Optimizing Logistics with AI:

A Conceptual Comparison Framework of Manual and Automated Route Planning at Bouwvervoer

Cem Yildiz

Bachelor Thesis

Industrial Engineering and Management

Author

C. Yildiz (Cem)

Supervisors

First supervisor University of Twente: dr. J.P.S. Piest (Sebastian)

Second supervisor University of Twente: dr. P.B. Rogetzer (Patricia)

Company Supervisor: A. Berndsen (Arjan)

Company

Bouwvervoer

Educational Institution

University of Twente

Drienerlolaan 5, 7522 NB, Enschede

Publication Date

22/03/2025

Management Summary

Bouwvervoer, a construction transport company with a fleet of approximately 100 vehicles, currently relies on a semi-digital manual planning system to coordinate its logistics operations. While digital tools support this approach, they heavily depend on human decision-making, introducing inefficiencies in route planning, resource utilization, and delivery performance. With the logistics industry rapidly evolving through Artificial Intelligence (AI) integration, Bouwvervoer initiated a pilot AI-based planning system. However, the company lacks a structured method to evaluate whether the AI-driven system outperforms the current manual setup and aligns with its strategic goals.

This thesis addresses that gap by developing a conceptual comparison framework to compare both systems systematically. The aim is to support data-driven decision-making on AI adoption in logistics planning. The framework is built upon eight Key Performance Indicators (KPIs), categorized under transportation effectiveness, employee satisfaction, and customer service quality. These KPIs, Total Distribution Cost, Capacity Utilization, % Empty km's / Total km, Turnover per Hour per Driver, On-Time Delivery Performance, Satisfaction with Working Hours, Fulfilment of Specific Driver Requests, and Order Accuracy, were selected through a literature review and input from company stakeholders.

A multiplicative analytical hierarchy process (MAHP) was applied to prioritize the KPIs using pairwise comparisons completed by six experienced planners. The KPI weights were then implanted in two analysis methods: the Weighted Composite Score (WCS) and Weighted Difference Analysis (WDA), both of which quantify system-level performance differences. These methods were implemented in an interactive Power BI dashboard alongside descriptive statistics and selected visualization techniques.

Although the dashboard was built using mock-up data due to operational data limitations, it demonstrates how the framework would function in practice. The dashboard allows users to analyze KPI trends, compute performance scores, and visually compare systems over time. The solution was evaluated through structured feedback from the company supervisor, who assessed the tool's usability, design, and decision-making impact. The evaluation confirmed that the dashboard offers valuable managerial insights but highlighted areas for refinement, including navigation improvements and more direct visual comparisons.

The developed framework is context-specific to Bouwvervoer but generalizable to similar mid-sized logistics firms exploring AI integration. To extend its applicability, future implementations should incorporate accurate operational data and engage a broader range of stakeholders for evaluation.

In conclusion, this thesis offers a structured tool for comparing manual and AI-driven logistics systems. It supports data-based decision-making and applies the groundwork for the broader adoption of AI technologies in logistics planning.

Table of Contents

List of Figures

List of Tables

List of Equations

List of Abbreviations

AHP – Analytic Hierarchy Process

AI – Artificial Intelligence

CI – Consistency Index

CR – Consistency Ratio

DEA – Data Envelopment Analysis

IT – Information Technology

KPI – Key Performance Indicator

MAHP – Multiplicative Analytic Hierarchy Process

MPSM – Managerial Problem-Solving Method

PCM – Pairwise Comparison Matrix

RI – Random Index

SCOR – Supply Chain Operations Reference

SLR – Systematic Literature Review

TOPSIS – Technique for Order of Preference by Similarity to Ideal Solution

WCS – Weighted Composite Score

WDA – Weighted Difference Analysis

1. Introduction

This chapter introduces the research context, company background, and the core problem that motivated the study. It begins by outlining the growing role of AI in the logistics industry. The chapter then narrows its focus to Bouwvervoer, a construction transport company with a manual planning system. As the company explores AI-driven planning solutions, it lacks a structured method to compare the manual system with the proposed AI system. This chapter defines the nature of that gap through a problem analysis, including a problem cluster and a measurement of the gap between norm and reality. These sections lay the foundation for developing a comparative performance framework.

1.1 General/Scientific Context

This section aims to provide a broad overview of the role of AI in the logistics industry. It highlights how AI has become a key driver of innovation, leading to improved operational efficiency and effectiveness. It provides a foundational understanding of why AI is crucial for companies aiming to stay competitive in today's fast-evolving logistics industry.

The logistics industry is rapidly developing with the advancement of digital technologies and AI. As companies aim for efficiency and sustainability, adopting AI in logistics has become a key feature distinguishing them from competitors. AI-based systems can optimize decision-making by leveraging large volumes (as much as possible) of data, thereby reducing operational costs, improving route efficiency, and minimizing environmental impacts. The role of AI in logistics has been widely recognized, particularly for its ability to automate complex processes that previously required human intervention. AI-driven logistics systems can analyze vast amounts of historical and real-time data to make more informed decisions. Studies find up to 46.15% increase in decision outcomes (Rinat et al., 2024). Also, Choi et al. (2016) gives hints about AI technologies integration into logistics will deepen, offering more sophisticated solutions such as autonomous vehicles, drone deliveries, and AI-driven supply chain management systems.

These innovations improve operational efficiency and create a competitive edge for companies that adopt them, driving the logistics industry toward a more digitized and automated future. However, as highlighted by Zhang (2019), the transformation of traditional logistics through AI presents significant challenges alongside these opportunities. There is a considerable gap between the logistics industry's current capabilities and the need for intelligent data collection and perception infrastructure, such as big data and logistics cloud systems. Additionally, the coverage and accuracy of logistics-related internet technologies are still insufficient, restraining the full realization of intelligent logistics. Furthermore, the logistics industry is currently struggle with a critical shortage of technology and talent. Despite the industry's ability to generate massive amounts of data and present rich business scenarios, the lack of skilled professionals with comprehensive knowledge of logistics and AI is an ongoing issue that is slowing progress.

These challenges illustrate that while AI presents possibilities for logistics, significant infrastructure advancements, talent development, and standardization are required to unlock its potential fully. Companies like Bouwvervoer that can successfully navigate these challenges and leverage AI's capabilities stand to gain a competitive advantage as the logistics industry evolves.

1.2 Company Context

This section outlines the historical and operational context of Bouwvervoer. It highlights the company's growth, the increasing complexity of its operations, and the challenges posed by dependence on manual logistics planning. This background sets the stage for exploring AI as a transformative solution.

Bouwvervoer is a family-owned logistics service specializing in transporting construction materials. It was established in 1920. The company is committed to continual improvement and operates with a fleet of approximately a hundred vehicles (bouwvervoer.nl, 2023). Bouwvervoer is at a critical point in its journey toward operational advancements and sustainability. As the company grows its customer base and the complexity of its operations grow as well, and faces challenges that make manual logistics planning insufficient. These challenges significantly affect operational efficiency, cost-effectiveness, and environmental sustainability. A possible solution to these challenges, AI, is increasingly getting integrated into logistics and supply chain management, and the quantity of data and continuous advancements in computing capabilities open new possibilities for improving decision-making in supply chains (Boute et al., 2022). The trend toward adopting AI and machine learning solutions to solve similar challenges highlights the need for Bouwvervoer to re-consider its existing planning systems.

1.3 Defining the Problem

In this section we address the challenges and opportunities Bouwvervoer faces with its current logistical framework and the exploration of the AI system. In section 1.3.1, we examine the critical issue—Bouwvervoer's need for an evaluation framework to transition from its semi-digital, manual system to the AI-driven system. This section highlights the absence of comprehensive guidance for comparing both systems' operational efficiency, cost-effectiveness, and sustainability impacts. The section 1.3.2 elaborates on the core problem: the lack of a comparative analysis framework for logistical planning systems. It shows the resulting intermediate problems, such as uncertainty in performance evaluation and inadequate data utilization, originating from this gap. Finally, in section 1.3.3 we focus on the current challenges regarding norm and reality.

Over the decades, Bouwvervoer has consistently aimed to leverage technology to enhance its logistical operations, adopting an optimization software since the 1990s. Despite these efforts and several attempts to achieve fully automated planning, the company has encountered challenges integrating these technologies into its existing workflows. This highlights the motivation behind exploring the AI system as a potential solution to the enduring complexities of manual planning.

Bouwvervoer's logistical operations have relied on a manual planning system in its core approach. Using a Gantt chart embedded in the system, planners decide which truck to use and which route for which customers. However, it is supported by digital tools and IT systems. This integration of digital assistance within a mainly manual framework creates a unique operational way of working. In this thesis, the term 'manual' does not mean 'on paper'; instead, it refers to a system where human decision-making commands, although assisted by digital tools. This combination of manual management with digital tools provides a specific efficiency level but also brings limitations and challenges.

Even with the support of digital tools, the manual planning process at Bouwvervoer faces operational challenges, such as inefficiencies in route planning and logistical coordination. Financially, this approach can lead to increased costs, which impacts fuel consumption and time management. Environmentally, the inability to achieve the most efficient and sustainable routes leads to increased carbon emissions, challenging Bouwvervoer's commitment to sustainability (bouwvervoer.nl, 2023).

Introducing AI system as a pilot project presents a new opportunity for optimization. However, a gap exists in the need for evaluation criteria to effectively compare the manual system with this new AI-driven solution. This gap highlights the need for a structured approach to evaluate and quantify the differences and potential improvements the AI system offers over the current mode of operation.

1.3.1 Identification of Action Problem

Bouwvervoer needs a clear framework or criteria to effectively compare and evaluate the performance of the current semi-digital, manual system against the AI-driven solution. Incorporating the human perspective of decision-makers, it becomes essential to develop criteria that reflect the company's strategic objectives and logistics planners' understanding and instincts.

This situation leads to the primary action problem, anything, or any situation different from how one wants it to be is an action problem (Heerkens, 2017): *Missing Guidance on Decision-Making to Switch to AI Planner.* The missing guidance pertains to several critical areas: evaluating both systems' operational efficiency, cost-effectiveness, and sustainability alignment under future conditions. More specifically, the guidance gap involves creating comparative frameworks that can assess potential differences, impact on key performance indicators, and alignment with Bouwvervoer's long-term strategic goals. This comparison requires understanding the AI and Manual systems and identifying which aspects of logistical planning and execution can be optimized or need adjustment when transitioning to an AI-driven approach.

As Bouwvervoer continues to grow, the scalability of manually managing logistics becomes a considerable concern. The manual system becomes more time-consuming and burdens planners, particularly in high-demand periods, increasing the possibility of errors or oversights. These challenges are not unique to Bouwvervoer but reflect broader trends in the logistics and transportation industry. Choi (2016) says third-party logistics service providers and traditional transportation and shipping companies face significant operational risks. highlights that the high level of uncertainty and threats to system reliability are critical issues in global logistics systems. Furthermore, Zhang (2019) highlights that as AI technology gains broader adoption, numerous modern logistics companies are exploring its use to improve various operations and boost overall logistical efficiency. In the larger context, improving route efficiency and reducing the number of trips matches with cost-saving measures and global ambitions towards environmental responsibility.

1.3.2 Problem Cluster and Motivation of Core Problem

The core problem is stated as following: *Missing Comparative Analysis Framework for Logistics Planning Systems.* This fundamental gap interferes with systematically evaluating and placing the manual planning system with the AI-driven system. This lack of comparative insight causes a series of intermediate problems shown in the problem cluster (*Figure 1*).

First, there is inadequate data utilization for decision-making, which means that organizations lack a structured approach to leverage available logistics data effectively. This limitation leads to unclarity in planning processes, making it difficult to quantify the potential improvements AI could present. Second, the undetermined readiness level of the AI system for operational use prevents decision-makers from measuring whether AI-based solutions are mature enough to replace or complement manual systems. AI models may require further testing, calibration, or integration with existing workflows before they are fully operational. Lastly, these two challenges are collectively contributed by the non-existence of a system/tool for comparing manual and AI planning methods. Without a structured method for comparison, logistics companies struggle to make informed decisions regarding AI adoption, leading to hesitation and suboptimal implementation.

These challenges combine into a critical action problem: the need for clear, data-driven guidance on transitioning to the AI-driven logistics planning system. With this guidance, Bouwvervoer can make a well-informed decision about embracing new technology to enhance operational efficiency, reduce costs, and boost sustainability efforts.

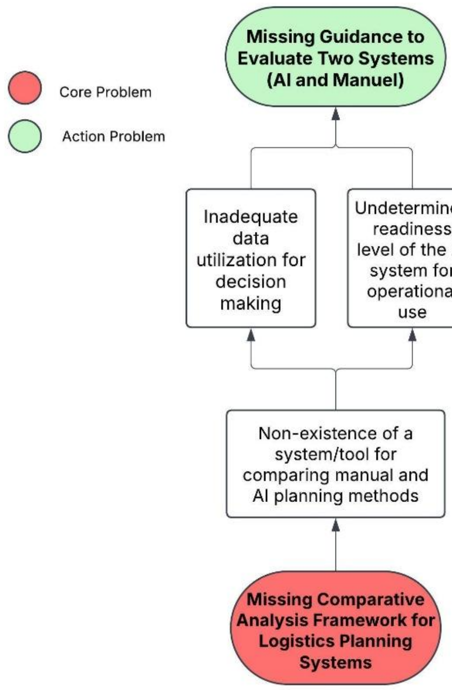

Figure 1: Problem Cluster

1.3.3 Measurement of Norm and Reality

The intermediate problems established in the problem cluster require defining the expected standards, or norms, against which the current reality can be assessed. The norm represents the ideal state where logistics planning is optimized. At the same time, the reality reflects the current state of operations at Bouwvervoer, which relies on semi-manual decision-making supported by digital tools. The absence of a comparative analysis framework creates a gap between these two states. While AI-driven logistics planning is designed to achieve maximum efficiency through full data utilization, automated route optimization, and minimized inefficiencies, Bouwvervoer's current system remains heavily dependent on human expertise, which, despite leveraging digital tools, introduces inconsistencies and inefficiencies in decision-making.

One of the key dimensions of measurement is operational efficiency. The norm in AI-driven logistics assumes that the system optimally utilizes the fleet, minimizes empty kilometres, and ensures commitment to planned schedules. The reality, however, shows that manual planning decisions, although based on experience, can lead to inefficiencies such as inconsistent vehicle utilization and an increased number of empty kilometres, affecting both costs and service quality. Cost-effectiveness is another essential aspect of comparison. AI planning is expected to distribute resources optimally, reducing operational expenses related to fuel consumption, vehicle wear, and labour costs. In contrast, manual planning, even when supported by digital tools, often results in inefficiencies that drive up transportation costs due to suboptimal route selections and human limitations in processing vast amounts of logistical data.

Another critical dimension of measurement is data utilization. AI-driven logistics systems integrate and analyse historical and real-time data to improve planning accuracy, ensuring that decisions are continuously improved based on evolving conditions. However, the reality at Bouwvervoer reflects a lack of a structured approach for effectively leveraging available data. Decisions are primarily based on planners' expertise and experience rather than systematic data-driven insights, limiting the potential for optimization. Scalability and adaptability further demonstrate the gap between norm and reality. AI-driven logistics systems are designed to scale up with growing demand. The reality at Bouwvervoer, however, is that the manual system becomes increasingly difficult to manage as operational complexity grows. The increasing workload on planners can result in bottlenecks, increased errors, and slower response times, eventually limiting the company's ability to meet market demands.

2. Methodology

This chapter outlines the study's methodological foundation. It introduces the MPSM framework and explains how its seven phases are tailored to Bouwvervoer's specific case. It then details the research design, including stakeholder involvement, research questions, chosen methods, and data collection strategies. Finally, the chapter reflects on key limitations and evaluates the validity and reliability of the methodological approach.

2.1 MPSM Approach and Fundamentals of the Research

This section outlines the methodology adopted for the research, following MPSM. The MPSM was selected as the research methodology because it is a structured, systematic approach to addressing complex problems, particularly good at managerial and business problems. MPSM's seven-phase approach enables detailed investigation, from problem identification to solution implementation and review.

The MPSM, a systematic method including seven distinct phases (Heerkens, 2017), is applied throughout this research. Each phase is specified to the context of Bouwvervoer's transition to AI-driven logistics planning and a summary of it can be seen in Table 1.

Phase	Description	Deliverable
1: Defining the Problem	Initial identification of challenges with the current manual system and the potential of the AI system. Analysis of Bouwvervoer's operations to identify core and action problems.	A comprehensive problem statement detailing the operational inefficiencies and the potential for AI-driven logistics planning.
2: Formulating the Approach	Outlining the research goals and aligning them with the MPSM process. Identifying fundamental topics to guide the research.	A research approach document outlining objectives, key questions, and the methodological framework.
3: Analysing the Problem	Deepening the understanding of the problem through revisiting definitions, identifying knowledge gaps, and uncovering new causes.	An analysis highlighting new insights into the logistical challenges and potential areas for AI integration.
4: Formulating Solutions	Developing potential solutions like a comparative analysis tool or framework using a multicriteria model to evaluate various metrics and methodologies.	A solutions framework document detailing proposed tools, evaluation criteria, and alternative methodologies.
5: Choosing the Solution	Evaluating proposed solutions to select the most effective and feasible option for comparative analysis.	A decision justifying the chosen solution based on its alignment with research goals and Bouwvervoer's needs.
6: Implementing the Solution	Applying the selected comparative analysis tool to evaluate manual versus AI-driven systems at Bouwvervoer.	Implementation of the comparative analysis tool, with initial findings and insights into both systems.
7: Evaluating the Solution	Assessing the effectiveness and functionality of the implemented solution. Identifying limitations and areas for improvement.	An evaluation report with a comprehensive assessment of the solution's impact, effectiveness, and recommendations for future enhancements. (Not fully delivered because of the limitations of the thesis)

Table 1: MPSM Phases

To effectively evaluate the AI-driven system for Bouwvervoer, an understanding of its functionalities and applications is essential. This includes studying AI applications in logistics to ensure proper implementation in the comparative analysis. This phase aligns with MPSM Phase 2, focusing on using AI-driven logistics planning to address the identified problems from a broad industry perspective.

A critical aspect is assessing Bouwvervoer's current use of its semi-digital, manual planning system. Falling under MPSM Phase 3 involves analysing the existing setup and identifying areas where AI can enhance efficiency. Research is conducted to understand the manual system's limitations and data usage.

Selecting the right KPIs is crucial for measuring and improving logistics performance. A comprehensive literature review is conducted to identify relevant KPIs in logistics and supply chain management. This task is part of MPSM Phase 3 and inform the development of the comparative analysis tool.

The development of the comparative analysis tool contain the findings from the above phases, including KPI selection and understanding of AI functionalities. This falls under MPSM Phase 4 and Phase 5. The tool is designed to provide clear, actionable insights for decision-making at Bouwvervoer.

In Chapter 5, a functional tool for comparative analysis between manual and AI-driven logistics planning systems is delivered. The tool display mock up data and KPIs effectively, aiding in efficient decision-making for Bouwvervoer’s logistics planning.

2.2 Research Design

This section outlines the methodological approach to investigate how Bouwvervoer can evaluate the transition from its manual logistics planning system to an AI-driven alternative. The research design follows the framework of the MPSM model and combines both exploratory and descriptive research strategies. It includes an analysis of key stakeholders, formulation of research questions and knowledge problems, and a description of the chosen research types and methods for data collection. Additionally, it reflects on the methodological limitations in the study, such as data constraints and the absence of a full-scale AI implementation.

2.2.1 Stakeholder Analysis

The primary subjects of this research are linked to the operational and strategic aspects of logistics planning at Bouwvervoer. This research's stakeholders include (can also be seen in *Figure 2*):

- **Researcher:** In the context of this project, the primary researcher is a bachelor's student responsible for designing the study, conducting a literature review, engaging with Bouwvervoer's management and IT experts for primary data collection, and analysing qualitative and quantitative data. The researcher's role is to evaluate the potential and challenges of adopting AI-driven logistics solutions at Bouwvervoer.

