*Proceedings of the 11th North American International Conference on Industrial Engineering and Operations Management, Milwaukee, Wisconsin, USA, June 9-11, 2026*

*Proceedings of the International Conference on Industrial Engineering and Operations Management*

Publisher: IEOM Society International, USA DOI: 10.46254/NA11.20260067
Published: June 9, 2026

# Optimization of Service Level through Machine Learning, Linear Programming, and Standard Work in a Dark Kitchen

**Esteban Puerta-Román, Cielo Luis-Barreto, Gino Viacava-Campos and Jorge Antonio Corzo-Chavez**
Carrera de Ingeniería Industrial
Universidad de Lima, Lima, Perú
20215199@aloe.ulima.edu.pe, 20211543@aloe.ulima.edu.pe,
Geviacav@ulima.edu.pe, jacorzo@ulima.edu.pe

## Abstract

This research addresses service-level inefficiencies in the dark kitchen sector, a rapidly growing business model within the food delivery industry. The analyzed operation initially exhibited a service level of 78.2%, significantly below the target service level for last-mile delivery operations, highlighting the need for operational improvement during peak demand periods. To address this gap, an integrated improvement model was developed and validated through a before-and-after pilot study conducted under real operating conditions. The proposed model combines three complementary engineering tools: (i) a Machine Learning–based decision-support system for delivery routing, (ii) a Linear Programming model for optimal staff allocation during peak hours, and (iii) process standardization through digital Standard Work in the packing verification stage. Performance was evaluated using key indicators such as on-time delivery, order integrity, and labor productivity. The results show a 15.2% improvement in on-time delivery rate (from 79% to 91%), a 64% reduction in packing error rate (from 11.8% to 4.2%), and a 36.8% increase in staff productivity (from 12.5 to 17.1 orders per staff-hour), leading to an overall OTIF improvement from 78% to 91%. The main contribution of this study lies in the integrated application of predictive analytics, mathematical optimization, and process standardization within a single operational framework, providing a practical and replicable solution for dark kitchens and similar service-based food delivery operations facing high demand variability.

## Keywords
Dark Kitchen, Machine Learning, Linear Programming, Service Level, Workforce optimization, Standard Work.

### 1. Introduction
The accelerated growth of dark kitchens has transformed the gastronomic sector's operating model by concentrating operations exclusively on home delivery services. While this format offers advantages such as lower fixed costs and greater operational flexibility, it also poses significant challenges regarding delivery punctuality, order accuracy, and efficient resource utilization. In highly competitive urban markets, these variables directly influence customer satisfaction and business sustainability.

Service levels in delivery-oriented restaurants are typically evaluated using the On-Time In-Full (OTIF) indicator, which integrates timely delivery and order completeness. In last-mile delivery operations, a minimum OTIF threshold of 90% is widely recognized as the baseline for acceptable service performance, below which customer satisfaction and retention are significantly impacted (Kasoju et al., 2025; Zosu et al., 2024). Despite the sector's increasing digitalization, many dark kitchens still present operational gaps during peak demand periods, particularly in routing decision-making, staff allocation, and quality control at the packaging stage. These deficiencies lead to delays, order errors, and additional costs associated with reprocesses, service recoveries (courtesies), and customer churn.

In response to this gap, this study proposes a structured methodological framework aimed at improving the service level in a dark kitchen through the integration of three complementary components: (i) a Machine Learning model for optimal delivery routing and sequencing, (ii) a Linear Programming model for optimal packaging staff allocation during peak hours, and (iii) process standardization via digital Standard Work. This approach seeks to reduce operational variability and enhance the reliability of the dispatch process.
This work contributes to the literature by presenting an applied model that combines predictive analytics, mathematical optimization, and process standardization within the dark kitchen context, offering practical implications for operational management in food delivery services.

