Discussion

Data
Limited data and limitations in data quality complicate proper demand forecasting. The new structure (centralization) has only been active since 2014. So there is limited data available. The first vehicles are coming off lease as this thesis is written. For further research it is important to analyze the behavior that comes with this event. Do countries indeed replace their vehicles with models of the same type? Do they strictly follow the five year lease term policy? Can the deviation in the number of vehicles replaced be captured in a discrete uniform distribution? Guaranteeing high service levels under higher demand uncertainty generally comes with high costs. This not only shows the importance of minimizing uncertainties through adequate communication and coordination with COs on vehicle needs. It also implies that adequate demand forecasting is important to meet the targeted service levels.

Emergencies
It is especially interesting to analyze what will happen when we move to a just-in-time inventory control policy if AFMS requires COs to request vehicles well in advance as this will result in almost deterministic demand for development missions. The disruptions in
the supply process are now mostly caused by emergencies. True emergencies are often times unpredictable but give a tremendous increase in demand and disrupt the supply chain. It is extremely difficult to predict emergencies, but demand shocks and disruptions in the allocation process can be captured in a model. Pedraza-Martinez and Van Wassenhove (2012) describe the challenge of having a dual object of relief and development. A logistic system is needed that at the same time makes short-term decisions under high uncertainty with time of response as the main objective as well as long-term decisions under low uncertainty with cost efficiency as the main objective. The structure of the system should be based on the proportion of relief and development work done. The right balances of prepositioning and global procurement needs to be determined.

Finance
Another important aspect of the vehicle supply process is the financial situation of AFMS. The current financial structure complicates the FMS. Vehicles (assets) are costly. Currently, vehicles are kept for (at least) 5 years while the budget cycle is 1 year. UNHCR is largely dependent on the money of donors. If AFMS has budget left at the end of the budget year, they might receive less funding in the following year. The budget cycle creates an incentive to use remaining resources on vehicles late in the year since the cars are expensive and thus easily drain the budget. The increased ordering, or “Christmas shopping” as AFMS is naming it, results in unallocated stocks, which increase holding costs. We should compare the current structure to a VSP without these financial constraints. To capture this financial situation it is important to follow the financial flows and include the financial constraints in the model. AFMS can be supported if it is shown what the financial credit should be to generate certain efficiencies.

7 Conclusion

Saving fleet costs can free up substantial resources, which can be spent on additional help for beneficiaries. Additionally, fleet has a major direct impact on beneficiaries. If fleet is not properly managed, offering timely help becomes challenging if not impossible.

UNHCR’s Global Fleet Management (GFM) project was introduced in 2014 to address shortcomings in the organization’s decentralized fleet management practices (from procurement to operation and asset disposal). The implementation of the centralized system is perceived to be successful in many aspects, but there is substantial room for improvement. In this thesis we have focused on the vehicle supply process (VSP) of the Global Fleet Management (GFM) program. The VSP involves all physical, administrative, and financial activities needed to fulfill a country office’s (CO’s) vehicle request. The process of delivering vehicles has grown haphazardly and become cumbersome. Optimizing the VSP is needed.

---

6AFMS receives budget from UNHCR based on the previous year.
In this thesis we have analyzed two supply chain interventions by means of a discrete event simulation model. The performance criteria include costs, service, and/or regularity of flows. First, we have considered inventory control policies that regulate the inventory position at each point in time. Second, we have analyzed a demand smoothing policy, which creates a more stable demand flow.

In our research we have evaluated three inventory control strategies. Reflecting current practice, the first strategy entails placing monthly orders so as to bring the upstream pipeline up to a constant base stock level $S$. To evaluate the strategy we tested multiple levels of $S$. The second strategy we evaluated is make-to-order, where we keep no stock and only procure after COs request vehicles. Third, we tested a guaranteed service level inventory control policy. We evaluated this policy with multiple target fill rates.

We have shown that following UNHCR’s current policy of keeping a constant upstream pipeline gives us quite poor performance. We analyzed our current situation and show that this is far from optimal. However, introducing a guaranteed service level inventory control policy with time-varying base stocks substantially lower delays while only marginally increase average inventory levels.

We have shown that introducing guaranteed service level inventory control policies are robust with respect to the demand distribution, meaning that more uncertainty results in keeping more stock, but not in a different trade-off position. The optimal curve shifts towards more inventory, but the policies defining the shape of the curve are still the same. This shows that the required inventory level is thus highly dependent on the uncertainty in demand. Proper forecasting and reducing uncertainty by clear communication is thus key.