- **Bouwvervoer’s management and operational staff, who are responsible for the daily planning of the routes:** Bouwvervoer's management and operational staff offer insights into daily logistics planning and operational challenges. They provide real-world perspectives on AI integration's efficiency and potential impact. Potential conflicts of interest for Bouwvervoer's management and operational staff in the project could come from resistance to change associated with AI implementation. Staff may fear AI-driven systems could automate tasks, leading to job losses or role alterations. Additionally, there might be concern about the transparency and fairness in how AI integration decisions are made, possibly affecting trust and cooperation within the team.

- **IT experts working on the AI system:** The IT experts working on the AI system are important in the project for their technical expertise in AI development and integration. Their role involves designing, developing, and implementing the AI system, ensuring it aligns with Bouwvervoer's operational requirements and objectives. They provide technical insights, support evaluating the system's effectiveness, and address technical challenges.

- **Bouwvervoer's existing IT supplier**: As the provider of the current logistics planning solutions, IT supplier's perspective is crucial for understanding the technical and operational baseline from which AI integration is considered. However, the potential shift to AI-driven system introduces a conflict of interest, given that both services serve similar operational needs but are rivals in the technology solutions space. This concerns about IT supplier's reduced involvement and the strategic implications for Bouwvervoer's choice in logistics planning technologies.

- **Potentially, other logistics companies that are interested in AI integration.**

  Their insights, expectations, and experiences provide a perspective necessary for the research's success. By engaging with these various stakeholders, the research aims to capture a view of the transition to AI-driven logistics planning, ensuring that the findings are relevant, practical, and beneficial for parties involved.

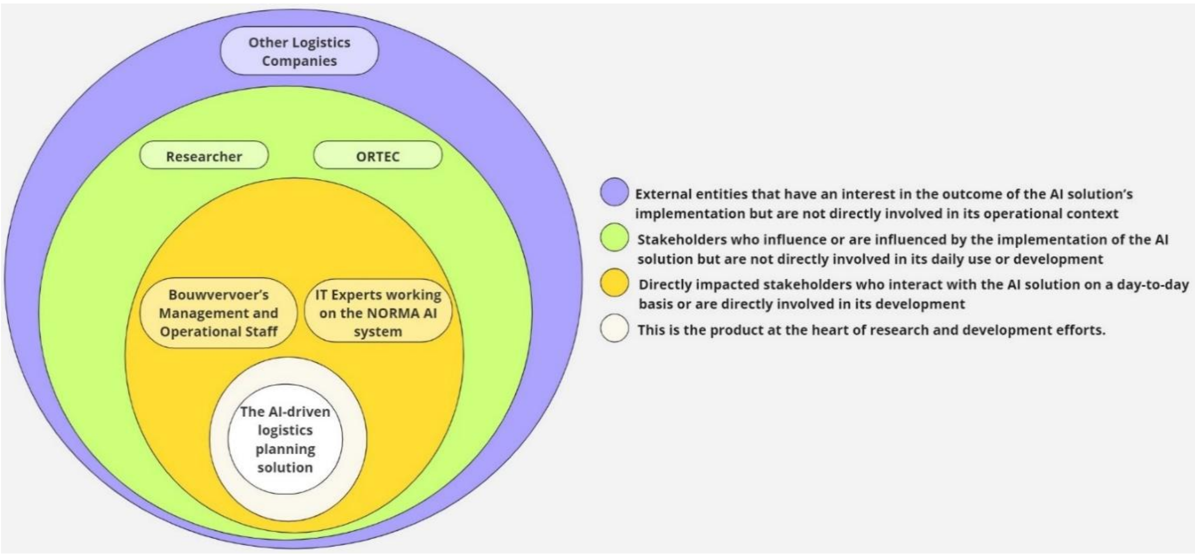

Figure 2: Stakeholder Onion Diagram

2.2.1 Research Questions and Knowledge Problems

Section 2.2.1 focuses on the critical questions driving this research, centred on Bouwvervoer's exploration of AI in logistics. The main question seeks to develop a framework for comparing Bouwvervoer's manual system with the AI-driven system. Sub-questions delve into identifying KPIs relevant to Bouwvervoer's goals, assessing the current data availability for supporting this analysis, analysing the importance of chosen KPIs, and exploring frameworks for KPI selection, data visualization, and analysis methodologies. The chosen questions can be seen in *Table 2*.

Question/Knowledge Problem	Research Type	Data Gathering Method	Method Nature
Main Research Question: How can Bouwvervoer effectively compare its existing manual logistics planning system with the AI-driven system?	Exploratory & Descriptive	Literature Review, Semi-Structured Interviews	Qualitative
RQ1: What are the most critical KPIs for evaluating the efficiency and effectiveness of logistics planning systems, specifically in Bouwvervoer?	Exploratory	Literature Review, Expert Interviews	Qualitative
RQ2: What is Bouwvervoer's perspective on the hierarchy of the selected KPIs for evaluating logistics planning systems?	Descriptive	Expert Interviews	Quantitative
RQ3: How does the current data availability at Bouwvervoer support or restrict the comparative analysis of the manual and AI systems?	Exploratory & Descriptive	Interviews, Internal Data Analysis	Qualitative & Quantitative
KQ1: What frameworks or models exist for selecting logistics and supply chain management KPIs that can be applied to AI-driven and manual systems?	Exploratory	Literature Review	Qualitative
KQ2: How can the importance of KPIs be quantified for logistics planning systems?	Exploratory	Literature Review	Qualitative
KQ3: What are the established data visualization and analysis methodologies in comparing logistics system performance?	Exploratory & Descriptive	Literature Review	Qualitative & Quantitative

Table 2: Research Questions and Knowledge Problems

2.2.3 Research Types

This study on Bouwvervoer's transition to AI-driven logistics planning adopts a mixed approach, combining both exploratory and descriptive research methods. The exploratory component is important in learning AI's application in logistics at Bouwvervoer. It allows for a flexible exploration of patterns, potential benefits, and challenges associated with the AI-driven system. On the other hand, the descriptive aspect of the research focuses on accurately portraying Bouwvervoer's current logistics planning processes. This methodical observation and description are important for setting a clear baseline against which the AI system's performance can be evaluated.

The study incorporates a secondary data analysis through a systematic literature review, including academic journals and industry reports in databases like Scopus and Web of Science. This review focuses on AI applications in logistics, comparative analysis frameworks, and KPI frameworks in transportation logistics to establish a solid foundation for evaluating AI's potential in Bouwvervoer's context. The study uses primary research through qualitative methods, specifically semi-structured interviews, and targeted surveys with Bouwvervoer's stakeholders, IT experts, and operational staff. These methods are chosen for their strength in producing detailed insights into the practical implications of AI integration, including the opportunities it presents. Semi-structured interviews allow for flexible discussions, while surveys provide quantifiable data.

The research methodology directly corresponds with MPSM, with the exploratory phase aimed at identifying and analysing the problem (MPSM Phases 1-3) and the descriptive phase focused on formulating and evaluating solutions (MPSM Phases 4-7). To operationalize, the study conduct interviews with stakeholders. Data from these primary sources are analysed using thematic analysis for interviews and statistical analysis for survey responses, directly providing into the evaluation framework's development. The data-gathering methods for each research question and knowledge problem are outlined and can be seen in *Table 2*.

2.2.4 Limitations of Research Design

While this research's design follows a structured approach using the MPSM, several limitations impact the overall scope of the study. These limitations are particularly significant in the final phase (Phase 7: "Evaluating the Solution"), which cannot be completed with real data due to time, resource, and operational constraints. This limitation affects the ability to fully assess whether the comparative framework developed in this thesis is effective for evaluating AI-driven logistics planning versus manual systems.

Data Availability and Relevance

The study's effectiveness is dependent on the availability and relevance of data. This includes historical and generated data from the manual system and data generated from the pilot implementation of the AI system. Limitations in the scope, depth, or quality of the available data impact the quality of the comparative analysis and the validity of the study's conclusions.

Time Constraints

Given the scope of the research, a time constraint of 10 weeks limit the depth of analysis, specifically in the exploratory parts of the study. The time available for developing and testing the comparative analysis tool or framework need to be increased to explore all potential functionalities and integrations. In a full-scale implementation, Phase 7 would involve testing and iterating on the comparison framework developed in this thesis. It would require integrating real-time data from both the manual system and the AI-driven system and several rounds of feedback from stakeholders. However, the current time frame does not permit such testing, leaving Phase 7 incomplete.

Lack of Full-Scale Implementation

Phase 7 of the MPSM is designed to evaluate whether the solution developed—here, the comparison framework—provides valid and actionable insights. This evaluation typically involve testing the framework in a full-scale operational environment, identifying gaps or areas of improvement, and refining the tool accordingly. However, since the AI system is not fully integrated into Bouwvervoer's operations, the comparison framework cannot be thoroughly tested.

Change Management and Employee Adaptation

While the research aims to consider the human element, particularly regarding change management and employee adaptation, the complexity of human behaviours and resistance to change appear. The study's recommendations are based on theoretical frameworks and stakeholder interviews, which do not fully capture the practical challenges of implementing change in an organizational setting.

Generalization of Results

The results of this research are based on focusing on a single organization's transition to AI-driven logistics planning. Therefore, generalizing results to other organizations is not fully possible.

2.3 Assessment of validity and reliability

Ensuring the validity and reliability of the methodology and findings is crucial in any research study. This section discusses the evaluation of the internal and external validity of the study as well as the reliability of the comparative analysis framework developed to assess the manual system versus the AI-driven system at Bouwvervoer.

It is important to consider both validity and reliability as essential metrics, as shown by Cooper and Schindler (2014). Validity refers to the capacity to which our tools accurately measure what they intend to do, which involves internal and external validity. Internal validity reviews that comparative analysis frameworks accurately capture the effectiveness of Bouwvervoer's manual versus AI-driven logistics systems. We can examine measurements such as KPIs to confirm that they reflect accurate data from the logistics systems. External validity concerns the generalizability of our findings. While our study is tailored to Bouwvervoer's unique operational setup, understanding the limitations of our research's relevancy to other contexts is necessary.

Reliability is equally important and contains the consistency of our research tools in providing results. Reliability stability ensures that repeated measurements under similar conditions produce consistent outcomes, an aspect we mainly focus on in testing the AI-driven logistics system's performance at Bouwvervoer. Equivalence in reliability involves confirming that different tools or observers produce similar results in measuring the effectiveness of logistics planning (Cooper & Schindler, 2014). However, the reliability of the framework cannot be thoroughly evaluated due to the limitations in Phase 7. This phase would typically involve repeated application of the framework in real-world settings to determine whether the results are consistent over time and across different datasets. The lack of real-time operational testing means that potential variability in the results, especially in dynamic logistics environments, has not been accounted for. As a result, while the framework can be theoretically sound, its reliability in practical applications remains uncertain.

3. Literature Review

3.1 KPI Selection

This chapter addresses RQ1 and KQ1 by defining the selection criteria for the KPIs used in the comparative framework. In developing the comparative framework for evaluating the logistics planning systems at Bouwvervoer, we focused on prioritizing KPIs that directly relate to the plan's operational efficiency and have a higher possibility of data availability in future. We chose cost management, employee satisfaction, and customer service quality as main categories to focus. After discussions with the company supervisor and carefully considering the thesis objectives, we selected a total of eight KPIs. Five of the KPIs are focused on transportation efficiency and effectiveness are the following: Capacity Utilization, On-time Delivery Performance, Total Distribution Cost, Turnover per Hour per Driver, and % Empty km's / Total km. These KPIs were chosen from the framework of Iankoulova (2012) to ensure a literature based foundation for our analysis. *Figure 3* shows the focused part of the framework. Iankoulova (2012) used the work of Krauth et al. (2005) to build the framework. We also looked at the framework of Krauth et al. (2005) and they focused on the internal perspective and management point of view which we intend to do so with these five KPIs:

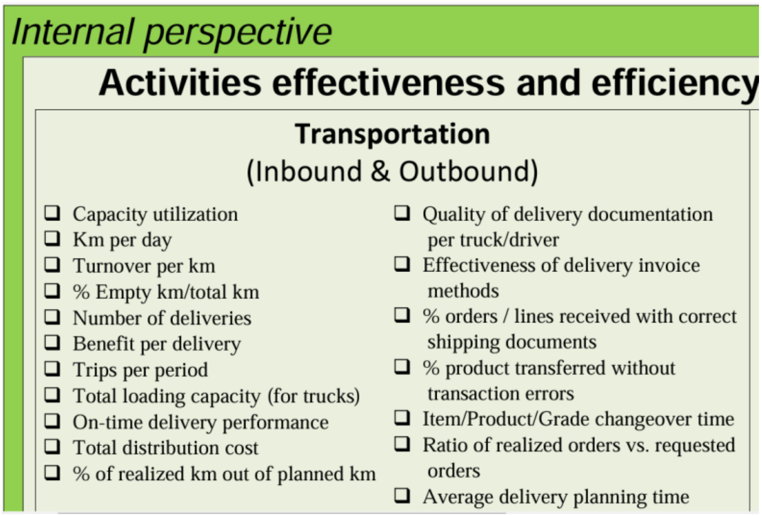

Figure 3: KPI framework of Iankoulova (2012) based on Krauth et al. (2005)

Capacity Utilization: This is a critical KPI for assessing how efficiently transportation resources are used. It measures the extent to which available vehicle capacity is utilized during each trip. By optimizing capacity utilization, companies can reduce the number of trips needed, lower fuel consumption, and minimize operational costs. In Iankoulova's (2012) framework, this KPI directly addresses the operational efficiency of the logistics process, ensuring that resources are used optimally. Poor utilization may result in unnecessary trips, wasted resources, and higher costs.

On-time Delivery Performance: This KPI measures delivery punctuality and is a crucial indicator of service reliability. It reflects the efficiency of logistics operations in meeting customer expectations, which is directly linked to customer satisfaction, which is why it will also be part of customer point-of-view metrics. This KPI helps identify potential bottlenecks or inefficiencies in the delivery process, such as traffic delays, poor route planning, or operational constraints. By focusing on this metric, we can evaluate how well the logistics system maintains reliable delivery schedules.

Total Distribution Cost: This KPI aggregates all the costs associated with the logistics and transportation process, including fuel, labour, maintenance, and administrative costs. This KPI is essential for assessing the financial efficiency of logistics operations. By monitoring distribution costs, we can identify inefficiencies and areas where cost-saving opportunities exist.

Turnover per Hour per Driver: This KPI evaluates drivers' productivity by measuring the revenue generated per hour worked. It offers insights into how effectively human resources are utilized in logistics operations. High hourly turnover indicates efficient use of labour, while low turnover may suggest inefficiencies such as poor route planning, extended idle times, or suboptimal scheduling.

% Empty km's / Total km KPI: This measures the kilometres driven with no cargo, directly indicating inefficiencies in route and capacity planning. Reducing the percentage of empty kilometres can lead to significant cost savings and environmental benefits by lowering fuel consumption and optimizing vehicle usage.

In addition to the KPIs focused on transportation effectiveness and efficiency, we selected four that address employee satisfaction and customer service quality. These KPIs were identified after discussions with Bouwvervoer's management and executive teams, who emphasized the importance of these areas for the company's long-term strategic goals. While transportation metrics are crucial for assessing operational efficiency, the company's leadership stressed that employee well-being and customer satisfaction are equally important for maintaining a competitive and sustainable business model.

The first KPI, **Satisfaction with Working Hours**, measures the level of contentment among employees regarding their working schedules. Employee satisfaction with working hours directly impacts crew morale, ownership of responsibility, and overall productivity. Susanto et al. (2022) show that when employees are satisfied with their working conditions, particularly work-life balance, they tend to be more productive, have higher job satisfaction, and are less likely to leave the organization. High satisfaction levels can reduce turnover, lowering recruitment and training costs. During discussions with management, this KPI was highlighted as a critical factor in improving workforce stability, which ultimately supports the company's operational success and comforts the management and executive teams about its ability to retain its workforce.

The second KPI, **Fulfillment of Specific Driver Requests**, assesses the company's ability to meet its drivers' needs and preferences, such as route preferences or schedule adjustments. This KPI was included because management emphasized that accommodating driver requests contributes significantly to job satisfaction. Meeting these requests improves the work environment. By ensuring drivers feel valued and heard, Bouwvervoer can improve its operational efficiency, as satisfied employees are more likely to perform at higher levels and demonstrate a more significant commitment to the organization.

Order Accuracy is the third KPI which is related to customer service quality. This KPI measures the accuracy with which customer orders are fulfilled, ensuring that the correct products are delivered as requested. Order accuracy is critical to customer satisfaction, as errors in order fulfillment can lead to customer dissatisfaction, increased return rates, and higher operational costs due to corrective actions. Management identified this KPI as fundamental to maintaining high levels of customer trust and loyalty, which should make the management and executive teams feel confident about the company's customer relationships.

Lastly, **On-time Delivery Performance** from the customer's perspective was also selected as a critical KPI. While this overlaps with the transportation-focused version of the same metric, the emphasis here is on the punctuality of deliveries as perceived by the customer. Management and executives stressed the importance of meeting customer expectations about delivery timelines, noting that late deliveries can severely damage the company's reputation and lead to financial penalties or lost business. In today's competitive logistics landscape, customer expectations for timely deliveries are higher than ever, and Bouwvervoer's ability to consistently meet these expectations is critical for maintaining long-term customer relationships.

Finally, we have created the SCOR KPI metrics from eight chosen KPIs. These KPIs have been organized into a three-level hierarchy, as illustrated in the *Figure 4*, covering transportation efficiency, employee satisfaction, and customer service quality. This hierarchical grouping is essential for the upcoming weight analysis of the KPIs and the development of questionnaires to assess their relative importance. We took inspiration from Ayyildiz et al. (2021) SCOR 4.0 performance metrics, which provided a strong foundation for structuring the levels and sub-groups of metrics (SCOR 4.0 model). The three-level structure enables a detailed and layered approach, assigning appropriate weights that are more accessible and performing more accurate evaluations during the analysis phase.

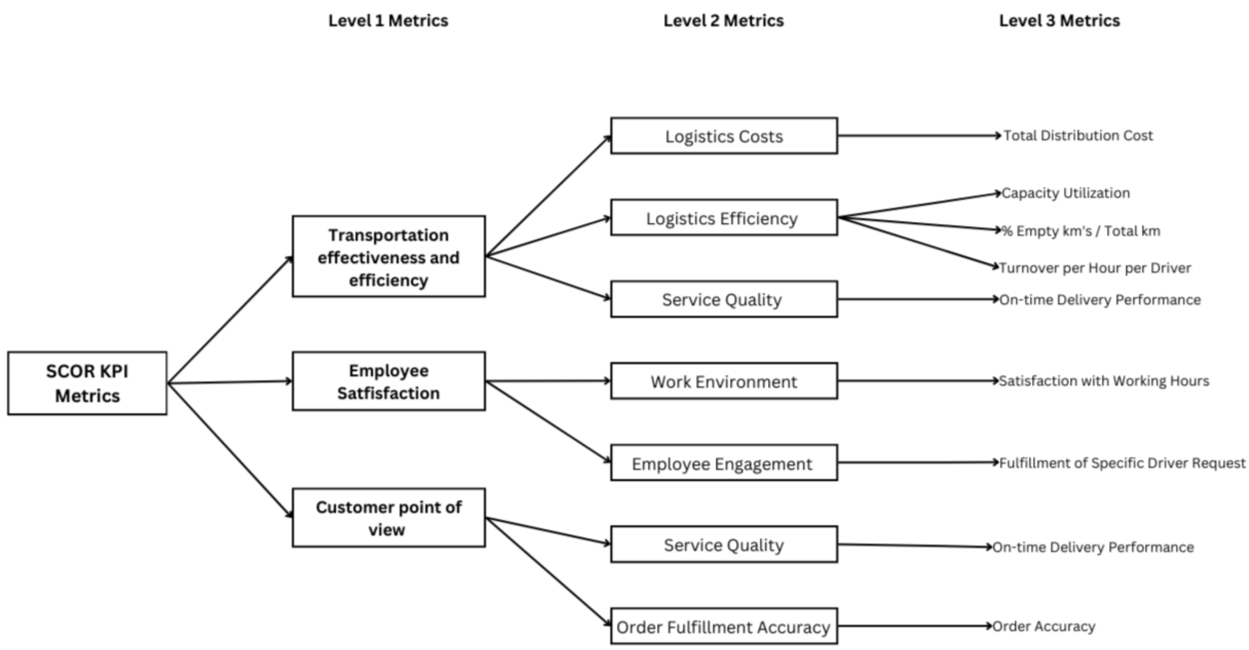

Figure 4: SCOR KPI metrics

3.2 KPI Weighting

Accurately quantifying KPIs is crucial for this thesis. Logistics systems rely on many KPIs to monitor, evaluate, and optimize performance across various dimensions. KQ2 comes into play in this sense, finding out how the quantification of importance of KPIs can be done, helping build the part where the framework's KPI weighting is done appropriately. We intend to review existing methods like AHP, DEA, and TOPSIS, and their variations to understand the subject's industry standards. We learn about how these methods are used in which situations so we can create an approach tailored to our case. Before we create a relationship between articles found about the variety of methods, we will evaluate each item in their groups and then wire relationships.