### 1.1 Objectives

The main objective of this research is to develop and evaluate the performance of an integrated operational model aimed at reducing service level inefficiencies in a dark kitchen environment. The proposed model combines delivery routing based on predictive analytics, optimal workforce planning, and process standardization to decrease delays, enhance staff efficiency, and reduce order errors:

- Develop and implement a predictive delivery routing and sequencing model using Machine Learning to improve delivery punctuality in scenarios of high demand variability.
- Formulate and apply an optimization model based on Linear Programming to enhance the allocation of packaging staff, thereby reducing bottlenecks and increasing labor productivity.
- Establish standardized digital verification procedures (through Standard Work) to minimize packaging errors and improve order completeness.

## 2. Literature Review

Following the operational diagnosis, this study presents a structured methodological framework designed to enhance service performance within a dark kitchen context. The framework consists of three coordinated components, each aimed at addressing a critical source of inefficiency. By combining predictive analytics, optimization techniques, and process standardization, the approach enables systematic improvements in delivery execution, labor utilization, and order accuracy. The following sections present the theoretical foundations supporting each component.

### 2.1. Customer Satisfaction and Service Reliability in Food Delivery

Customer satisfaction plays a central role in achieving customer loyalty within the restaurant industry. Empirical evidence indicates that perceived service quality has a significant direct and indirect effect on consumer loyalty in restaurant services, mediated by customer satisfaction (Ahmed et al., 2023). Consistently, Rezende G. et al. (2024) identify service quality as one of the most influential determinants of customer satisfaction. In the context of online food delivery services, reliability—accurate and timely order fulfillment—and empathy—personalized attention to customer needs—are key drivers of both satisfaction and loyalty (Arli D. et al., 2024). Similarly, Martinez et al. (2023) show that satisfaction with fundamental attributes such as taste and price has a stronger effect on loyalty than preferences for specific menu items, highlighting the importance of an overall positive experience. Moreover, innovation and a satisfactory customer experience significantly enhance satisfaction and encourage repeat patronage and positive word-of-mouth in fast-food restaurants (Manhas et al., 2024). These findings establish the operational rationale for improving on-time delivery performance, as service reliability is a primary driver of customer satisfaction in food delivery contexts.

### 2.2. Delivery Routing Optimization Using Machine Learning

Several studies have examined the application of Machine Learning to optimize order delivery in the restaurant sector. Jorge D. et al. (2024) propose a simulation–optimization tool that reduces delivery times by 4.5% by considering vehicle heterogeneity and traffic conditions; however, demand variability is excluded, limiting the model's applicability. Similarly, Huq F. et al. (2025) develop an optimization system combining mixed-integer linear programming with Machine Learning models (LSTM and XGBoost) to predict demand and travel times, significantly improving resource allocation, reducing costs, and increasing customer satisfaction by 20%. In contrast, Kancharla et al. (2024) incorporate uncertainty in preparation times and future orders using Sample Average Approximation and Adaptive Large Neighborhood Search with Machine Learning, achieving 25% more profitable routes and over 30% efficiency gains. Additionally, Yang et al. (2024) employ probabilistic models and greedy algorithms to predict driver order acceptance behavior, reducing delivery times. Cosmi et al. (2024) propose a deterministic MILP model prioritizing service quality by minimizing delays and rejections.

## 2.3. Workforce Allocation Through Linear Programming

Recent literature highlights that accurate workforce allocation is critical for maintaining adequate service levels in environments with high operational variability, such as dark kitchens. Linear Programming (LP) models provide a structured and quantitative approach for allocating human resources under operational constraints, including staff availability, demand fluctuations, and shift structures, while optimizing objectives such as cost minimization and service reliability. Bhuiyan and Mazumder (2024) demonstrate the applicability of LP for employee allocation in an emerging economy context, achieving significant cost savings and improved workforce efficiency without requiring advanced technologies. Their findings support the use of LP-based models as a practical and accessible method for optimizing staff assignment decisions in operations with variable demand, which aligns with the peak-hour staffing challenges observed in the present study.