We also proposed a policy for smoothening future demand. Demand spikes are gradually lowered by allowing certain vehicles to come off lease in an earlier or later stage. This policy makes sure that the variability in demand is reduced such that inventory control policies can be more easily implemented. We have shown that demand smoothing results in lower levels of workload variability at the hub. We evaluated longer smoothing horizons, this creates more complexity but substantially decreases workload variability. Implementing smoothing seems very promising when increasing complexity by smoothing over longer horizons is not a bottleneck.

We have provided an in-depth discussion of the performance before and after the introduction of our interventions. We have shown how these interventions enact a positive continual improvement loop that can lead to professional fleet management. It is up to UNHCR to make the trade-off between costs and service. However, based on our data we have shown that we can reduce lead times by roughly 40%, while at the same time reducing holding costs and workload variability at the hub by moving from a relatively myopic fixed upstream pipeline policy to a more advanced inventory system. In this research we have shown that delay can be substantially lowered with marginal cost increments. Implementing a guaranteed service level inventory control policy also provides UNHCR with a formal policy to provide to newly hired staff. Implementing smoothing has important
implications for UNHCR. Smoothing does require some communication and coordination efforts with COs, as they have to agree on submitting their orders somewhat sooner or later. A little bit of such coordination, as we show, already makes a substantial positive impact on holding costs. Moreover, the resulting decrease in workload variability at the hub can substantially save costs for overtime and/or hiring additional personnel.

Instituting a guaranteed service level inventory control policy takes time and requires strong commitments at the highest level of the organization, as well as significant investments. However, once fully in place, the policy has the potential to save UNHCR millions of dollars through more effective and efficient management of one of UNHCR’s most valuable assets: its light vehicles.
8 References


A  Appendix

A.1  Vehicle Model Coverage

![Land Cruiser Models](a) Coverage Land Cruiser Models

![Prado Models](b) Coverage Prado Models

Figure 18: The Coverage of the Specific Models within their Parent Model.

A.2  Confidence Intervals for Various Demand Distributions

Table 3: The Confidence Interval and the Width of the Interval for the Average Delay in Weeks for the Various Demand Distributions with their Parameters between Brackets for the Guaranteed Service Level of 50.0%.

<table>
<thead>
<tr>
<th>Distribution</th>
<th>CI Width</th>
<th>Width</th>
<th>Distribution</th>
<th>CI Width</th>
<th>Width</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deterministic(10)</td>
<td>[4.15, 4.49]</td>
<td>0.34</td>
<td>Lumpy(40)</td>
<td>[4.98, 5.20]</td>
<td>0.22</td>
</tr>
<tr>
<td>Poisson(10)</td>
<td>[3.83, 4.17]</td>
<td>0.34</td>
<td>Lumpy(30)</td>
<td>[5.57, 5.87]</td>
<td>0.30</td>
</tr>
<tr>
<td>Sinus(3)</td>
<td>[3.75, 4.09]</td>
<td>0.35</td>
<td>Lumpy(20)</td>
<td>[3.91, 4.17]</td>
<td>0.27</td>
</tr>
<tr>
<td>Sinus(6)</td>
<td>[3.39, 3.73]</td>
<td>0.34</td>
<td>Erratic(8,12)</td>
<td>[3.41, 3.75]</td>
<td>0.33</td>
</tr>
<tr>
<td>Sinus(9)</td>
<td>[3.33, 3.65]</td>
<td>0.33</td>
<td>Erratic(5,15)</td>
<td>[2.81, 3.11]</td>
<td>0.30</td>
</tr>
</tbody>
</table>
A.3 Vehicles Coming Off Lease Separated by Vehicle Model

Figure 19: Number of Vehicles Coming Off Lease per Month from 2019 until 2023 Aggregated from Weeks to Months based on Data from AFMS Separated by Model.

A.4 Sensitivity Analysis: Smoothened Demand versus The Benchmark

Figure 20: The Pareto Curves Showing the Trade-off Between Delay Versus Inventory Level for the Benchmark Case versus the Case in which Vehicles Coming Off Lease are Completely Smoothened.
A.5 Sensitivity Analysis: Demand Variation without Smoothing

(a) The Average Delay Versus Average Inventory Level.

(b) The Standard Deviation of Delay Versus Average Inventory Level.

Figure 21: The Pareto Curve Shift Showing the Trade-off Between Delay Versus Inventory Level for Various Demand Variation Distributions without Smoothing.