These three techniques mentioned above have similarities and differences. AHP is a multi-criteria decision-making tool that uses pairwise comparisons and expert judgment to prioritize KPIs, providing a structured framework for decision-making (Sirisawat et al., 2024). DEA is a non-parametric method used for efficiency analysis, benchmarking multiple decision-making units based on their inputs and outputs, making it particularly effective in quantitative settings (Danladi et al., 2024). TOPSIS evaluates alternatives based on their proximity to an ideal solution and offers an approach to handling complex, multi-dimensional decision-making problems (Celik & Akyuz, 2018).

We aim to compare and evaluate the effectiveness, practicality, accuracy, and flexibility of AHP, DEA, and TOPSIS in quantifying KPIs for logistics planning systems. This review seek insights into each method's strengths and limitations by synthesizing findings from 13 studies.

3.2.2 Discussion of Methods

This section discusses methods for quantifying and prioritizing KPIs in logistics planning. We conducted an SLR to determine the most suitable approach for KPI weighting in this study, analysing 13 relevant academic articles.

a. AHP

AHP has been demonstrated to be highly effective in various logistics contexts due to its structured framework for evaluating and prioritizing KPIs. The method allows for a systematic pairwise comparison of criteria (Sirisawat et al., 2024). The literature reviewed highlighted AHP's ability to integrate qualitative and quantitative data, providing insights into KPI performance. For instance, AHP effectively prioritized KPIs such as patient satisfaction and operational efficiency in healthcare logistics, leveraging expert judgments to ensure reliability (Sirisawat et al., 2024). Additionally, AHP's use of the Balanced Scorecard framework further shows its effectiveness in performance measurement (Regragui et al., 2018). AHP is flexible and adaptable to various contexts. The method can be easily tailored to specific needs and integrated with frameworks like the Balanced Scorecard. Studies highlighted AHP's flexibility in the healthcare and supply chain contexts, where it effectively addressed diverse evaluation criteria (Regragui et al., 2018). AHP is practical in terms of its ease of implementation and manageable resource requirements. The method involves straightforward mathematical calculations that can be performed using standard statistical software. AHP's reliance on expert judgments makes it clear that the evaluations depend on practical experience, which makes the method reliable when done with the correct people. Literature shows AHP's practicality in healthcare and supply chain contexts, highlighting again its integration with frameworks like the Balanced Scorecard (Chorfi et al., 2015). AHP's ease of use and low requirement for coding capabilities make it suitable for logistics teams with limited technical expertise. AHP is accurate in quantifying KPIs due to its structured approach and the use of consistency ratios to validate expert judgments. The method's ability to convert qualitative judgments into quantitative scores is functional.

b. DEA

DEA effectively evaluates efficiency across diverse logistics contexts. DEA's ability to handle multiple inputs and outputs simultaneously allows for a comprehensive performance assessment. The method effectively benchmarks performance by comparing entities against best-performing peers, providing clear efficiency scores and actionable insights. For example, DEA was utilized to assess the efficiency of transport corridors for soybean and corn exports in Brazil, identifying the most efficient corridors and guiding future infrastructure investments (Alves Junior et al., 2021). Also, DEA was effectively implemented to improve road transport efficiency in a Spanish retail company, demonstrating significant improvements in operational performance (García-Arca et al., 2018). The main issue in adapting in this thesis is DEA's practicality. DEA requires an understanding of linear programming and access to specialized software, but it does not require extensive data collection. However, the need for specialized software may present a challenge for logistics teams with limited coding capabilities. DEA's accuracy comes from its detailed mathematical framework, which ensures precise efficiency assessments. The method's ability to handle multiple inputs and outputs without requiring specific functional forms improves accuracy. DEA's use in evaluating port and transport efficiency demonstrated its capability to provide detailed and reliable efficiency scores (Danladi et al., 2024). DEA's flexibility is apparent in its adaptability to different contexts and ability to integrate with other models, such as the Network Equilibrium Model. The method can be applied to various logistics scenarios. Literature shows DEA's flexibility in evaluating transport corridors and road transport efficiency (Alves Junior et al., 2021).

c. TOPSIS

TOPSIS ranks and selects alternatives based on their closeness to an ideal solution. TOPSIS is particularly useful in multi-criteria decision-making scenarios, evaluating alternatives against ideal and negative-ideal solutions. In maritime transportation engineering, TOPSIS effectively ranked decision-making problems, providing precise and reliable results even in complex, multi-dimensional scenarios (Celik & Akyuz, 2018). The method's ability to provide precise rankings highlights its effectiveness. While TOPSIS is straightforward in its computational process and the clarity of its results, its practicality in this thesis context is limited due to the requirement for a clearly defined adverse scenario. The method is relatively simple and only requires limited computational resources. Studies have shown its practicality in supply chain contexts, providing clear and actionable rankings (Jothimani & Sarmah, 2014). TOPSIS can be implemented using spreadsheet software, making it accessible for logistics teams with limited technical resources. The problem is that TOPSIS involves identifying ideal and negative-ideal solutions for each criterion, which may not be feasible in logistics settings. The absence of a negative scenario can complicate the application of TOPSIS, making it less practical. TOPSIS accurately ranks alternatives based on their proximity to ideal solutions. The method's ability to provide clear and consistent rankings ensures accuracy even in complex scenarios. Studies have demonstrated TOPSIS's accuracy in various contexts (Celik & Akyuz, 2018). TOPSIS is flexible in handling multiple criteria and adapting to various decision-making scenarios. The simplicity of its implementation process advances its flexibility.

d. Decision

The SLR reveals that AHP, DEA, and TOPSIS each have unique strengths in quantifying KPIs for logistics planning. AHP excels in structured evaluation and expert involvement, DEA excels in comprehensive efficiency assessment and benchmarking, and TOPSIS excels in precise and consistent ranking. All three methods are practical, accurate, and flexible, making them valuable tools for different aspects of logistics performance measurement. Given the context of logistics and the limitations related to coding capabilities and data availability, AHP emerges as the preferred choice. AHP is superior regarding practicality and ease of use considering other methods mentioned. AHP's ability to be implemented with minimal coding requirements and its effective integration with existing frameworks make it highly suitable for limited technical expertise. Therefore, AHP is selected as the optimal method to use in weighting of the KPIs.

3.2.4 Multiplicative AHP

AHP is a comprehensive and structured decision-making technique that helps analyse complex decisions by breaking them down into a multi-level hierarchical structure. AHP incorporates qualitative and quantitative decision-making, allowing decision-makers to compare elements at each hierarchy level.

AHP's primary advantage lies in its use of an intuitive semantic scale to compare and rank alternatives and to determine criteria weights. However, when new alternatives are introduced, AHP can be subject to rank reversal. Additionally, decision-makers' responses sometimes need to be more consistent. AHP addresses this issue by calculating the consistency ratio (CR) and ensuring it is below 10% to confirm judgment coherence (Chorfi et al., 2015). Tu et al. (2023) proposes two methods to avoid the issue: Axiom-Based Prioritization and MAHP. According to the article, axiom-based prioritization performs better than MAHP in handling rank reversal issues, but it is much more complex to apply. Rank reversal occurs when the addition or removal of alternatives affects the ranking of the existing alternatives. Since this is not an essential consideration for our case and alternatives (KPIs) will be consistent during the execution of the thesis, we decided to use the MAHP model.

Ramanathan (1996) and Tu et al. (2023) explain how to use MAHP. The process starts by defining the decision goal and structuring the problem into a hierarchy, including the overall goal, criteria, sub-criteria, and alternatives. The relative importance of each element is then assessed through pairwise comparisons, forming a PCM where alternatives are evaluated against each other. Entries in the PCM, denoted as a_{ij}, represent the factor by which alternative i is preferred over alternative j concerning a criterion. The reciprocal value is used when the comparison is inversed, ensuring that $a_{ij} = {}^{1}/_{a_{ji}}$.

Then calculation of local weights comes. The local weights of alternatives are derived by calculating the geometric mean of each row in the PCM:

$$w_i = \left(\prod_{j=1}^{n} a_{ij} \right)^{1/n}$$

Equation 1: Geometric Mean

These raw weights are then AI-systemlized to sum to one, yielding the AI-systemlized weights w_i' as follows:

$$w_i' = \frac{w_i}{\sum_{k=1}^{n} w_k}$$

Equation 2: AI-systemlization

This AI-systemlization facilitates the comparability of weights across different criteria.

Next step is the aggregation of weights for multiple criteria. The purpose of the aggregation process in AHP is to combine weights across different levels of a decision hierarchy, typically when there are overarching criteria and more specific sub-criteria. This aggregation is essential in complex decision-making scenarios. However, in scenarios where the decision structure comprises a single layer of criteria, the weights derived from the initial PCM calculations provide a direct measure of each criterion's relative importance without further aggregation. However, in our case, we have a multi-level hierarchy (Level 1, 2, and 3) also as shown in *Figure 4*. Therefore, once we calculate weights for each level, we can aggregate them using geometric mean aggregation to get the final weights for the detailed metrics in Level 3.

These weights conclusively represent the logistic experts' assessments and preferences regarding the importance of the evaluated KPIs. This simplicity in structure might give a more straightforward application of the AHP but also provides clarity and focus on decision-making.

Also, the consistency of each PCM can beassessed using the CR, which is calculated as:

$$CR = \frac{CI}{RI}$$

Equation 3: Consistency Ratio

Here, CI is the Consistency Index derived from the eigenvalue method or approximation, and RI is the Random Index, which depends on the number of alternatives. A CR less than 0.1 is typically considered acceptable, indicating that the matrix has an acceptable level of consistency. If the CR exceeds this threshold, the judgments within the PCM may require reassessment and adjustment. Adjustments may involve altering some evaluations to reduce subjectivity and increase logicality.

3.3 Analysis and Visualization Methods

This section answers KQ3, introduces several statistical and analytical methods to assess individual and composite KPIs, focusing on approaches that combine, compare, and visualize performance metrics. The review also incorporates principles of data visualization, as effective data representation is critical to translating statistical findings into understandable, actionable insights.

3.3.1 Analysis Methods

Weighted Composite Score

In Devellis (2016) on measurement scales, composite scores are introduced to combine multiple items into a single measure, enabling the capture of complex, theoretical constructs that individual items alone might not fully represent. The WCS, a composite score with different weighted metrics, aggregates multiple performance indicators into a single overall score that reflects the relative importance of each KPI. This single score can then identify general trends, make comparisons, or simplify complex data into actionable insights.

Descriptive and Comparative Statistics for Individual KPIs

Descriptive statistics form the basis for understanding data distributions and tendencies for each KPI in isolation. As Dong (2023) emphasizes, descriptive statistics is fundamental in statistical analysis, providing essential methods to summarize and organize data variability in meaningful ways. Metrics such as mean, median, variance, and standard deviation are crucial to comprehending the typical performance levels and variability across KPIs. For instance, standard deviation quantifies the spread of values around the mean. This allows analysts to detect KPIs where the AI system offers more consistent results than the manual system or where optimization may be required to reduce variability. Dong (2023) highlights that these basic descriptive measures serve as a "building block" for more complex statistical analyses, making them essential to initial data assessments and comparative analysis stages.

Such comparative insights add a layer of individual KPI analysis that complements the overall composite score. Therefore, descriptive statistics provide Bouwvervoer with clarity and precision in the initial KPI assessment.

Weighted Difference Analysis

Weighted Difference Analysis evaluates the difference in performance between two systems by adjusting each KPI difference according to its assigned weight. Given its relative importance, this approach provides a perspective by highlighting how each KPI individually contributes to the overall performance gap. The methodology is particularly beneficial in comparative frameworks, where organizations must understand the impact of adopting new systems in specific operational areas.

The study by Soriano-Gonzalez et al. (2023) on KPIs in urban mobility logistics for sustainable cities offers relevant support for this approach. In examining critical KPIs for Barcelona's mobility logistics, the authors underscore the need for a framework that allows each KPI to be weighted according to its impact on strategic goals, a fundamental principle in WDA. The authors' use of the Eltis method, a standardized KPI measurement framework, supports the need for weighted approaches in performance comparison. In this context, WDA can be seen as an extension of these standardized frameworks, offering a way to incorporate KPI-specific insights into a comprehensive performance comparison.

3.3.3. Visualization Methods

Initial Knowledge

Data visualization is a critical tool in transforming data into actionable insights. It helps decision-makers to quickly identify patterns, trends, and relationships that may remain hidden. Visualization allows datasets to be understood through visuals, where information is presented in a digestible, accessible format, aiming to reduce the viewer's cognitive load and improve interpretability. Paczkowski (2021) emphasizes that effective data visualization relies on principles that help perceive and understand visual representations. One critical goal of data visualization is to reduce or eliminate "chartjunk"—outside elements in a graph that do not contribute to the primary message. Tufte (2016) popularized this concept, arguing that visualization should be as clear and direct as possible to facilitate understanding. Elements like excessive colors, unrelated graphics, or complicated legends can confuse the main insights. Well-designed visualizations put away distractions, focusing the viewer's attention on the data's key message. The visualization design must also consider the data characteristics, particularly the continuity (discrete or continuous) and dimensionality (single or multiple series) of the data. This provides that the chosen visualization method best represents the data's structure, effectively revealing relationships, trends, and distributions.

Data Size Considerations

The size of the dataset plays a significant role in determining which visualization techniques are suitable and how effective they can be in obtaining insights. According to Paczkowski (2021), data visualization strategies vary based on dataset size, ranging from Small to Big Data applications. Smaller datasets, often under 10^6 data points, allow for detailed analysis using traditional visualization tools, while larger datasets require adaptations to handle complexity and volume.

For our focus on logistics KPIs, it is reasonable to categorize the data as small to medium based on the expected number of observations. In logistics performance tracking, typical datasets may contain fewer than 10,000 observations, aligning with the size range that human visual perception can process effectively. The human eye can process a maximum of 10^6 to 10^7 points, making large-scale visualizations challenging to interpret. The dataset sizes categorized as small to medium comfortably fit into this perceptual range, indicating that traditional visualization techniques will work for our case without extensive modifications. Given this small to medium data scale, techniques like scatter plots, line charts, boxplots, and bar charts are likely acceptable.

Visualization Flowchart

With the data size identified as small to medium, we can now move on to building a flowchart, constructed to streamline the selection of visualization techniques for KPIs.

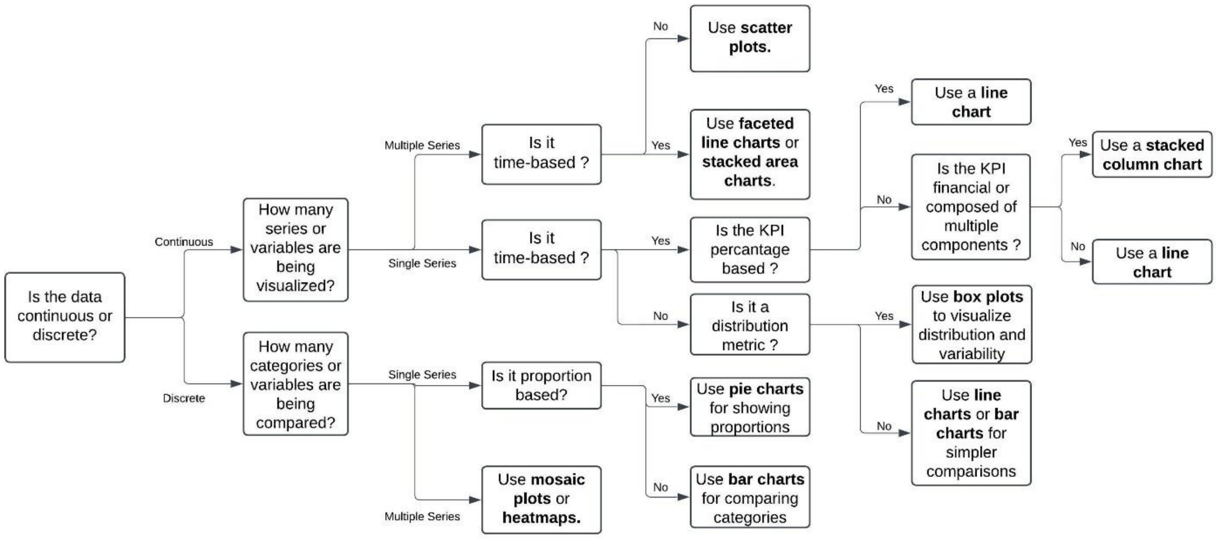

Figure 5: Visualization Technique Selection Aid Flow Chart

This flow chart *(Figure 5)* guides users through decisions based on data attributes, such as continuity (continuous or discrete), the number of series (single or multiple), and the presence of a time dimension. This approach simplifies the visualization selection process. It ensures that each KPI is represented in an interpretable way, allowing decision-makers to draw meaningful conclusions and take informed actions.

3.4 Conclusion

This chapter has addressed RQ1 by identifying the most relevant KPIs for evaluating logistics planning performance at Bouwvervoer and KQ1 by establishing a structured framework for KPI selection. The eight chosen KPIs, five focused on transportation effectiveness and efficiency, and three addressed employee satisfaction and customer point of view, provide a basis for comparing manual and AI-driven planning systems. Additionally, the chapter explored KQ2, determining how KPIs can be quantified and weighted through multi-criteria decision-making techniques. The systematic literature review of 13 articles examined AHP, DEA, and TOPSIS as potential methodologies for KPI weighting. The comparative analysis demonstrated that while each method has distinct strengths, AHP emerged as the most suitable choice due to its structured evaluation process, practical implementation, and integration with existing decision-making frameworks. MAHP was introduced as a tool to improve consistency in KPI prioritization while addressing limitations such as rank reversal. Finally, KQ3 was answered by selecting an approach to analysing and visualizing KPI performance. WCS and WDA were introduced as key methods for evaluating logistics efficiency, along with descriptive and comparative statistical techniques. A visualization selection flowchart was developed to guide the representation of KPI data.

This chapter set the foundation for developing a comparative analysis framework by defining KPIs, selecting a weighting methodology, and analytical tools for performance assessment.

4. Conceptual Framework Creation

This chapter presents the development of the conceptual framework designed to compare the manual logistics planning system with the AI-driven system at Bouwvervoer. The framework is structured around KPIs, a MAHP for KPI weighting, and analysis tools that assess system performance.

4.1 Selection of KPIs and Causes of (Under) Performance

KPIs have been selected based on their relevance to Bouwvervoer's operational goals and data availability from manual and AI-driven systems. Each KPI addresses a critical aspect of logistics planning, including efficiency, resource management, reliability, and cost-effectiveness. By measuring and analysing these KPIs, we can provide a comprehensive assessment of the performance of both systems and offer data-driven insights for decision-making. Appendix 2 summarizes the chosen KPIs, detailing their definitions, significance, how they are measured, and their practical applications in the logistics process.

In this chapter we also discuss possible causes for unsatisfactory performance in the chosen KPIs. One key factor contributing to underperformance in capacity utilization at Bouwvervoer is inefficient load planning. When vehicle capacity is not maximized, operational costs increase as more trips are required to deliver the same volume of goods. Additionally, incomplete or unclear data regarding cargo amounts in the current systems may worsen the issue, making it challenging to optimize truck loading fully.

Delays in deliveries cause underperformance in on-time delivery performance. The causes of these delays can include suboptimal route planning, traffic congestion, and unexpected disruptions, such as accidents or weather conditions. Additionally, insufficient coordination between scheduling and actual delivery execution can result in bottlenecks.