## 2.4. Process Standardization in Service Operations

Process standardization is widely recognized as an effective mechanism for reducing operational variability and improving service reliability in high-throughput service environments. Standard Work, a core lean principle, establishes the most efficient and consistent sequence of tasks for a given process, enabling repeatable performance under variable demand conditions. Van Assen (2023) underscores the role of the service concept as a missing link in lean service implementation, emphasizing that standardized procedures must be aligned with the operational context to yield sustained improvements in service quality. In the dark kitchen environment, where peak demand creates time pressure that frequently leads to the omission of verification steps, the implementation of digital Standard Work checklists provides a structured countermeasure against operator-dependent errors, directly supporting the improvement of order completeness indicators.

# 3. Methods

The present study proposed an integrated service level improvement model applied to a dark kitchen, structured under a sequential input–intervention–output logic and based on the combined use of engineering tools. The model was validated through pilot tests conducted in a real operating environment.

As shown in Figure 1, the model consisted of three complementary components, designed to address the main root causes affecting the company's service level performance. These causes were primarily associated with inefficiencies in order delivery routing (42.3%), improper allocation of packing staff during peak demand periods (34.62%), and the omission of the verification step during such critical periods (23.08%).

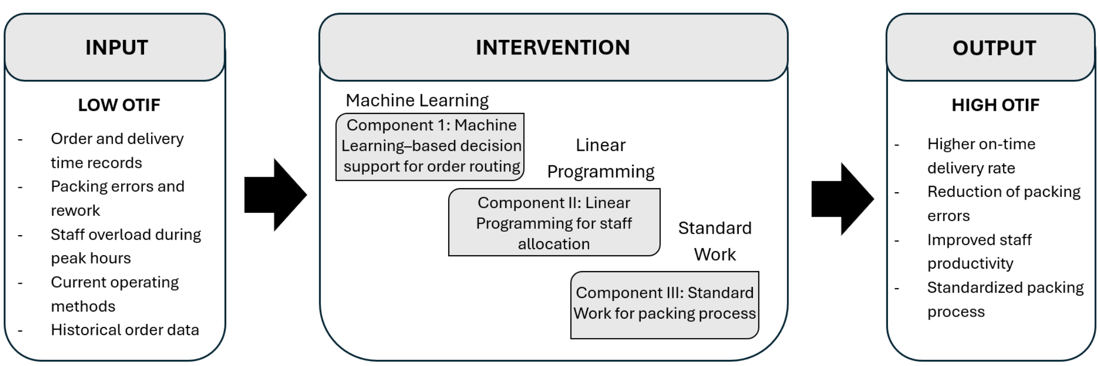



Figure 1. Proposed Model

**Model indicators**

The effectiveness of the proposed model was evaluated using a set of operational performance indicators, selected according to the main service-level problems identified in the dark kitchen. These indicators were defined to measure the impact of each component of the model under a before–after pilot comparison. Table 1 presents the indicators used to assess the performance of the proposed model and their relationship with each intervention component.

Table 1. Proposed Model Indicators

| Component | Key Indicator | Description | Unit |
|---|---|---|---|
| C1: Machine Learning–based decision support for order routing | On-time delivery rate | Percentage of orders delivered within the promised time window | % |
| | Average delivery time | Average time from order confirmation to delivery | min |
| C2: Linear Programming for staff allocation | Staff productivity | Number of orders processed per staff member per hour | orders/staff-hour |
| | Staff utilization | Degree of staff workload during peak hours | % |
| C3: Standard Work for packing process | Packing error rate | Percentage of orders with packing errors | % |
| | Checklist compliance rate | Percentage of orders verified using the QR checklist | % |
| Overall Performance | On-time in-full (OTIF) | Percentage of orders delivered on time and without errors | % |

**Component I: Machine Learning–based decision support for order routing**

Component I was designed to support order routing and dispatch decisions in the dark kitchen through a Machine Learning–based decision-support approach. A supervised tree-based regression model (specifically, a Decision Tree regressor, selected for its interpretability in operational dispatch contexts and its suitability for structured tabular data) was used to estimate delivery time based on historical order data, including delivery zone, order volume, and service time patterns. These delivery time estimates were used as inputs to prioritize and sequence orders during peak demand periods, supporting dispatch decisions under high demand variability. The use of delivery time–related indicators is consistent with recent studies on meal delivery routing problems, where on-time performance and delivery duration are key measures of routing effectiveness (Kancharla et al., 2024).