Underperformance in the total distribution cost KPI is often the result of inefficient resource allocation. High fuel costs, underutilized labor, and poor vehicle maintenance management can contribute to escalated distribution costs. Another factor driving costs up is the increase in empty or underloaded trips, which wastes fuel and driver hours.

Lower turnover per hour per driver may result from inefficient scheduling, poorly optimized routes, or extended idle times during deliveries. In some cases, drivers may face delays due to traffic, loading issues, or coordination challenges, which reduce their productivity. Another potential issue is the inability to match drivers with trips that maximize their productivity. Poor communication and coordination between dispatchers and drivers also can contribute to lower performance in this KPI.

A high percentage of empty kilometers indicates underperformance in route optimization. This issue often arises when trucks are sent out without fully utilizing their cargo capacity or when return trips are made without coordinating for return loads. The inability to match available shipments with return routes results in wasted fuel and a chance to optimize, increased operational costs, and higher carbon emissions.

Employee dissatisfaction with their working hours can be attributed to inconsistent scheduling, long shifts, or a lack of flexibility in adapting personal preferences. Underperformance in this area may be caused by a lack of communication between management and drivers, where employees feel their needs and preferences should be considered. With structured surveys or feedback mechanisms to regularly assess employee satisfaction, Bouwvervoer can avoid upsetting its workforce.

Order accuracy can be improved due to communication breakdowns between warehouses, logistics planners, drivers, and inefficiencies in order fulfillment processes. Errors in packing or documentation can lead to incorrect deliveries, which frustrates customers and increases operational costs due to returns and re-delivery efforts.

4.2 Data Availability and Gaps

A critical factor in evaluating the manual and AI-driven logistics planning systems is the availability of reliable data for each KPI. However, as detailed in Appendix 3, significant gaps exist in the dataset, making it impossible to conduct a direct, data-driven comparison between the two systems. The lack of data affects the manual and AI-driven systems, limiting the feasibility of performing a fully quantitative assessment. While some datasets provide partial information for the manual system, Appendix 3 shows many crucial elements that are missing or incomplete.

Drawing empirical conclusions on AI adoption benefits remains unfeasible without a unified dataset covering both systems. As a result, this study relies on a conceptual comparison framework rather than an empirical one, focusing on theoretical performance assessment methods rather than direct data analysis. Addressing these data gaps would require significant efforts in data collection, integration, and validation before a fully operational comparison could be conducted.

4.3 MAHP Questionnaire

This section introduces the MAHP methodology, which assign weights to the selected KPIs presented earlier. The MAHP is a decision-making tool that enables us to structure complex problems into a hierarchy and prioritize elements by calculating relative weights through pairwise comparisons. This approach is particularly useful in our context as it provides a quantitative means to evaluate and compare the relative importance of different KPIs. The MAHP plays role in deriving the weights of the eight KPIs identified in Section 3.1 and detailed in Section 4.1. These weights reflect the strategic priorities of Bouwvervoer's stakeholders and provide insight into which areas are most essential. The decision-makers involved in this process, including logistics planners and operational managers, will provide their insights through a structured MAHP questionnaire. This process involves pairwise comparisons of the selected KPIs using a standardized 9-point scale, enabling us to quantify the relative importance of each KPI.

4.3.1 Data Collection

The MAHP questionnaire was distributed to a group of key stakeholders within Bouwvervoer. These participants were selected based on their roles and expertise in the company’s logistics operations. The questionnaire was distributed via email as a structured digital survey, allowing participants to complete the pairwise comparisons conveniently. Given the technical nature of the MAHP process, each respondent was provided with detailed instructions on how to evaluate the KPIs using the nine-point pairwise comparison scale. A total of six participants from Bouwvervoer, each with more than ten years of experience in logistics planning, were selected to participate in the MAHP survey. Their roles and years of experience make them well-suited to evaluate the performance metrics that are most relevant to Bouwvervoer’s operations. Response rate of the participants was 100%.

4.3.2 MAHP Design

The MAHP questionnaire is structured to evaluate the relative importance of the selected KPIs by organizing them into a multi-level hierarchy. This approach allows for a systematic breakdown of the decision-making process, ensuring each metric is weighted according to its significance. The hierarchical structure consists of three levels: Level 1 includes three main categories of metrics, Level

2 expands into specific sub-metrics, and Level 3 contains the most detailed KPIs that directly reflect logistics performance.

The hierarchical structure, as shown in *Figure 4* in Chapter 3, begins with Level 1, which consists of three main categories:

1. Transportation Effectiveness and Efficiency
2. Employee Satisfaction
3. Customer Point of View

In Level 2, each of these categories is further broken down into more specific sub-metrics, such as Logistics Costs, Logistics Efficiency, and Service Quality under Transportation Effectiveness, and similar expansions for the other categories. Finally, Level 3 consists of detailed KPIs like Total Distribution Cost, Capacity Utilization, and On-time Delivery Performance, which are linked to their respective sub-metrics and categories.

For each level in the hierarchy, pairwise comparisons are conducted to determine the relative importance of the metrics under consideration. For example, within Transportation Effectiveness, the importance of Logistics Costs is compared to Logistics Efficiency and Service Quality. Respondents rate these comparisons using a logarithmic scale, reflecting how much more critical one metric is relative to another. A score of 1 represents equal importance, while scores like 3, 5, 7, and 9 reflect increasing degrees of importance for one metric over another. Intermediate values (2, 4, 6, 8) are also used in this questionnaire for finer distinctions.

This comparison process is repeated at each level of the hierarchy. At Level 3, the detailed KPIs such as Capacity Utilization, % Empty km's / Total km, and Turnover per Hour per Driver are compared in pairs within their sub-metric categories. The exact process is applied to the Employee Satisfaction and Customer Point of View categories, ensuring all KPIs are evaluated about their peers. The final questionnaire given to participants can be seen in Appendix 4. All calculations done and explained in Chapter 4.3.3 can be found in Appendix 5.

Matrices

Once the pairwise comparisons have been completed, the results are used to form matrices at each level. A matrix for a particular level compares all elements within that level against each other. For instance, in Level 2, under Transportation Effectiveness, the matrix include logistics costs, efficiency, and service quality comparisons. This matrix captures the relative importance of each sub-metric in achieving transportation effectiveness.

Similarly, matrices are constructed for other main categories in Level 2 and for the detailed KPIs in Level 3. These matrices serve as the basis for calculating the metrics' weights.

Calculation of Weights

After the pairwise comparison matrices were constructed, the geometric mean method was used to calculate the local weights for each element. The geometric mean of each row in the pairwise comparison matrix was calculated to derive the raw weights for each criterion. These raw weights were then AI-systemlized so that the sum of the weights for each level equaled one. This process was carried out for all levels of the hierarchy, resulting in the AI-systemlized weights for Level 1, Level 2, and Level 3 elements.

Aggregation of Weights Across Levels

The next step involved aggregating the weights across levels to obtain the final priority weights for the detailed KPIs in Level 3. This was done by multiplying the weights at each level from Level 1 to Level 3. For example, the final weight for Total Distribution Cost was calculated by multiplying the weight of Transportation Effectiveness (Level 1), the weight of Logistics Costs (Level 2), and the weight of Total Distribution Cost (Level 3). This process ensured that the priorities of the detailed metrics were influenced by their corresponding higher-level categories, providing a comprehensive evaluation of the KPIs.

Handling of Multiple Weights for On-time Delivery Performance

A unique situation arose with On-time Delivery Performance (OTDP), which appeared under Transportation Effectiveness and Customer Point of View. This resulted in two distinct weights for OTDP. To address this, two approaches were considered: calculating a weighted average of the two weights to produce a single final weight or keeping the two weights separate to reflect the different impacts OTDP has in each category. Both approaches were evaluated, and the decision on which to apply was based on the specific focus of the analysis. Because the thesis focuses on practical applications we chose to go with the first approach.

Consistency Check

Finally, the CR was calculated for each pairwise comparison matrix to ensure that the judgments were consistent and logical. The CR was determined by first calculating the largest eigenvalue of the matrix and then using the CI and the RI, which was taken from Saaty (1987). Matrices with less than three elements had a RI of zero, therefore were not calculated. All CR values were below the accepted threshold of 0.1, indicating that the judgments made by the participants were consistent and no further adjustments were needed.

4.3.3 Results and Analysis

Applying the MAHP resulted in a detailed evaluation of the KPIs relevant to the decision-making framework. We used Excel to carry out the calculation, which can be seen in Appendix 5. Structuring the problem into a hierarchy, performing pairwise comparisons, and aggregating the results across levels gave the final priority weights for each KPI. These weights explain the relative importance of each KPI within the context of operational and strategic objectives.

Final Weights of KPIs

The final weights of the KPIs are summarized in *Table 3*. These weights reflect the judgments made by the participants, aggregated through the MAHP process to give a clear ranking of the KPIs based on their relative importance.

KPI	Weight
Total Distribution Cost	0.12302
Capacity Utilization	0.12809
% Empty km’s / Total km	0.06698
Turnover per Hour per Driver	0.07520
On-Time Delivery Performance	0.21741
Satisfaction with Working Hours	0.07064
Fulfilment of Specific Driver Request	0.10250
Order Accuracy	0.21616

Table 3: KPI Weights

These final weights represent the priorities assigned to each KPI by the participants, with the sum of the weights totaling 1.

Consistency of Judgments

Throughout the MAHP process, the CR was calculated for each pairwise comparison matrix to ensure that the participants' judgments were consistent. In all cases, the CR values were below the acceptable threshold of 0.1, indicating that the participants' judgments were logical and coherent. This consistency check further validated the results' reliability, ensuring no significant adjustments were needed.

Insights on KPI Prioritization

The following bar chart *(Figure 6)* visually represents the assigned weights for each KPI, highlighting their relative importance in Bouwvervoer's logistics planning. Additionally, the comparative heatmap *(Figure 7)* displays these KPIs within their categorized framework, emphasizing the strategic distribution of priorities across Transportation Effectiveness, Employee Satisfaction, and Customer Point of View.

As seen in the bar chart, On-time Delivery Performance and Order Accuracy emerge as top priorities, aligning with Bouwvervoer's emphasis on customer satisfaction and operational reliability. KPIs related to cost efficiency and resource utilization, such as Total Distribution Cost and Capacity Utilization, are represented in blue, signifying their role in Transportation Effectiveness. Meanwhile, KPIs under Employee Satisfaction are shown in green, reflecting moderate but essential importance to employee well-being and work-life balance.

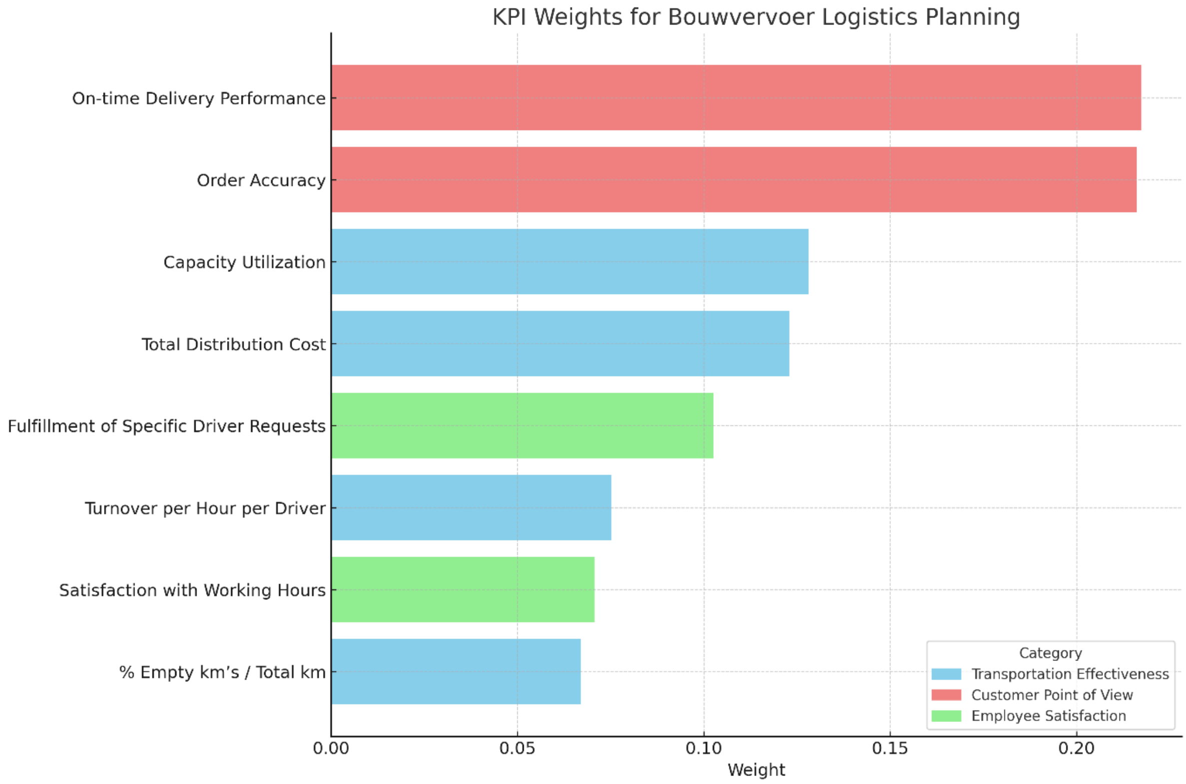

Figure 6: Bar Graph of KPI Weights

The comparative heatmap offers a structured view of KPI priorities within their respective categories. In the Transportation Effectiveness category, KPIs such as Capacity Utilization and Total Distribution Cost are weighted significantly, supporting their essential roles in achieving cost-efficient and resource-optimized operations. By contrast, Customer point-of-view KPIs (On-time Delivery Performance and Order Accuracy) stand out significantly, reflecting a solid strategic focus on customer satisfaction and service reliability. The Employee Satisfaction category, represented by metrics such as Satisfaction with Working Hours and Fulfilment of Specific Driver Requests, shows a moderate priority level, indicating the company's balanced approach to maintaining workforce well-being alongside operational goals. This hierarchical visualization reveals how Bouwvervoer's stakeholders have layered priorities, with the highest focus on customer-facing metrics followed by cost and efficiency considerations.

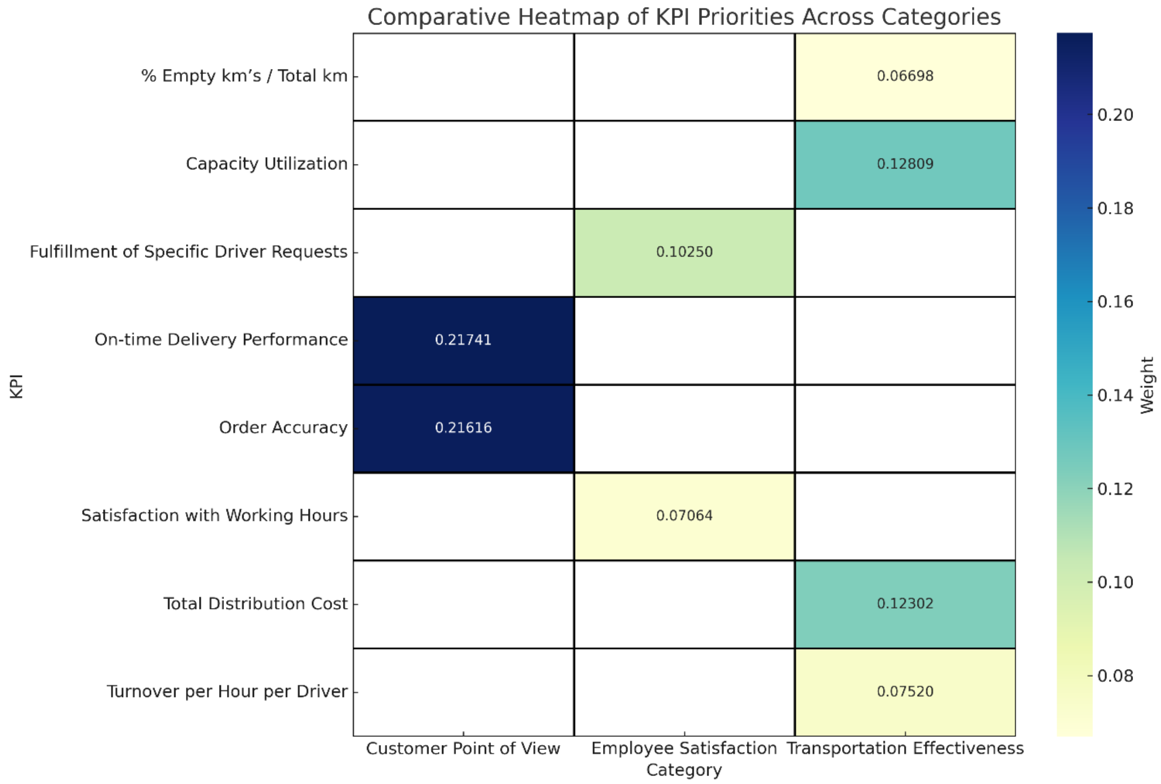

Figure 7: Heatmap of KPI Weights

Implications for Decision-Making

The final results provide decision-makers with clear insights into which KPIs should be prioritized to improve operational performance and customer satisfaction. With **On-time Delivery Performance** and **Order Accuracy** emerging as top priorities, organizations can focus on improving these areas to ensure reliable service delivery. Optimizing **Capacity Utilization** and managing **Total Distribution Costs** will further enhance efficiency and profitability in transportation operations.

4.4 Analysis Tools

This chapter builds upon the literature discussed in Section 3.3 to establish a structured approach for analysing and visualizing KPI performance. WCS, as introduced by Devellis (2016), is used to aggregate multiple KPIs into a single performance measure to assess overall system performance. Descriptive and Comparative Statistics, as outlined by Dong (2023), serve as a foundation for understanding data distributions and variability, offering insights into the consistency of performance across different KPIs. WDA, discussed by Soriano-Gonzalez et al. (2023), enables a detailed examination of performance gaps between the manual and AI-driven systems by adjusting KPI differences according to their assigned importance. Additionally, data visualization techniques are crucial in improving the interpretability of KPI performance metrics. The principles outlined by Paczkowski (2021) and Tufte (2016) guide effective visual representation, ensuring that complex data is presented in an accessible and meaningful manner.

4.4.1 Weighted Composite Score

The WDA is calculated by combining multiple metrics, each adjusted by a set weight that reflects its relative importance. Each measure (or KPI) is first multiplied by its assigned weight to compute this score. These weighted values are then summed to yield a final composite score, representing an overall performance measure. Formally, if x_i represents each metric and w_i its corresponding weight, the Weighted Composite Score S can be calculated by means of *Equation 4*.

$$S = \sum_{i=1}^{n} w_i \times x_i$$

Equation 4: Weighted Composite Score

Where n is the total number of metrics.

This score provides an accessible, high-level overview of each system's performance, highlighting which system offers a stronger alignment with the company's logistical and operational goals. Scenario analyses will further explore how varying performance conditions and weight adjustments affect this score, offering insights into potential sensitivity to different KPIs.

4.4.2 Descriptive Statistics

The descriptive statistics approach involves calculating statistical measures for each KPI, including the mean, median, variance, and standard deviation. These statistics reveal patterns in performance, highlighting areas where each system may demonstrate greater consistency or volatility.

Process and Key Metrics:

- **Mean**: Provides the average value of each KPI, offering a baseline for performance comparison.
- **Median**: Shows the central tendency, reducing the impact of outliers and providing a stable comparison point.
- **Variance and Standard Deviation**: These metrics quantify the spread or variability around the mean, identifying KPIs with inconsistent performance or a higher likelihood of fluctuation. Lower variability suggests more excellent reliability, which is often desirable for KPIs linked to operational predictability.

Descriptive statistics can identify differences between the manual and AI-driven systems for each KPI beyond their average values. For example, while both systems may achieve similar average on-time delivery rates, the standard deviation could highlight one system as more reliable.

4.4.3 Weighted Difference Analysis

WDA examines how each KPI individually contributes to the overall performance gap between the manual and AI-driven systems. The analysis calculates the weighted difference for each KPI by determining the absolute difference in each KPI's value between the manual and AI-driven systems. Then, this difference is multiplied by the KPI's weight.

The weighted difference for a given KPI Di is expressed as:

$$D_i = w_i \times \left| x_i^{AI} - x_i^{Manual} \right|$$

Equation 5: Weighted Difference

Where w_i is the weight of the i-th KPI, and x_i^{AI} and x_i^{Manual} represent the KPI values for the AI-driven and manual systems, respectively.

For instance, if "On-time Delivery Performance" has a substantial weighted difference, it may suggest that the AI system improves punctuality over the manual system. Conversely, suppose a KPI like

"Total Distribution Cost" shows little difference. In that case, the analysis suggests that both systems have similar cost efficiency. This analysis helps prioritize specific KPIs for operational adjustments, especially in areas where the weighted difference is significant. For instance, a high-weighted difference in "Capacity Utilization" might start actions to optimize AI algorithms for better resource management.

4.4.4 Visualization of KPIs

Each KPI is paired with a visualization technique selected using the flowchart developed in section 3.3.3. This flowchart guided the decision-making process for each visualization, providing alignment with each KPI's data characteristics, such as continuity, temporal nature, and series type.

The *Figure 9* below outlines the chosen visualization techniques for each KPI, along with the rationale for each choice:

KPI	Visualization Technique	Rationale
Capacity Utilization	Line Chart	A line chart works because capacity utilization is a continuous, percentage-based metric. This visualization makes it easy to observe efficiency changes between the two systems
On-time Delivery Performance	Line Chart	As a time-based metric, a line chart captures trends in delivery punctuality, allowing for a clear comparison of reliability over time
Total Distribution Cost	Stacked Column Chart	Since distribution costs include multiple components, a stacked column chart allows a breakdown of cost elements over time, making it easier to identify specific areas for potential cost savings
Turnover per Hour per Driver	Bar Chart	Since this KPI is a straightforward comparison of productivity rates per driver, a bar chart will allow for an easy comparison of hourly turnover
% Empty km's / Total km	Line Chart	If this metric is tracked over time, a line chart can capture trends and patterns in route efficiency. This visualization shows whether the AI-driven system leads to a reduction in empty kilometres over time
Satisfaction with Working Hours	Pie Chart	A pie chart displays the proportion of employee satisfaction levels and provides a clear view of the workforce perspective regarding each system's impact on work-life balance
Fulfillment of Specific Driver Requests	Pie Chart	A pie chart displays the percentage of fulfilled driver requests across both systems
Order Accuracy	Line Chart	A line chart captures the trend in order accuracy over time

Table 4: Chosen Visualization Techniques for KPIs

4.5 Conclusion

This chapter builds upon the foundation established in Chapter 3, where the KPIs were initially identified by refining their definitions, analyzing potential causes of underperformance, and setting up the necessary analysis tools. Through development of a conceptual comparison framework, this chapter establishes a methodology for evaluating the manual and AI-driven logistics planning systems at Bouwvervoer. This chapter has answered RQ2 by determining the relative importance of each KPI through the MAHP. The established KPI weights is the basis for future comparisons and performance assessments within the framework. Additionally, this chapter has addressed RQ3 by assessing data availability and gaps. The analysis in Appendix 3 highlights the significant missing data in manual and AI-driven systems, making an empirical, data-driven comparison unfeasible at this stage. Furthermore, this chapter extends the analytical framework introduced in Chapter 3 by integrating WCS, WDA, and Descriptive Statistics. Applying these techniques provides that the framework is prepared to provide quantifiable insights into system performance. The selection of visualization techniques, guided by the literature, further improves the interpretability of KPI results.

Overall, this chapter establishes the foundational elements of the conceptual comparison framework, setting the stage for the implementation of a practical solution. By defining KPIs, assigning weights, and selecting appropriate analysis tools, this framework is ready to be translated into an actionable tool. The next step involves creating a Power BI dashboard to apply the framework and visualize performance metrics.

5. Solution Implementation

This chapter presents the implementation of the conceptual framework through a roadmap and an interactive dashboard. These tools form the core of the comparative evaluation framework designed to analyze and compare the performance of manual and AI-driven logistics planning systems at Bouwvervoer. The roadmap provides a structured visual representation of the process, while the dashboard is a practical interface for managers to interact with performance data. Together, these solutions offer an approach to decision-making, helping Bouwvervoer to evaluate which system works better with its operational objectives.

5.1 Roadmap for System Comparison

The roadmap in *Figure 8* outlines the sequential data flow and activities required for the comparative framework. It serves as a guiding tool, addressing all relevant components of the comparison process. The roadmap starts with data input from both the manual planning system and the AI-driven system, facilitated by a unified database that ensures consistency and comparability.

Subsequently, the planning systems generate logistics plans independently. The manual system relies on human decision-making supported by digital tools, while the AI system automates route and resource planning. Each system's performance is evaluated based on eight predefined KPIs. The roadmap incorporates visualization and analysis stages where individual KPI performance is observed. WCS and WDA calculations are subsequently performed to aggregate performance metrics into interpretable scores. These calculations allow managers to determine whether the AI-driven system offers superior performance compared to the manual system. Finally, the roadmap reaches a final, which is a comparison of the two systems' KPIs together.

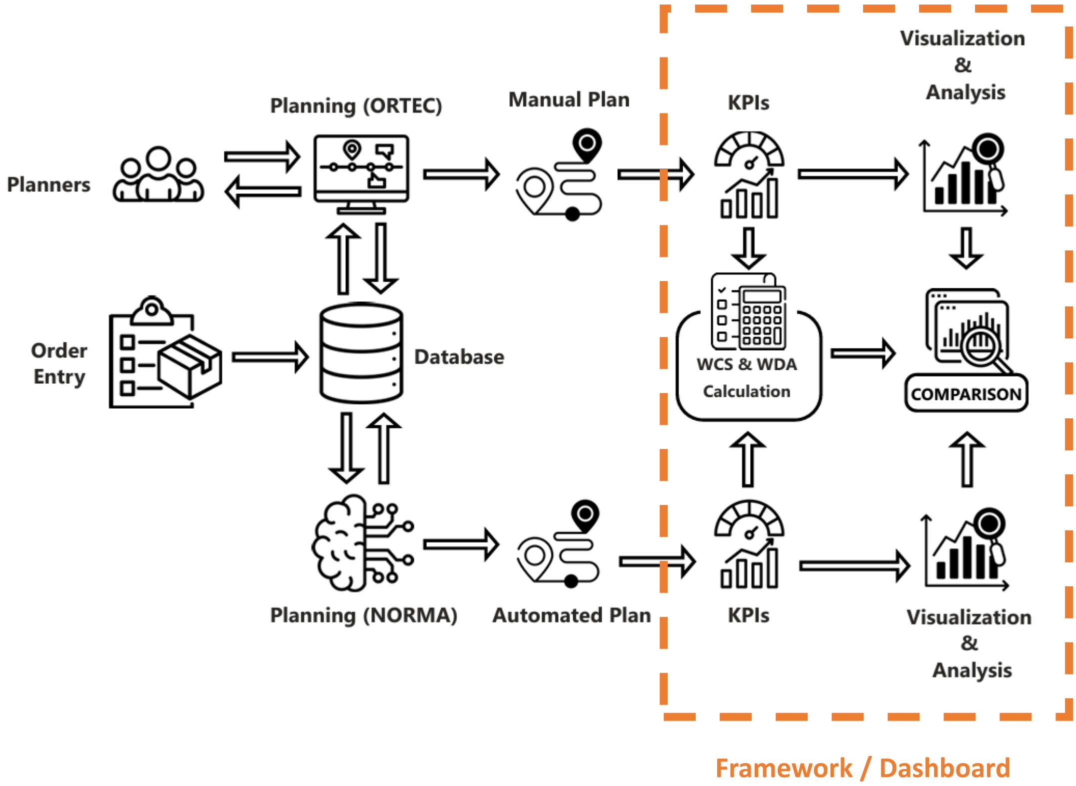

Figure 8: Comparison Roadmap

5.2 Dashboards as a Central Tool for Visualization

Dashboards have emerged as a central tool for visualizing logistics system performance. Piela (2017) highlights the importance of dashboards in presenting KPIs to multiple stakeholders, especially in environments that require real-time updates. Dashboards offer flexibility, as they can be customized for different organizational levels—operational, tactical, and strategic—allowing users to access relevant KPIs appropriately. This adaptability is essential in logistics, where real-time visibility of performance metrics can influence daily operational decisions, while static views may support long-term strategic planning.

Dashboard design should prioritize simplicity and clarity to avoid overwhelming users with too much data. Piela (2017) also emphasizes the need for interactive and dynamic visualizations, which allow users to drill down into data and obtain more granular insights. Proper visualization techniques can vastly improve decision-making processes in logistics by making KPIs more accessible and easier to understand.

The Power BI dashboard developed for this study operationalizes these principles. Explicitly designed for Bouwvervoer's managers, the dashboard integrates simplicity and interactivity, ensuring its usability for decision-makers at the strategic and operational levels.

5.3 Mock-Up Data

To facilitate the development and demonstration of the Power BI dashboard, a seven-week mock-up dataset was created to simulate the performance of the manual and AI-driven logistics planning systems. This synthetic dataset was necessary due to the data limitations outlined in Appendix 3, which revealed significant gaps in the availability of accurate operational data—particularly from the manual and still-developing AI systems. The purpose of the mock-up data was to enable the implementation of the conceptual comparison framework in a functional environment, showcasing its analytical and visualization capabilities.

The mock-up data was structured to reflect two key dynamics. First, the AI-driven system was designed to outperform the manual system across all eight KPIs. Second, the mock-up data incorporated higher variability in the weekly performance of the AI-driven system. This design choice reflects a common trade-off in automated systems. While they often achieve superior overall results, they may be more sensitive to external factors, leading to more significant fluctuations in performance over time. In contrast, the manual system displayed more stable but lower performance trends. While the mock-up data is artificial and do not fully reflect the complexities of actual operations, it is a valuable tool for demonstrating the dashboard's functionality and potential.

5.4 Dashboard Solution

The Power BI dashboard is an interactive decision-support tool designed to evaluate and compare the performance of manual and AI-driven logistics planning systems at Bouwvervoer. The dashboard integrates KPI monitoring, weighted analysis, and performance comparison within a structured architecture, ensuring that key stakeholders can easily interpret and act upon logistics performance insights. The system has been developed explicitly for managerial use.

The dashboard is structured as a multi-layered system that follows a logical data pipeline, enabling tracking, analysing, and comparing performance indicators. It is built upon four key functional components: KPI definition and weighting, individual KPI analysis, WCS & WDA calculation, and system-wide comparison.

5.4.1 Navigation and KPI Overview

At the dashboard's core, a structured navigation panel with the same design as the Comparison Roadmap (*Figure 8*) shown in Section 5.1 allows users to interact with various analysis pages. The KPIs (overview) page, shown in *Figure 9*, provides a foundational understanding of the eight selected KPIs, outlining their definitions. The dashboard presents this information using text-based descriptions, tables summarizing KPI weights, and a pie chart visualizing the relative importance of each KPI. Users can directly track individual KPI performance from this section by clicking on buttons on right below side of the page.

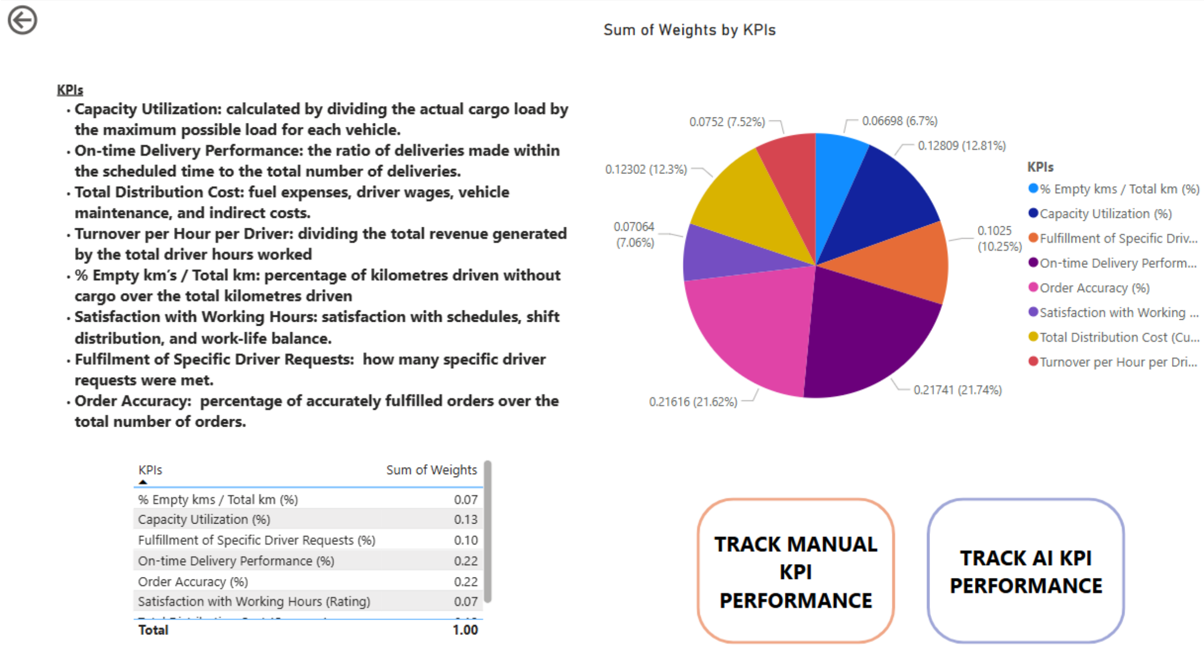

Figure 9 : KPIs Page

5.4.2 KPI Tracking and Visualization

The dashboard's KPI tracking functionality allows users to explore the performance trends of individual KPIs across the seven-week data period. Each KPI has a dedicated page that can be guided through the Visualization & Analysis page shown in *Figure 10*.

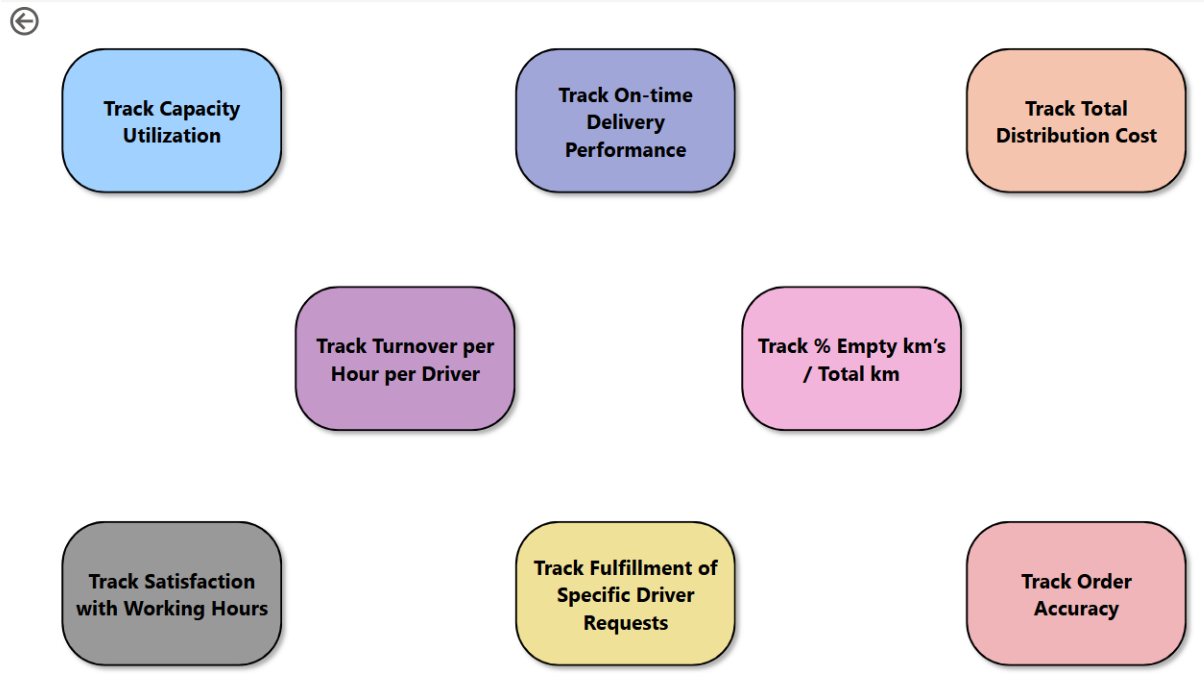

Figure 10 : Visualization & Analysis Page

When clicked upon one of the KPI page buttons, the opened page displays time-series visualizations, summary statistics, and system-specific insights. *Figure 11* shows one of the examples out of sixteen KPI pages.

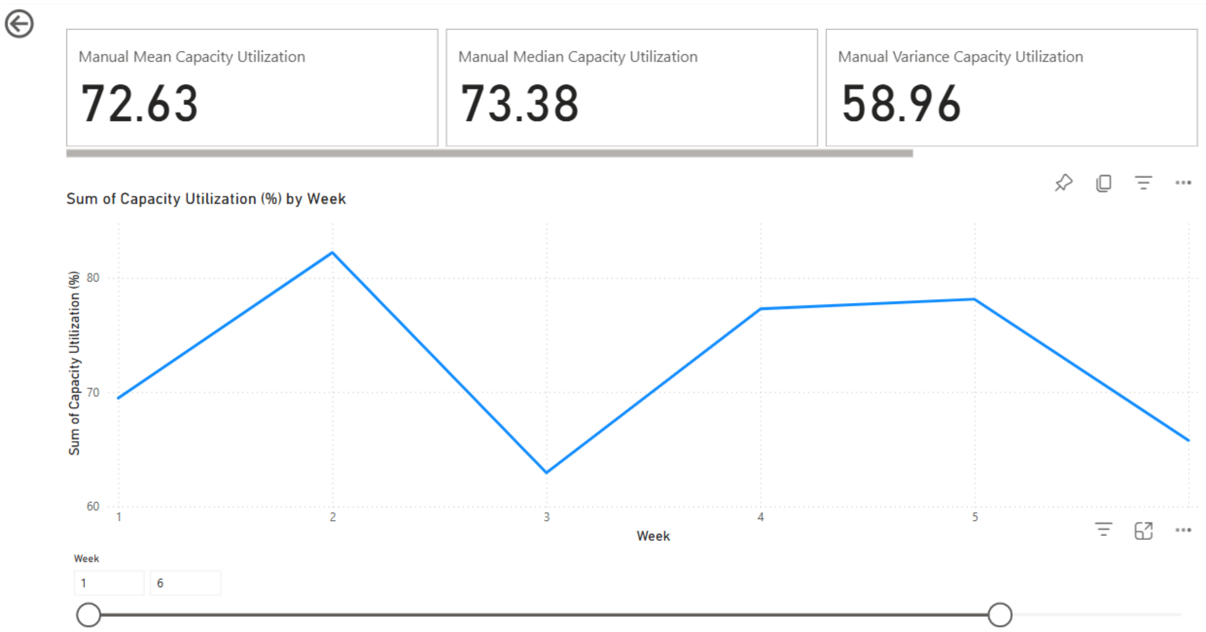

Figure 11 : Track Manuel Capacity Utilization Page

Key statistical measures, such as mean, median, variance, and standard deviation, are presented at the top of the page, offering a snapshot of the system's stability. Below, a line graph visualizes weekly trends, enabling users to identify fluctuations, patterns, and potential anomalies in logistics efficiency.

An important design choice was incorporating separate KPI pages for manual and AI-driven systems, allowing for independent assessments before direct comparisons. This structure ensures that users understand each system's strengths and weaknesses before evaluating performance differences.

5.4.3 WCS and WDA Calculation

The dashboard integrates a WCS and WDA calculations page, shown in *Figure 12*, translating raw KPI performance data into a structured evaluation framework. The WCS and WDA values are presented through numeric displays, comparison tables, and summary insights. A key feature of this section is its categorization of improvement, where overall performance change is classified into qualitative categories (e.g., "Modest Improvement") to provide more precise interpretations of the AI system's impact.