**Component II: Linear Programming for staff allocation**

Component II focused on improving the allocation of packing staff during peak demand periods through the application of Linear Programming. The model supported staff assignment decisions by determining the optimal distribution of available operators across packing activities, considering demand levels and workload constraints. This approach aimed to reduce workload imbalance and operational bottlenecks during peak hours. Similar Linear Programming–based approaches have been successfully applied to employee allocation problems, where productivity and utilization indicators are commonly used to evaluate workforce performance (Bhuiyan & Mazumder, 2024).

The LP model was formulated as follows. Let $x_t$ denote the number of operators assigned to packing during demand time slot t (t = 1, 2, …, T), $D_t$ the expected order demand during slot t, p the average operator productivity (orders/staff-hour), and S the total available operator pool:

Minimize $\quad Z = \Sigma_t \max(0, D_t - p \cdot x_t)$

Subject to:

$p \cdot x_t \geq D_t$ for all t [packing capacity meets demand]

$\Sigma_t x_t \leq S \cdot T$ [total staff availability across all slots]

$1 \leq x_t \leq S$ for all t [allocation bounds]

$x_t \in \mathbb{Z}^+$ for all t [integer constraint]

In the pilot implementation, T = 3 hourly slots per peak session were considered. Demand levels $D_t$ were estimated from historical order records, and productivity p was calibrated to the observed baseline of 12.5 orders/staff-hour. The model output guided the reallocation of operators to the packing cell during peak hours.

**Component III: Standard Work for packing process**

Component III implemented Standard Work to standardize the packing verification process through a digital checklist. This component ensured that all orders passed through a mandatory verification step prior to dispatch, particularly during peak demand periods. The objective was to reduce packing errors by ensuring consistent execution of operational tasks under high workload conditions. In service operations, process standardization is widely recognized as an effective approach to improve service reliability and reduce operational variability (Van Assen, 2023).

## 4. Data Collection

### Initial Data

Initial data collection was conducted to identify the main operational issues affecting service level performance in the dark kitchen. This phase combined the analysis of operational records with direct on-site observations carried out over a period of two consecutive weeks.

Observations focused primarily on peak demand periods, which were concentrated on Saturdays and Sundays between 3:00 p.m. and 6:00 p.m., identified as the time window with the highest operational workload. During this period, historical records of orders, delivery times, packing errors, and staff allocation were reviewed in order to detect recurring service-level failures.
The initial analysis revealed that service-level issues were persistent and systematically concentrated during peak demand periods rather than occurring randomly throughout the operation.
To ensure the statistical validity of the collected data, minimum sample sizes were estimated at a 95% confidence level ($Z = 1.96$).
For proportion-based indicators (OTD, OTIF, checklist compliance, and packing error rate), Cochran's formula was applied:

$$n_0 = Z^2 \cdot p(1-p) / e^2$$

assuming a conservative proportion $p = 0.5$ and a margin of error $e = 0.06$, resulting in a minimum sample size of 267 orders.
For average delivery time, the sample size was estimated using $n = (Z\sigma/E)^2$, with an observed standard deviation of 11 minutes and a precision of ±2 minutes, yielding a minimum of 116 orders.

A total of 305 orders were analyzed, exceeding the required sample sizes and ensuring statistical validity of the results.

### Identification and Prioritization of Root Causes

Based on the observed service-level failures during the two-week data collection period, each incident was classified according to its underlying root cause. The frequency of occurrence of each cause was quantified to identify the factors with the highest operational impact.

A total of 130 service-level failure incidents were recorded and classified during the observation period. The analysis showed that over 80% of these incidents were attributed to three main root causes: inefficient order routing (42.3%), improper staff allocation during peak demand periods (34.62%), and omission of the packing verification step (23.08%).