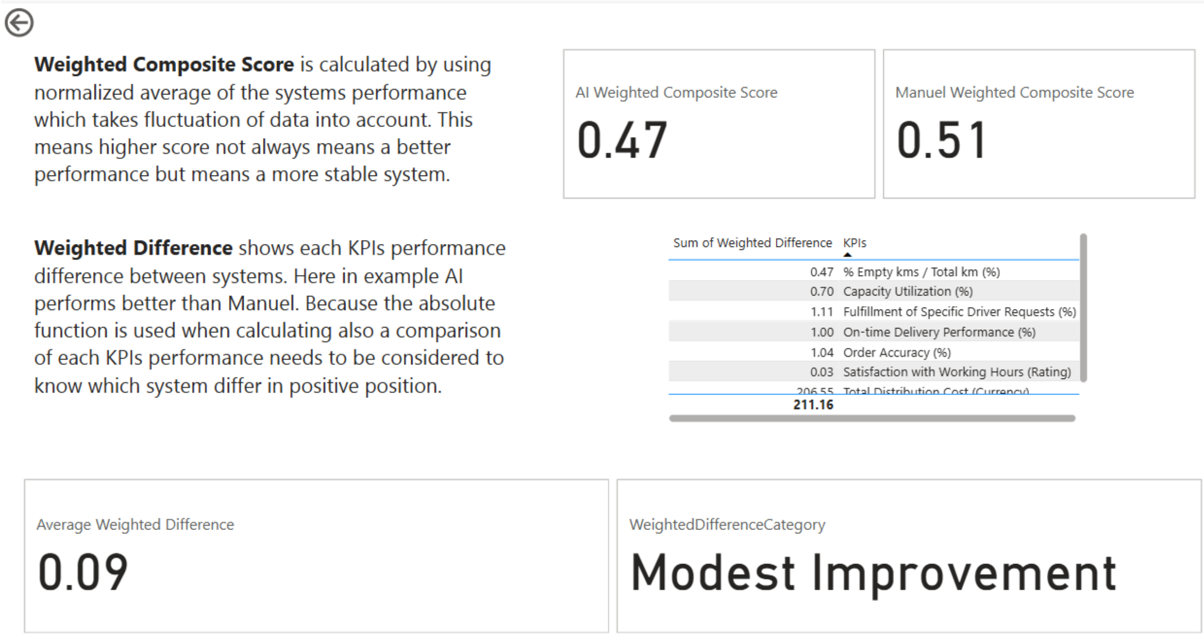

Figure 12 : WCS and WDA Calculations Page

5.4.4 System-Wide Comparison

The final component of the dashboard is the Comparison page shown in *Figure 13*, which aggregates all KPI evaluations into a direct comparison between the manual and AI-driven logistics planning methods. This section enables side-by-side evaluations and interactive filtering, providing users with a comprehensive decision-support tool. The dashboard uses comparative line graphs to present individual KPI performance trends for both systems. Users can select a specific KPI to visualize its historical performance across the seven weeks, enabling targeted assessments.

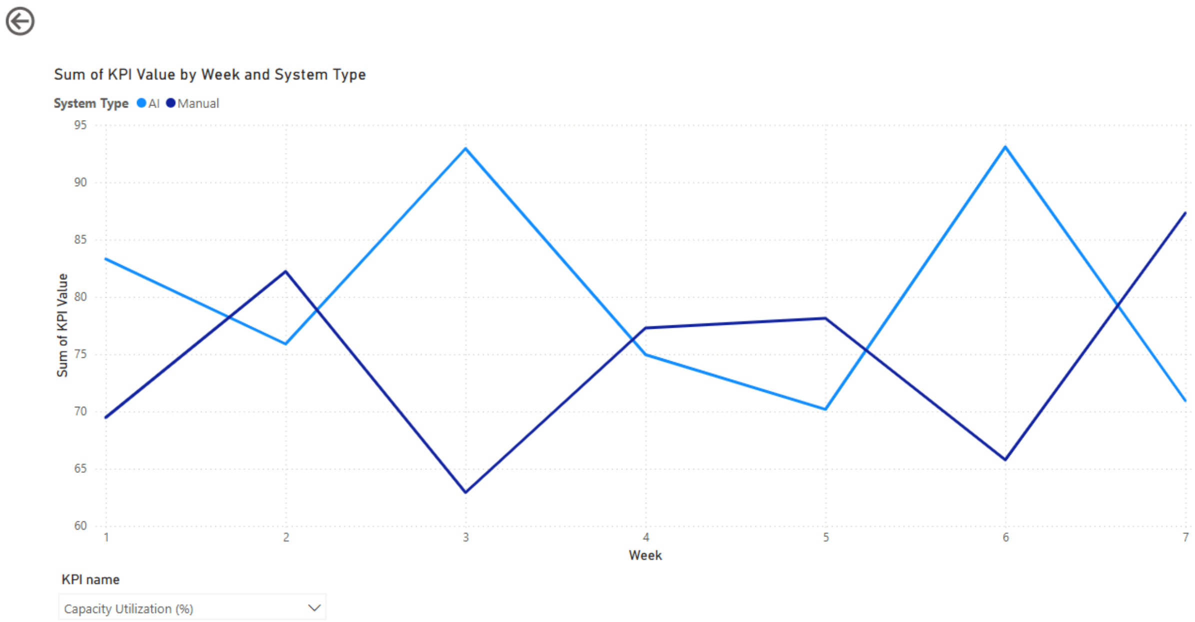

Figure 13: Comparison Page

6. Solution Evaluation

The evaluation chapter assesses the Comparison Roadmap and KPI Dashboard developed as part of the thesis. The assessment is based on feedback collected through an evaluation form completed by the company supervisor. This chapter critically examines the tools' usability, functionality, and strategic impact, highlighting their contributions to decision-making processes at Bouwvervoer while identifying areas for improvement. The feedback form used for evaluation can be found in Appendix 6.

6.1 Feedback Summary

The primary objective of this assessment was to determine the dashboard's usability, functionality, and overall effectiveness in supporting decision-making processes at Bouwvervoer. By gathering insights from the company supervisor, this evaluation highlights strengths and identifies areas for improvement.

The feedback was categorized into four main areas: usability and design, functionality and technical performance, impact on decision-making, and aesthetic and accessibility. Each section summarizes key observations, ratings, and recommendations for future improvements. While the dashboard successfully provides a structured framework for performance analysis, feedback suggests changes that could further improve its user experience, navigation efficiency, and comparative analysis capabilities.

It is important to note that the evaluation was conducted by only one participant, the company supervisor, limiting the scope and comprehensiveness of the assessment. While their insights provide valuable feedback, the absence of multiple users' perspectives restricts the ability to measure how different stakeholders might interact with and interpret the dashboard.

6.1.1 Usability and Design

The usability of the dashboard was rated 4/5, indicating that the tool is generally easy to navigate. The company supervisor noted that a key improvement would be ensuring users can return to the "home" page with one click from any section. Additionally, the dashboard layout received a 3/5 score for intuitiveness. It was observed that the lack of explicit markers identifying KPI origins (manual or AI) and their placement within the broader workflow sometimes caused confusion. Addressing these concerns by incorporating a more dynamic and visually explicit navigation framework could enhance usability further.

6.1.2 Functionality and Technical Performance

The dashboard scored 4/5 in terms of the helpfulness of its visualizations, such as line charts and weighted composite scores, for understanding trends and differences between the manual and AI-driven systems. A specific suggestion was to include comparative graphs overlaying manual and AI performance within a single visualization for easier direct comparisons. The technical bugs or limitations were not experienced, suggesting that the dashboard meets basic functionality expectations.

6.1.3 Impact on Decision-Making

The dashboard's impact on decision-making processes was rated 4/5. The company supervisor indicated that once populated with operational data, the dashboard could significantly support performance evaluations.

6.1.4 Aesthetic and Accessibility

The dashboard's aesthetic design was considered appealing and professional, with a rating of 4/5. Its accessibility to users with limited technical expertise also scored 4/5, reflecting that the design successfully works for various users.

6.2 Conclusion

The evaluation highlights the dashboard's success in meeting its core objectives of improving data visualization and supporting decision-making at Bouwvervoer. While the tool demonstrates functionality and usability, addressing the identified areas for improvement will further improve its strategic impact. With these refinements, the dashboard will provide an even more effective platform for evaluating the transition from manual to AI-driven logistics planning.

7. Conclusion

This final chapter brings together the key findings and contributions of the research, reflecting on how the study addressed its objectives and research questions. It begins by summarizing the methodology and results, including KPIs' selection and weighting, the comparison framework's development, and its implementation in a Power BI dashboard. The chapter then discusses the theoretical and practical implications of the framework, especially its relevance for Bouwvervoer and broader logistics decision-making. Finally, the study's limitations are assessed, and future research directions are presented.

7.1 Summary

This thesis aimed to develop a conceptual comparison framework for evaluating logistics planning systems, explicitly comparing a manual route planning system with an AI-driven alternative at Bouwvervoer. Given data availability constraints, the study constructed a structured methodology to assess logistics performance using KPIs, AHP-based weighting, and comparative analysis tools implemented in a Power BI dashboard.

To answer RQ1 and KQ1, the research conducted a literature review to define a set of eight KPIs. The selected KPIs included Capacity Utilization, On-time Delivery Performance, Total Distribution Cost, Turnover per Hour per Driver, % Empty km / Total km, Satisfaction with Working Hours, Fulfillment of Specific Driver Requests, and Order Accuracy.

To address RQ2 and KQ2, the study explored multiple weighting methodologies and selected MAHP as the most suitable approach. The MAHP framework provided structured, expert-driven prioritization of KPIs based on pairwise comparisons conducted with the company's logistics planners. The final weighting results indicated that On-time Delivery Performance and Order Accuracy were the most critical KPIs, reflecting the company's strong emphasis on service reliability. Cost-related KPIs such as Total Distribution Cost and Capacity Utilization were also highly weighted, demonstrating the financial and operational efficiency considerations in logistics decision-making.

The study conducted a data availability analysis in response to RQ3. This revealed significant operational data gaps for manual and AI-driven logistics systems, as detailed in Appendix 3. The lack of comprehensive historical records made it infeasible to compare the two systems directly. Consequently, a seven-week mock-up dataset was created to simulate logistics performance trends. The dataset was structured to reflect variability in AI-driven decision-making while maintaining the stability of manual system performance. This ensured the Power BI dashboard could effectively demonstrate the framework's analytical capabilities.

KQ3 focused on the tools necessary to analyze and visualize system performance in a way that supports decision-making. In response, the thesis reviewed and selected a set of established analytical methods: WCS to provide an overall performance metric, WDA to show system-level performance gaps by KPI, and Descriptive Statistics to allow examination of weekly stability and trends. These tools were supported by visualization techniques, including line charts, bar charts, pie charts, and stacked columns, which were selected based on the data characteristics of each KPI and implemented through a flowchart.

The conceptual framework was then implemented in Power BI answering. The dashboard was designed as an interactive tool for logistics managers, featuring:

- KPI tracking pages that visualize individual performance metrics.
- WCS and WDA calculations for system-level assessment.

- A comparative analysis function allows users to assess performance trends between the manual and AI-driven systems.

7.2Theoretical Contributions

This research contributes to comparative performance assessment. Developing a KPI-based comparison model bridges a gap in the literature, as most existing studies focus on AI's impact on logistics without offering structured comparative methodologies. By integrating KPI-based performance measurement, WCS, and WDA, this study provides an approach for systematically evaluating logistics performance.

The MAHP application introduces a weighting mechanism tailored for logistics KPIs. This multi-criteria decision-making tool improves the objectivity of KPI prioritization, ensuring that expert judgment is systematically incorporated into performance evaluation. Additionally, this study advances the role of data visualization in logistics analytics by demonstrating how interactive dashboards can improve decision-making transparency. The combination of statistical analysis, multi-criteria decision-making, and visualization techniques provides a foundation for future comparative studies in AI-driven logistics optimization.

7.3 Practical Implications for Bouwvervoer and Industry

The findings of this research have direct implications for Bouwvervoer and the logistics industry. The developed framework provides a decision-support tool that allows companies to evaluate AI adoption based on quantitative evidence rather than intuition or trial-and-error approaches. The results suggest that an immediate shift to AI-driven logistics planning may not be optimal. Instead, a hybrid approach that combines AI optimization with human oversight can lead to more stable performance while mitigating the risks associated with full automation.

The research highlights the importance of real-time KPI tracking in ensuring AI's alignment with operational goals. Companies must continuously monitor AI-driven logistics performance, making adjustments where necessary to maintain efficiency. Moreover, the study highlights the need for industry-specific AI customization. Since AI does not operate uniformly across all logistics contexts, firms should tailor AI algorithms to fit their specific operational constraints and strategic objectives.

7.4 Limitations and Future Directions of the Study

While this study provides valuable insights into the development and evaluation of a logistics performance dashboard, several limitations concerning validity, reliability, and generalizability must be acknowledged. These limitations impact the degree to which findings can be applied to broader contexts beyond the specific case of Bouwvervoer, and offer key insights into areas for future research and improvements.

7.4.1 Validity and Reliability Concerns

One of the main limitations of this study is the restricted size of the evaluation, as feedback was obtained from only a single participant, the company supervisor. A one-person evaluation naturally limits the reliability of the findings, as it does not account for various perspectives from other key stakeholders, such as planners and drivers. A more complete assessment involving multiple users across different roles would increase internal validity by ensuring that the dashboard meets the needs of various end-users. Furthermore, since no operational data was available during the evaluation, the feedback was based entirely on mock-up data. This restricts the validity of the evaluation, as user perceptions may change when interacting with real-world logistics data.

7.4.2 Generalizability to Other Logistics Sectors and Companies

Another significant limitation is that the study is highly context-specific, meaning the results cannot be easily generalized to other logistics companies or sectors. The dashboard was designed with Bouwvervoer's specific operational structure, KPI framework, and managerial priorities, making it less relevant to companies with different logistics models, fleet structures, or business goals. Additionally, the choice of KPIs and their weighting were tailored to Bouwvervoer's unique evaluation criteria, meaning that other firms with different KPIs may require significant modifications to the framework before its application.

The data gaps identified in the study also constrain generalizability. Since the conceptual framework relies on structured data collection and integration, companies with different data availability and structure may not be able to implement the same approach without adjustments. Furthermore, logistics firms that operate in different geographic regions, market conditions, or regulatory environments may have different performance goals that the current dashboard does not account for.

7.4.3 Need for Further Validation and Testing

Given these limitations, future research should focus on expanding the scope of evaluation by involving multiple stakeholders in the validation process. Conducting usability tests with a broader sample of employees from different functional areas would improve the reliability and credibility of the results. Additionally, implementing the dashboard with accurate operational data rather than mock-up data would allow for a more accurate assessment of its practical effectiveness, ultimately improving its adaptability to different logistics environments.

In conclusion, while the developed dashboard offers a structured and methodologically sound approach to comparing manual and AI-driven logistics systems, its evaluation remains poor and context-specific. Despite these limitations, the research lays the foundation for future advancements. By addressing these limitations, future research can refine the conceptual comparison framework into a more robust, widely applicable decision-support tool for logistics management.

References

Alves Junior, P. N., Melo, I. C., Branco, J. E., Bartholomeu, D. B., & Caixeta-Filho, J. V. (2021). Which green transport corridors (GTC) are efficient? A dual-step approach using network equilibrium model (NEM) and data envelopment analysis (DEA). *Journal of Marine Science and Engineering*, *9*(3), 247. https://doi.org/10.3390/jmse9030247

Belhadi, A., Touriki, F. E., & El fezazi, S. (2017). Prioritizing the solutions of lean implementation in smes to overcome its barriers. *Journal of Manufacturing Technology Management*, *28*(8), 1115–1139. https://doi.org/10.1108/jmtm-04-2017-0066

Boute, N. (2022). Artificial Intelligence in Logistics and Supply Chain Management. In *Applications of Artificial Intelligence in Industry and Economy* (pp. 49-65). Springer. https://doi.org/10.1007/978-3-030-95764-3_3

Bouwvervoer. (2024). *Company website*. Retrieved from https://www.bouwvervoer.nl

Celik, E., & Akyuz, E. (2018). An interval type-2 fuzzy AHP and Topsis methods for decision-making problems in maritime transportation engineering: The case of ship loader. *Ocean Engineering*, *155*, 371–381. https://doi.org/10.1016/j.oceaneng.2018.01.039

Choi, T.-M., Chiu, C.-H., & Chan, H.-K. (2016). Risk management of logistics systems. *Transportation Research Part E: Logistics and Transportation Review*, *90*, 1–6. https://doi.org/10.1016/j.tre.2016.03.007

Chorfi, Z., Berrado, A., & Benabbou, L. (2015). Selection of key performance indicators for supply chain monitoring using MCDA. *2015 10th International Conference on Intelligent Systems: Theories and Applications (SITA)*, 1–6. https://doi.org/10.1109/sita.2015.7358395

Cooper, D. R., & Schindler, P. S. (2014). *Business Research Methods*. McGraw-Hill Education.

Danladi, C., Tuck, S., Tziogkidis, P., Tang, L., & Okorie, C. (2024). Efficiency analysis and benchmarking of container ports operating in lower-middle-income countries: A DEA approach. *Journal of Shipping and Trade*, *9*(1). https://doi.org/10.1186/s41072-024-00163-2

Devellis, R. F. (2016). *Scale development: Theory and applications*. Sage Publications, Inc.

Dong, Y. (2023). Descriptive statistics and its applications. *Highlights in Science, Engineering and Technology*, *47*, 16–23. https://doi.org/10.54097/hset.v47i.8159

García-Arca, J., Prado-Prado, J. C., & Fernández-González, A. J. (2018). Integrating kpis for improving efficiency in road transport. *International Journal of Physical Distribution & Logistics Management*, *48*(9), 931–951. https://doi.org/10.1108/ijpdlm-05-2017-0199

Heerkens, G. (2017). *Managerial Problem-Solving Method (MPSM)*.

Holden, R., Xu, B., Greening, P., Piecyk, M., & Dadhich, P. (2016). Towards a common measure of greenhouse gas related logistics activity using data envelopment analysis. *Transportation Research Part A: Policy and Practice*, *91*, 105–119. https://doi.org/10.1016/j.tra.2016.06.001

Iankoulova, I. (2012) Business intelligence for horizontal cooperation: measuring the performance of a transportation network sharing cooperation between logistics companies.

Jothimani, D., & Sarmah, S. P. (2014). Supply Chain Performance Measurement for Third Party Logistics. *Benchmarking: An International Journal*, *21*(6), 944–963. https://doi.org/10.1108/bij-09-2012-0064

Krauth, E., Moonen, H. & Popova, V., 2005. Performance Measurement and Control in Logistics Service Providing.

Kumar, M., Sharma, M., Raut, R. D., Mangla, S. K., & Choubey, V. K. (2022). Performance Assessment of Circular Driven Sustainable Agri-food supply chain towards achieving sustainable consumption and production. *Journal of Cleaner Production*, *372*, 133698. https://doi.org/10.1016/j.jclepro.2022.133698

Ladyman, J. (2012). *Understanding Philosophy of science*. Routledge.

Paczkowski, W. R. (2021). *Business Analytics*. https://doi.org/10.1007/978-3-030-87023-2

Ramanathan, R. (1997). Stochastic decision making using multiplicative AHP. *European Journal of Operational Research*, *97*(3), 543–549. https://doi.org/10.1016/s0377-2217(96)00285-8

Regragui, H., Sefiani, N., & Azzouzi, H. (2018). Improving performance through measurement: The application of bsc and AHP in Healthcare Organization. *2018 International Colloquium on Logistics and Supply Chain Management (LOGISTIQUA)*, 51–56. https://doi.org/10.1109/logistiqua.2018.8428262

Rinat, K., Koli, S., Sobti, R., Ledalla, S., & Arora, R. (2024). Data-driven decision making: Real-world effectiveness in industry 5.0 – an experimental approach. *BIO Web of Conferences*, *86*, 01061. https://doi.org/10.1051/bioconf/20248601061

Sirisawat, P., Rodbundith, T. S., & Hasachoo, N. (2024). Identifying the hospital logistics key performance indicators for public hospitals in remote areas of Thailand. *International Journal of Pharmaceutical and Healthcare Marketing*. https://doi.org/10.1108/ijphm-06-2023-0053

Soriano-Gonzalez, R., Perez-Bernabeu, E., Ahsini, Y., Carracedo, P., Camacho, A., & Juan, A. A. (2023). Analyzing key performance indicators for Mobility Logistics in smart and sustainable cities: A case study centered on Barcelona. *Logistics*, *7*(4), 75. https://doi.org/10.3390/logistics7040075

Thomé, A. M., Scavarda, L. F., & Scavarda, A. J. (2016). Conducting systematic literature review in Operations Management. *Production Planning & Control*, *27*(5), 408–420. https://doi.org/10.1080/09537287.2015.1129464

Tu, J., & Wu, Z. (2023). Analytic hierarchy process rank reversals: Causes and solutions. *Annals of Operations Research*. https://doi.org/10.1007/s10479-023-05278-6

Tufte, E. (2016). The Visual Display of Quantitative Information (1983). *Diagrammatik-Reader*, 219–230. https://doi.org/10.1515/9783050093833-035

University of Twente. (2015). Organisation, Regulations, and Codes of Conduct. Retrieved from https://www.utwente.nl/en/service-portal/organisation-regulations-and-codes-of-conduct/codes-of-conduct

Van De Poel, I., & Royakkers, L. (2011). *Ethics, Technology, and engineering: An Introduction*. John Wiley & Sons.

Vermaas, P. E., Kroes, P., & Franssen, M. (2011). *A philosophy of technology: From Technical Artefacts to Sociotechnical Systems*. Morgan & Claypool Publishers.

Wibowo, M. A., & Sholeh, M. N. (2015). The analysis of Supply Chain Performance Measurement at Construction Project. *Procedia Engineering*, *125*, 25–31. https://doi.org/10.1016/j.proeng.2015.11.005

Zhang, Y. (2019). *Adoption of artificial intelligence in logistics: An exploratory study*. Journal of Physics: Conference Series, 1325(1), 012085.

Appendix

Appendix 1 - Systematic Literature Review

Aim of Research:

The objective of this systematic literature review is to answer the following questions: What methodologies are currently used to quantify, weigh, or prioritize key performance indicators (KPIs) in the context of logistics, supply chain management, or route planning? This research gathers insights into the various approaches and techniques used to assess the importance of KPIs. By examining current practices and trends, the review seeks to understand these methodologies' effectiveness, challenges, and impacts on applications to better understand KPI integration in logistics operations.

Database Selection:

The research will use Scopus and Web of Science databases for this systematic literature review. These platforms were chosen due to their extensive scientific and academic literature archives, containing various disciplines focusing on technology and AI-related studies. Also, they are recommended by the University of Twente. These databases will help access relevant, high-quality academic papers, articles, and case studies relevant to AI-driven logistics systems.

Inclusion and Exclusion Criteria:

Setting explicit inclusion and exclusion criteria in a systematic literature review is crucial for filtering relevant research. These criteria guide the selection process. This study's criteria focus on identifying studies that address the quantification, weighting, or prioritization of KPIs in logistics, supply chain management, or route planning. Including peer-reviewed and English-language publications ensures the credibility and accessibility of the research data.

Criteria Type	**Criteria**
Inclusion	Focus: Articles must focus on the quantification, weighting, or prioritization of KPIs in the context of logistics, supply chain management, or route planning.
	Methodological Insight: Studies that provide clear methodologies or frameworks for quantifying the importance of KPIs.
	Peer-Reviewed: Only include peer-reviewed journal articles to ensure research validity.
	Publication Date: Prioritize studies published in the last 10 years to ensure relevance to current technologies and methodologies.
	Language: English-language publications only.
Exclusion	Focus: Research not addressing the quantification, weighting, or prioritization of KPIs in logistics systems.
	Language: Non-English research materials

Inclusion and Exclusion Table

Search Matrix:

Key Concept	Related Terms/Synonyms	Narrower Terms	Broader Terms
KPI Quantification	KPI Weighting, KPI Prioritization	KPI Selection Criteria, Performance Indicator Quantification	Performance Metrics, Business Intelligence
Logistics	Route Planning, Supply Chain	Transportation Management Systems, Warehouse Management Systems	Logistics Technology, Supply Chain Management Systems
Methodologies	Analytical Techniques	AHP, Fuzzy AHP, DEA, TOPSIS	Multicriteria Decision Making, Quantitative and Qualitative Analysis

Search Matrix Table

Search Query:

Search Query	Source	Number of Results	Date of Search	Notes
("quantifying KPI importance" OR "KPI weighting" OR "KPI prioritization" OR "route planning KPIs importance" OR "logistics KPI weighting" OR "performance indicator quantification" OR "KPI analysis in logistics" OR "KPI selection criteria" OR "multi-criteria decision making") AND ("logistics" OR "route planning" OR "supply chain")	Web of Science	1853	5-12-2023	Initial broad search for KPI quantification methodologies.
KPI* OR "KPI selection”) AND (AHP* OR "Fuzzy AHP" OR "multicriteria decision making") AND (logist* OR "supply chain" OR "route planning")	Web of Science	28	5-12-2023	Focused search using AHP methodology.
("KPI weigh*" OR "KPI select*" OR "priority setting in logistics") AND ("AHP" OR "Fuzzy AHP" OR "multicriteria decision making") AND ("logist*" OR "supply chain" OR "route planning")	Scopus	1	5-12-2023	Narrowed search in Scopus for AHP in logistics.
(KPI* OR "KPI selection") AND ("Data Envelopment Analysis" OR DEA*) AND (logist* OR "supply chain" OR "route planning")	Scopus	30	5-12-2023	Search using DEA methodology.

(KPI* OR "KPI selection") AND ("Data Envelopment Analysis" OR DEA*) AND (logist* OR "supply chain" OR "route planning")	Web of Science	40	5-12-2023	Expanded search in Web of Science for DEA in logistics.
(KPI* OR "KPI selection") AND (TOPSIS OR "Order Preference by Similarity to Ideal Solution") AND (logist* OR "supply chain" OR "route planning")	Scopus	7	5-12-2023	Search using TOPSIS methodology.
(KPI* OR "KPI selection") AND (TOPSIS OR "Order Preference by Similarity to Ideal Solution") AND (logist* OR "supply chain" OR "route planning")	Web of Science	10	5-12-2023	Expanded search in Web of Science for TOPSIS in logistics.

Screening Process:

In this systematic literature study, we used Mendeley, a well-known reference manager, to organize and streamline the screening process. Mendeley ensured that the pieces of literature were handled methodically and effectively. We imported references from previously stated databases. In the initial screening stage, titles and abstracts were compared to our inclusion and exclusion lists. The publication date, study design, and relevance to the research topic were among the criteria. We performed a near full-text review of the papers that made it beyond the first screening.

Literature Overview Matrix:

Method	Article Title	Authors	Year of Publication	Context of Logistics Planning	KPIs Used	Findings	Limitations
AHP	Identifying the hospital logistics key performance indicators for public hospitals in remote areas of Thailand	Sirisawat, P., Rodbundith, T. S., & Hasachoo, N.	2024	Hospital logistics in remote areas	Delivery time, inventory levels, cost efficiency	AHP effectively prioritized KPIs for improving hospital logistics.	Specific to the healthcare sector; may not be directly transferable to other logistics contexts.
AHP	Improving performance through measurement: the application of BSC and AHP in healthcare organization	Regragui, H., Sefiani, N., & Azzouzi, H.	2015	Healthcare logistics	Patient satisfaction, operational efficiency, cost management	AHP combined with BSC provides a framework for performance measurement.	Healthcare-focused; may require adaptation for other logistics environments.
AHP	Selection of Key Performance		2018	General supply chain	Delivery reliability,	AHP helps in selecting and	Needs context-specific adjustments

	Indicators for Supply Chain monitoring using MCDA			management	cost efficiency, lead time	prioritizing KPIs effectively.	for different industries.
AHP	The Analysis of Supply Chain Performance Measurement at Construction Project	Wibowo, M. A., & Sholeh, M. N.	2015	Construction project logistics	Project completion time, cost management, resource utilization	AHP is useful in evaluating and improving supply chain performance in construction projects.	Specific to construction; may need modifications for other sectors.
DEA	Which Green Transport Corridors (GTC) Are Efficient? A Dual-Step Approach Using Network Equilibrium Model (NEM) and Data Envelopment Analysis (DEA)	Alves Junior, P. N., Melo, I. C., Branco, J. E., Bartholomeu, D. B., & Caixeta-Filho, J. V.	2021	Green transport corridors	Sustainability, cost efficiency, delivery reliability	Combining NEM and DEA effectively evaluates the efficiency of green transport corridors.	Complex methodology may be challenging to implement.
DEA	Efficiency analysis and benchmarking of container ports operating in lower-middle-income countries: A DEA approach	Danladi, C., Tuck, S., Tziogkidis, P., Tang, L., & Okorie, C.	2024	Container ports in lower-middle-income countries	Operational efficiency, throughput, cost efficiency	DEA is effective in benchmarking and analyzing the efficiency of container ports.	Requires large and accurate data sets for reliable results.
DEA	Integrating KPIs for improving efficiency in road transport	García-Arca, J., Prado-Prado, J. C., & Fernández-González, A. J.	2018	Road transport logistics	Fuel efficiency, delivery time, cost efficiency	DEA helps in integrating and improving KPIs for road transport efficiency.	May not capture all qualitative aspects of performance.
DEA	Towards a common measure of greenhouse gas	Holden, R., Xu, B., Greening, P., Piecyk, M., &	2016	Greenhouse gas emissions in logistics	GHG emissions, fuel efficiency,	DEA provides a common measure for evaluating greenhouse gas-	Focuses primarily on environmental KPIs, may overlook other important logistics

	related logistics activity using data envelopment analysis	Dadhich, P.			operational efficiency	related activities in logistics.	aspects.
TOPSIS	An interval type2 fuzzy AHP and TOPSIS methods for decision making problems in maritime transportation engineering	Celik, E., & Akyuz, E.	2018	Maritime transportation engineering	Safety, cost efficiency, operational efficiency	Combining fuzzy AHP and TOPSIS improves decision-making in maritime transportation by addressing uncertainty.	Complexity in implementing fuzzy logic and the need for expert input.
TOPSIS	KPI and Logistics Dashboard Design Using Neutrosophic Statistics	Goyes García, J. F., Carrión Hurtado, L. H., León Yacelga, M. A., & Enríquez Chugá, J. F.	2021	General logistics management	Delivery times, inventory management, cost efficiency	TOPSIS combined with neutrosophic statistics enhances the visualization and decision-making process in logistics.	Requires advanced statistical knowledge and integration with neutrosophic logic.
TOPSIS	Performance assessment of circular driven sustainable agri-food supply chain towards achieving sustainable consumption and production	Kumar, M., Sharma, M., Raut, R. D., Mangla, S. K., & Choubey, V. K.	2022	Sustainable agri-food supply chain	Sustainability, resource efficiency, cost management	TOPSIS effectively prioritizes sustainability KPIs in agri-food supply chains.	May require extensive data collection and validation.
TOPSIS	Prioritizing the solutions of lean implementation in SMEs to overcome its barriers: An integrated fuzzy	Belhadi, A., Touriki, F. E., & El fezazi, S.	2017	Lean implementation in SMEs	Operational efficiency, cost reduction, process improvement	Integrating fuzzy AHP and TOPSIS helps in prioritizing lean implementation solutions effectively.	Complexity in integrating fuzzy AHP with TOPSIS and the need for expert judgment.

	AHP-TOPSIS approach						
TOPSIS	Supply chain performance measurement for third party logistics	Jothimani, D., & Sarmah, S. P.	2014	Third party logistics	Delivery reliability, cost efficiency, customer satisfaction	TOPSIS provides a comprehensive framework for measuring supply chain performance in third party logistics.	May not fully capture qualitative aspects of performance and requires data.

KQ1 Literature Table

Conclusion:

The systematic literature review aimed to identify and evaluate the current methodologies used to quantify, weigh, or prioritize key performance indicators (KPIs) in logistics, supply chain management, and route planning. It was found that methodologies such as the Analytical Hierarchy Process (AHP), Data Envelopment Analysis (DEA), and Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) are often used for KPI evaluation and decision-making. Each of these methods has its strengths and limitations, making it suitable for different contexts and applications. AHP is highly effective for scenarios requiring expert judgment and pairwise comparisons. DEA excels in quantitative settings with multiple decision-making units, and TOPSIS is advantageous for complex, multi-dimensional decision-making problems. Effectiveness, practicality, accuracy, and flexibility were essential criteria for evaluating these methods. After review, we found MAHP, an improved version of AHP, to be the suitable method for weighting KPIs in this thesis.

Appendix 2 - KPI Table

KPIs	DEFINITION	SIGNIFICANCE	HOW IT MEASURED	APPLICATION
Capacity Utilization	Capacity Utilization refers to how effectively the available transport capacity is used during operations. It measures the proportion of vehicle capacity that is utilized during each trip	Optimizing capacity utilization can significantly reduce the number of trips needed, lowering fuel consumption and operational costs. Underutilization of capacity often leads to inefficiencies and higher per-unit transportation costs	This KPI is expressed as a percentage, calculated by dividing the actual cargo load by the maximum possible load for each vehicle. This data can be collected from load manifests and fleet management systems	This KPI helps Bouwvervoer assess whether trucks are being used optimally, indicating areas where better load planning could reduce empty or partially filled trips
On-time Delivery Performance	This KPI tracks the percentage of deliveries that arrive within the scheduled time window, a critical measure of logistics	Timely deliveries are crucial for maintaining customer satisfaction and ensuring operational effectiveness. Delays can result in penalties,	This KPI is the ratio of deliveries made within the scheduled time to the total number of deliveries. The time data is usually	By focusing on this KPI, Bouwvervoer can identify bottlenecks in the delivery process, such as traffic congestion or

	reliability	lost business opportunities, and decreased client trust	recorded through systems like electronic proof of delivery (ePOD) or GPS-enabled fleet tracking	inefficient route planning, and work to improve delivery times
Total Distribution Cost	This KPI aggregates all the costs associated with distribution, including fuel, labour, maintenance, and other operational expenses	Minimizing total distribution costs is essential for improving overall financial performance. High distribution costs can reduce profit margins and make the logistics process unsustainable	Total Distribution Cost includes multiple factors such as fuel expenses, driver wages, vehicle maintenance, and indirect costs like administration. Data is typically sourced from financial records and fleet management systems	Monitoring this KPI allows Bouwvervoer to perform cost analyses, identify inefficiencies, and explore opportunities for cost savings, such as optimizing fuel usage or scheduling more efficient routes
Turnover per Hour per Driver	Measures the revenue generated per hour drivers work, assessing driver productivity in logistics operations	High turnover per driver hour indicates efficient use of labour resources. Low turnover could suggest inefficiencies like poor route planning, extended idle times, or inadequate scheduling	Calculated by dividing the total revenue generated by the total driver hours worked	Allows Bouwvervoer to assess driver productivity, identify inefficiency in labour use, and improve overall workforce management
% Empty km's / Total km	Measures the proportion of kilometres driven without cargo, indicating inefficiencies in route and capacity planning	Reducing the percentage of empty kilometres improves fuel efficiency and cost savings while minimizing environmental impact through lower emissions	This KPI is expressed as a percentage of kilometres driven without cargo over the total kilometres driven	Help Bouwvervoer in identifying inefficiencies in route planning and capacity utilization, reducing unnecessary trips and environmental impact, and optimizing vehicle use
Satisfaction with Working Hours	Measures employee satisfaction regarding their working schedules, particularly regarding work-life balance and flexibility	High employee satisfaction contributes to improved morale, higher productivity, and lower turnover, reducing recruitment and training costs	They are measured through employee surveys or feedback forms, focusing on satisfaction with schedules, shift distribution, and work-life balance. Data can be collected through periodic internal surveys	It helps Bouwvervoer maintain a satisfied workforce, reduce turnover, and improve overall operational stability
Fulfillment of Specific Driver Requests	Assesses how well the company meets drivers' requests, such as preferred	Meeting drivers' requests improves job satisfaction and contributes to a more	Measured through driver feedback and request fulfilment logs, tracking how	Assists Bouwvervoer in building a positive work environment

	routes or schedule adjustments	engaged and productive workforce. Employee engagement often leads to higher commitment and better operational performance	many specific driver requests were met	by addressing driver needs
Order Accuracy	Measures the percentage of customer orders delivered accurately regarding correct items, quantities, and conditions	Accurate order fulfillment is crucial for customer satisfaction and retention	They are calculated as the percentage of accurately fulfilled orders over the total number of orders	Enables Bouwvervoer to assess operational precision in order fulfilment and identify areas where processes need improvement to maintain high customer satisfaction and reduce operational errors

Appendix 3 - KPI Data Availability

KPIs	REQUIRED DATA	AVAILABILITY	GAPS IDENTIFIED
Capacity Utilization	Load manifests, vehicle capacity, trip records, cargo weight/volume	**Available**: in AI-SYSTEM data provides order weight and volume **Partially available:** in Manual, "Planning Order Dataset twee weken 2024" provides cargo amount (clarification needed on type of cargo)	**For Manual:** Clarification needed on cargo amounts in "Planning Order Dataset twee weken 2024"
On-time Delivery Performance	Delivery schedules, timestamps, GPS tracking data	**Available:** in AI-SYSTEM data provide planned arrive **Available:** in Manual, Delivery schedules exist in "Planning Orderset 2023" and "Planning Order Dataset twee weken 2024" for each trip	**For both:** No real-time data (GPS) on traffic delays or other factors affecting punctuality
Total Distribution Cost	Fuel costs, labor costs, maintenance, administrative expenses	**Partially Available:** in Manual, fuel consumption data available in "Fuel Consumption" dataset **Not available:** in AI-SYSTEM AI data	**For Manual:** Need data on labor and maintenance costs **For AI-SYSTEM:** Not available at all
Turnover per Hour per Driver	Driver work hours, revenue, trip data	**Not available:** in AI-SYSTEM AI data **Available:** in Manual, trip-specific revenue and work hour data exist in "Planning Order Dataset twee weken 2024."	**For AI-SYSTEM:** Not available at all
% Empty km's / Total km	Trip logs, cargo load data, route and GPS data	**Partially Available:** Trip logs, and order weight and volume available in "AI-system Trip Steps," but no cargo load data. **Partially Available:** in Manual, Delivery schedules exist in "Planning	**For Both:** Need cargo load data to accurately calculate empty kilometers

		Orderset 2023" and "Planning Order Dataset twee weken 2024" for each trip	
Satisfaction with Working Hours	Employee surveys, shift schedules	**Partially Available:** In Manual, shift schedules available in "Overnight" dataset. **Not Available:** in AI-SYSTEM AI data	**For Manual:** Employee satisfaction survey data is missing. **For AI-SYSTEM:** Not available at all
Fulfillment of Specific Driver Requests	Request logs, feedback forms, scheduling data	**Not Available:** In AI-system AI data **Partially Available:** In Manual, driver preferences exist, but no data on fulfillment of these preferences	**For Both**: Need system to track fulfillment of driver requests
Order Accuracy	Order logs, customer feedback, returns data	**Partially Available:** In AI-SYSTEM data order logs available **Partially Available:** In Manual, order logs available in the "Planning Order" dataset	**For Both:** No data on customer feedback or returns to verify accuracy

Expert Evaluation of Key Performance Indicators (KPIs) for Route Planner Efficiency and Effectiveness

This questionnaire aims to gather expert insights on the relative importance of Key Performance Indicators (KPIs) used to evaluate the efficiency and effectiveness of route planning in logistics. We also consider employee satisfaction and customer's point of view. Your input will contribute to a more accurate and effective prioritization of these KPIs

SCOR KPI Metrics

Level 1 Metrics
Transportation Effectiveness and Efficiency: Measures the ability to transport goods in a timely and cost-effective manner.
Employee Satisfaction: Indicates how satisfied employees are with their work environment, conditions, and overall job satisfaction.
Customer's Point of View: Reflects customer satisfaction and their perception of the service quality.

Level 2 Metrics
Logistics Costs: Represents the overall costs associated with logistics operations.
Logistics Efficiency: Measures how effectively logistics resources are utilized.
Service Quality: Evaluates the quality of the logistics service provided to customers.
Work environment:

- **Satisfaction with Working Hours:** Indicates how satisfied employees are with their working hours.

Employee Engagement:

- **Fulfillment of Specific Driver Requests:** Measures how well specific requests or needs of drivers are met.

Service Quality:

- **On-time Delivery Performance:** Evaluates the punctuality of deliveries according to the schedule.

Order Fulfillment Accuracy:

- **Order Accuracy:** Measures the accuracy of orders delivered to customers.

Level 3 Metrics
% of Realized km out of Planned km: Measures the proportion of kilometers actually traveled compared to what was planned.
Capacity Utilization: Indicates how effectively the available capacity is used in the logistics operations.
% Empty km's / Total km: Measures the proportion of kilometers driven without cargo compared to the total kilometers driven.
Turnover per Hour per Driver: Measures the revenue generated per hour worked by each driver.

Instructions

- Complete the Expert Background Information section to provide context for your responses.
- For each pairwise comparison of KPIs, use the linear scale to indicate how much more important one KPI is compared to the other in the Pairwise Comparisons section.

Confidentiality

Your responses will be kept confidential and used solely for the purpose of this analysis. The results will be aggregated and anonymized to ensure that individual responses cannot be identified

Expert Background Information

Please provide the following information about yourself to help us understand your perspective and expertise.

Name

First Name Last Name

Current Role/Position

Organization

Years of Experience in Logistics/Related Field

Less than 1 year
1-3 years
3-5 years
5-10 years
More than 10 years

Pairwise Comparisons of KPIs

For each pair of KPIs, please indicate how much more important one KPI is compared to the other using the linear scale provided. Consider the importance of each KPI in the context of route planner efficiency and effectiveness.

Pairwise Comparison Scale

9 - Extremely more important
8 - Very strongly to extremely more important
7 - Very strongly more important

6 - Strongly to very strongly more important
5 - Strongly more important
4 - Moderately to strongly more important
3 - Moderately more important
2 - Slightly more important
1 - Equal importance

Q1
Transportation Effectiveness and Efficiency vs. Employee Satisfaction
Context: When planning routes, consider the trade-off between achieving high transportation efficiency and ensuring employee satisfaction. How do you weigh the importance of transporting goods efficiently and effectively (Transportation Effectiveness and Efficiency) against maintaining high levels of employee satisfaction?
Question: On a scale of 9 to 1 to 9, how much more or less important is "Transportation Effectiveness and Efficiency" compared to "Employee Satisfaction"?
Linear scale from 9 (Transportation Effectiveness and Efficiency is extremely more important) to 1 (Equal importance) to 9 (Employee Satisfaction is extremely more important)

Transportation Effectiveness and Efficiency (LEFT SIDE) vs. Employee Satisfaction (RIGHT SIDE)

	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9
Transportation Effectiveness and Efficiency / Employee Satisfaction																	

Level 1 Metrics

Q2
Transportation Effectiveness and Efficiency vs. Customer's Point of View
Context: Consider the balance between achieving high transportation efficiency and meeting customer expectations. How do you weigh the importance of transporting goods efficiently and effectively (Transportation Effectiveness and Efficiency) against maintaining a high level of customer satisfaction (Customer's Point of View)?
Question: On a scale of 9 to 1 to 9, how much more or less important is "Transportation Effectiveness and Efficiency" compared to "Customer's Point of View"?
Linear scale from 9 (Transportation Effectiveness and Efficiency is extremely more important) to 1 (Equal importance) to 9 (Customer's Point of View is extremely more important)

Transportation Effectiveness and Efficiency (LEFT SIDE) vs. Customer's Point of View (RIGHT SIDE)

	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9
Transportation Effectiveness and Efficiency / Customer's Point of View																	

Q3
Employee Satisfaction vs. Customer's Point of View
Context: When planning routes, consider the trade-off between maintaining high employee satisfaction and meeting customer expectations. How do you weigh the importance of maintaining high levels of employee satisfaction (Employee Satisfaction) against ensuring customer satisfaction (Customer's Point

of View)?
Question: On a scale of 9 to 1 to 9, how much more or less important is "Employee Satisfaction" compared to "Customer's Point of View"? Linear scale from 9 (Employee Satisfaction is extremely more important) to 1 (Equal importance) to 9 (Customer's Point of View is extremely more important)

Employee Satisfaction (LEFT SIDE) vs. Customer's Point of View (RIGHT SIDE)

	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9
Employee Satisfaction / Customer's Point of View																	

Q4
Logistics Costs vs. Logistics Efficiency
Context: Consider the balance between minimizing logistics costs and achieving high logistics efficiency. How do you weigh the importance of controlling logistics costs (Logistics Costs) against maximizing the efficiency of logistics operations (Logistics Efficiency)?
Question: On a scale of 9 to 1 to 9, how much more or less important is "Logistics Costs" compared to "Logistics Efficiency"?Linear scale from 9 (Logistics Costs are extremely more important) to 1 (Equal importance) to 9 (Logistics Efficiency is extremely more important)

Logistics Costs (LEFT SIDE) vs. Logistics Efficiency (RIGHT SIDE)

	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9
Logistics Costs / Logistics Efficiency																	

Level 2 Metrics

Q5
Logistics Costs vs. Service Quality
Context: Consider the trade-off between minimizing logistics costs and maintaining high service quality. How do you weigh the importance of controlling logistics costs (Logistics Costs) against providing high-quality service (Service Quality)?
Question: On a scale of 9 to 1 to 9, how much more important is "Logistics Costs" compared to "Service Quality"?Linear scale from 9 (Logistics Costs are extremely more important) to 1 (Equal importance) to 9 (Service Quality is extremely more important)

Logistics Costs (LEFT SIDE) vs. Service Quality (RIGHT SIDE)

	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9
Logistics Costs / Service Quality																	

Q6
Logistics Efficiency vs. Service Quality
Context: When planning routes, consider the trade-off between achieving high logistics efficiency and maintaining high service quality. How do you weigh the importance of maximizing logistics efficiency (Logistics Efficiency) against providing high-quality service (Service Quality)?

Question: On a scale of 9 to 1 to 9, how much more important is "Logistics Efficiency" compared to "Service Quality"?Linear scale from 9 (Logistics Efficiency is extremely more important) to 1 (Equal importance) to 9 (Service Quality is extremely more important)

Logistics Efficiency (LEFT SIDE) vs. Service Quality (RIGHT SIDE)

	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9
Logistics Efficiency / Service Quality																	

Employee Satisfaction Metrics

Q7
Satisfaction with Working Hours vs. Fulfillment of Specific Driver Requests

Context: Consider the balance between satisfying general working hours preferences and fulfilling specific driver requests. How do you weigh the importance of ensuring satisfaction with working hours (Satisfaction with Working Hours) against meeting specific driver requests (Fulfillment of Specific Driver Requests)?
Question: On a scale of 9 to 1 to 9, how much more important is "Satisfaction with Working Hours" compared to "Fulfillment of Specific Driver Requests"?
Linear scale from 9 (Satisfaction with Working Hours is extremely more important) to 1 (Equal importance) to 9 (Fulfillment of Specific Driver Requests is extremely more important)

Satisfaction with Working Hours (LEFT SIDE) vs. Fulfillment of Specific Driver Requests (RIGHT SIDE)

	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9
Satisfaction with Working Hours / Fulfillment of Specific Driver Requests																	

Customer's Point of View Metrics

Q8
On-time Delivery Performance vs. Order Accuracy
Context: When planning routes, consider the trade-off between ensuring on-time deliveries and maintaining order accuracy. How do you weigh the importance of ensuring deliveries are on time (On-time Delivery Performance) against maintaining high accuracy in orders (Order Accuracy)?
Question: On a scale of 9 to 1 to 9, how much more important is "On-time Delivery Performance" compared to "Order Accuracy"?
Linear scale from 9 (On-time Delivery Performance is extremely more important) to 1 (Equal importance) to 9 (Order Accuracy is extremely more important)

On-time Delivery Performance (LEFT SIDE) vs. Order Accuracy (RIGHT SIDE)

9 8 7 6 5 4 3 2 1 2 3 4 5 6 7 8 9

On-time Delivery Performance / Order Accuracy

Level 3 Metrics

Logistics Efficiency KPIs

Q9
% of Realized km out of Planned km vs. Capacity Utilization
Context: Consider the balance between achieving the planned route distances and maximizing the use of available capacity. How do you weigh the importance of achieving planned kilometers (% of Realized km out of Planned km) against effectively utilizing available capacity (Capacity Utilization)?
Question: On a scale of 9 to 1 to 9, how much more important is "% of Realized km out of Planned km" compared to "Capacity Utilization"?Linear scale from 9 (% of Realized km out of Planned km is extremely more important) to 1 (Equal importance) to 9 (Capacity Utilization is extremely more important)

% of Realized km out of Planned km (LEFT SIDE) vs. Capacity Utilization (RIGHT SIDE)

	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9
% of Realized km out of Planned km / Capacity Utilization																	

Q10
% of Realized km out of Planned km vs. % Empty km's / Total km
Context: Consider the balance between achieving the planned route distances and minimizing empty kilometers. How do you weigh the importance of achieving planned kilometers (% of Realized km out of Planned km) against reducing empty kilometers (% Empty km's / Total km)?
Question: On a scale of 9 to 1 to 9, how much more or less important is "% of Realized km out of Planned km" compared to "% Empty km's / Total km"?Linear scale from 9 (% of Realized km out of Planned km is extremely more important) to 1 (Equal importance) to 9 (% Empty km's / Total km is extremely more important)

% of Realized km out of Planned km (LEFT SIDE) vs. % Empty km's / Total km (RIGHT SIDE)

	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9
% of Realized km out of Planned km VS % Empty km's / Total km																	

Q11
% of Realized km out of Planned km vs. Turnover per Hour per Driver
Context: Consider the trade-off between achieving the planned route distances and maximizing revenue per driver hour. How do you weigh the importance of achieving planned kilometers (% of Realized km out of Planned km) against maximizing revenue per driver hour (Turnover per Hour per Driver)?
Question: On a scale of 9 to 1 to 9, how much more or less important is "% of Realized km out of Planned km" compared to "Turnover per Hour per Driver"?Linear scale from 9 (% of Realized km out of Planned km is extremely more important) to 1 (Equal importance) to 9 (Turnover per Hour per Driver is extremely more important)

% of Realized km out of Planned km (LEFT SIDE) vs. Turnover per Hour per Driver (RIGHT SIDE)

9 8 7 6 5 4 3 2 1 2 3 4 5 6 7 8 9

% of Realized km out of Planned km / Turnover per Hour per Driver

Q12
Capacity Utilization vs. % Empty km's / Total km
Context: When planning routes, consider the balance between maximizing the use of available capacity and minimizing empty kilometers. How do you weigh the importance of effectively utilizing available capacity (Capacity Utilization) against reducing empty kilometers (% Empty km's / Total km)?
Question: On a scale of 9 to 1 to 9, how much more important is "Capacity Utilization" compared to "% Empty km's / Total km"?Linear scale from 9 (Capacity Utilization is extremely more important) to 1 (Equal importance) to 9 (% Empty km's / Total km is extremely more important)

Capacity Utilization (LEFT SIDE) vs. % Empty km's / Total km (RIGHT SIDE)

9 8 7 6 5 4 3 2 1 2 3 4 5 6 7 8 9

Capacity Utilization vs. % Empty km's / Total km

Q13
Capacity Utilization vs. Turnover per Hour per Driver
Context: Consider the trade-off between maximizing the use of available capacity and maximizing revenue per driver hour. How do you weigh the importance of effectively utilizing available capacity (Capacity Utilization) against maximizing revenue per driver hour (Turnover per Hour per Driver)?
Question: On a scale of 9 to 1 to 9, how much more important is "Capacity Utilization" compared to "Turnover per Hour per Driver"?Linear scale from 9 (Capacity Utilization is extremely more important) to 1 (Equal importance) to 9 (Turnover per Hour per Driver is extremely more important)

Capacity Utilization (LEFT SIDE) vs. Turnover per Hour per Driver (RIGHT SIDE)

9 8 7 6 5 4 3 2 1 2 3 4 5 6 7 8 9

Capacity Utilization / Turnover per Hour per Driver

Q14
% Empty km's / Total km vs. Turnover per Hour per Driver
Context: Consider the trade-off between minimizing empty kilometers and maximizing revenue per driver hour. How do you weigh the importance of reducing empty kilometers (% Empty km's / Total km) against maximizing revenue per driver hour (Turnover per Hour per Driver)?
Question: On a scale of 9 to 1 to 9, how much more important is "% Empty km's / Total km" compared to "Turnover per Hour per Driver"?Linear scale from 9 (% Empty km's / Total km is extremely more important) to 1 (Equal importance) to 9 (Turnover per Hour per Driver is extremely more important)

% Empty km's / Total km (LEFT SIDE) vs. Turnover per Hour per Driver (RIGHT SIDE)

9 8 7 6 5 4 3 2 1 2 3 4 5 6 7 8 9

% Empty km's / Total km vs. Turnover per Hour per Driver

Appendix 5 - MAHP Calculations

Short Code for KPI

KPI	Short Code
Level 1 Metrics	
Transportation effectiveness and efficiency	TEE
Employee Satisfaction	ES
Customer point of view	CPV
Level 2 Metrics	
Logistics Costs	LC
Logistics Efficiency	LE
Service Quality	SQ
Work Environment	WE
Employee Engagement	EE
Order Fulfillment Accuracy	OFA
Level 3 Metrics	
Total Distribution Cost	TDC
Capacity Utilization	CU
% Empty km's / Total km	EKM
Turnover per Hour per Driver	TPHD
On-time Delivery Performance	ODP
Satisfaction with Working Hours	SWH
Fulfillment of Specific Driver Request	FSDR
Order Accuracy	OA

Pairwise Comparison Matrices and Geometric Mean

Level 1

	P1	P2	P3	P4	P5	P6	Geometric mean
TEE vs ES	3	7	8	7	3	3	4.685705311
TEE vs CPV	3	0.14286	7	3	0.25	0.33333	0.953184293
ES vs CPV	1	0.14286	7	0.25	0.333333333	0.25	0.524557532

Level 2 Transportation Effectiveness

	P1	P2	P3	P4	P5	P6	Geometric mean
LC vs LE	1	0.25	0.166667	0.25	0.111111	1	0.324026883
LC vs SQ	3	0.333333	7	0.25	0.25	1	0.871290562
LE vs SQ	3	5	0.111111	1	0.5	1	0.970070115

Level 2 Employee Satisfaction

	P1	P2	P3	P4	P5	P6	Geometric Mean
WE vs EE	6	0.166667	0.125	6	0.142857	1	0.689171333

Level 2 Customer POV

	P1	P2	P3	P4	P5	P6	Geometric Mean
SQ vs OFA	3	0.142857	0.125	8	4	1	1.093991298

Level 3 Logistic Efficiency

	P1	P2	P3	P4	P5	P6	Geometrci Mean
CU vs EKM	1	1	8	6	7	1	2.636674665
CU vs TPHD	0.333333333	8	8	1	0.16667	1	1.235429341
EKM vs TPHD	3	6	8	0.16667	0.14286	1	1.227963714

Raw Weights (From Geometric Mean)

Level 1

	TEE	ES	CPV	SUM
TEE	1	4.685705311	0.953184293	6.638889604
CPV	0.213415043	1	0.524557532	1.737972575
CPV	1.049115063	1.906368586	1	3.955483649

Level 2 Transportation Effectiveness

	LC	LE	SQ
LC	1	0.324026883	0.871290562
LE	3.086163688	1	0.970070115
SQ	1.147722751	1.030853321	1

Level 2 Employee Satisfaction

	WE	EE
WE	1	0.689171333
EE	1.451017987	1

Level 2 Customer POV

	SQ	OFA
SQ	1	1.093991298
OFA	0.914084053	1

Level 3 Logistic Efficiency

	CU	EKM	TPHD
CU	1	2.636674665	1.235429341
EKM	0.3792656	1	1.227963714
TPHD	0.809435203	0.814356311	1

AI-systemlizing Raw Weights

Level 1

NORMALIZING		weights
Wtee	1.64684	0.48598
Wes	0.48195	0.14222
Wcpv	1.25992	0.3718
SUM	3.38871	

Level 2 Transportation Effectiveness

NORMALIZING		weights
Wlc	0.656016381	0.207933908
Wle	1.44125456	0.456826539
Wsq	1.057656448	0.335239553
SUM	3.154927389	

Level 2 Employee Satisfaction

NORMALIZING		weights
Wwe	0.830163437	0.407994
Wee	1.20458208	0.592006
SUM	2.034745517	

Level 2 Customer POV

NORMALIZING		weights
Wsq	1.04594039	0.522443097
Wofa	0.95607743	0.477556903
SUM	2.00201782	

Level 3 Logistic Efficiency

NORMALIZING		Weights
Wtdc		1
Wcu	1.482375241	0.473935207
Wekm	0.775133183	0.247820454
Wtphd	0.870293056	0.278244339
SUM	3.127801479	
Wodp		1
Wswh		1
Wfsdr		1
Wodp		1
Woa		1

Final Weights Before and After Multiple ODP Handled

	Final	new Final		
Wtdc	0.101051136	0.123020265		
Wcu	0.105217065	0.128091892		
Wekm	0.055017944	0.066979178		
Wtphd	0.061772268	0.075201933		
Wodp	0.162918775	0.217406058		
Wswh	0.058026292	0.07064156		
Wfsdr	0.084197194	0.102502174		
Wodp	0.194243992	0.217406058	Wop*final*	0.17858
Woa	0.177555335	0.216156941	New sum	0.82142

Consistency Check

Level 1

long max cal.			CI	RI	CR
TEE	1.50679	3.100531974	0.05027	0.58	0.086665
ES	0.44097	3.100531974			
CPV	1.15278	3.100531974			

Level 2 Transportation Effectiveness

long max cal.			CI	RI	CR
LC	0.648049046	3.116610718	0.058305359	0.58	0.100526
LE	1.423750487	3.116610718			
SQ	1.044811185	3.116610718			

Level 3 Logistic Efficiency

long max cal.			CI	RI	CR
CU	1.47110834	3.10403	0.05201	0.58	0.08968
EKM	0.769241727	3.10403			
TPHD	0.86367833	3.10403			

Appendix 6 - Evaluation Form Feedback

Monday, January 20, 2025

Dashboard Evaluation Form: Comparison Roadmap and KPI Dashboard

Thank you for taking the time to provide feedback on the Comparison Roadmap and KPI Dashboard. This form is designed to evaluate the usability, functionality, and impact of the dashboard and gather your suggestions for improvement.

This form includes:

Usability and Design Feedback: Ease of navigation, visual design, and accessibility.
Functionality and Technical Performance: Effectiveness of features, visualizations, and data integration.
Impact and Strategic Use: How the dashboard supports decision-making and aligns with organizational goals.
Open Comments: Space to share detailed suggestions or observations.

General Usability

How easy is it to navigate through the roadmap and dashboard? 4 / 5

What changes or improvements would make navigation easier?
Always being able to return to "home" with one click

Does the layout of the dashboard make it intuitive to understand the flow of data and KPIs? 3 / 5

Any suggestions for improving the layout or flow?
Not always, since ther is not always a link to the "home" page on the screen. In example, I do see different blocks on a screen, with different KPI's. But there is no sign, except the name of the "tab", what these kpi's are about (manual of AI) and where we are in the process regarding to the layout on the home screen.

Is the information presented on the dashboard relevant to your operational needs? 5 / 5

Are there any additional details or metrics you would like to see?
Not at this moment

Functionality

Are the visualizations (e.g., line charts, weighted composite scores) helpful in understanding trends and differences between systems? 4 / 5

Any feedback on specific visualizations?
It could be helpfull to see manual and AI performance in one graph to compare

Which KPIs or analyses would you add or remove?
None

Are there any technical limitations or bugs in the dashboard that you've observed?
Not that I am aware of

Impact

Do you feel the dashboard helps in decision-making? 4 / 5

Can you share specific examples of how it supports decision-making?
Yes, once being filled with operational data from the current manual plan and AI test plan

How likely are you to use this tool in routine evaluations of logistics performance? 4 / 5

What would make it more likely for you to use this regularly?
One overview with all kpi's where AI and Manual plan are being compared

Are there any technical features you would add to enhance functionality
No, MS BI offers sufficient possibilities

Aesthetic and Accessibility

Is the visual design of the dashboard appealing and professional? 4 / 5

What changes to the design could improve its appeal?
It is sufficient for its needs

Are all elements (charts, tables, slicers) accessible to users with limited technical expertise? 4 / 5