A Pareto analysis was performed at the root-cause level to prioritize improvement efforts and focus the proposed model on the most critical operational issues. Figure 2 presents the Pareto analysis of the main root causes of service-level failures identified during the diagnostic phase.

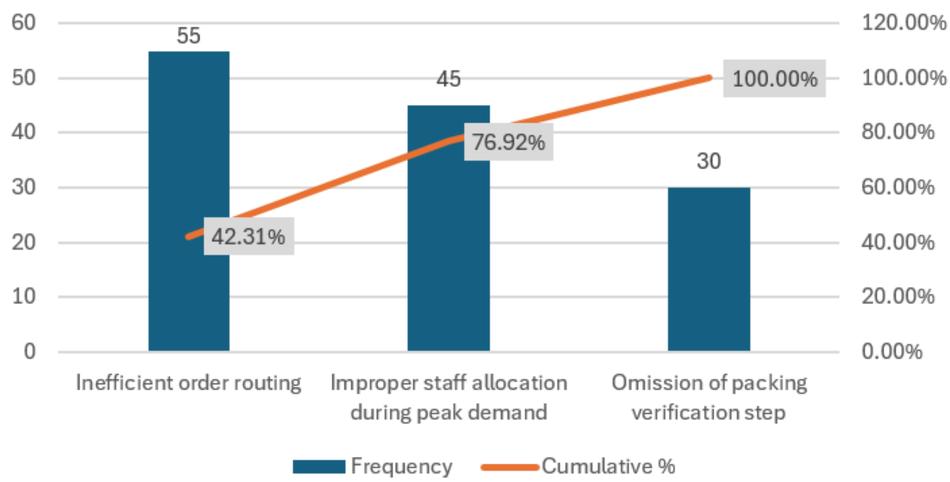



Figure 2. Pareto analysis of service-level failure root causes

As shown in Figure 2, the three identified root causes jointly accounted for the majority of observed service-level failures and were therefore selected for detailed analysis in the following subsections.

**Root Cause 1: Inefficient Order Routing**

To evaluate the impact of order routing on delivery performance, delivery time records were analyzed for all orders processed during peak demand periods over the two-week observation window. A total of 305 orders were evaluated.

The results showed that the average delivery time increased from 32 minutes during off-peak periods to 49 minutes during peak demand, with a standard deviation of ±11 minutes during peak hours. In addition, 21% of orders exceeded the promised delivery time window during peak periods, compared to 7% during off-peak hours.

These results indicate that routing decisions were not scalable under high order volume and relied primarily on operator judgment rather than structured criteria. The observed variability confirms that inefficient order routing is a major contributor to service-level degradation during peak demand periods.

**Root Cause 2: Improper Staff Allocation during Peak Demand**

To assess staff allocation efficiency, packing activity records and direct on-site observations were analyzed during peak demand periods. During these periods, an average of approximately 150 orders were processed within a three-hour window.

Although multiple operators were present in the operation, the packing activity was not formally reinforced during peak hours. In practice, packing tasks were predominantly handled by a single operator, while the remaining staff were assigned to other operational activities. This resulted in an effective packing capacity that did not scale with the increased order volume.

As a consequence, average packing productivity decreased to 12.5 orders per staff-hour during peak periods, compared to 18.2 orders per staff-hour during off-peak periods. Additionally, a recurring backlog of approximately 15 to 20 pending orders was consistently observed during peak demand periods. These findings indicate that the absence of a systematic, demand-based method for assigning packing staff during peak hours led to capacity bottlenecks, workload imbalance, and reduced operational performance, directly affecting service level indicators.

**Root Cause 3: Omission of Packing Verification Step**

To evaluate compliance with the packing verification process, direct observations were conducted during peak demand periods. A total of 180 orders were audited for verification compliance.

The analysis showed that verification was correctly performed in 92% of orders during off-peak periods, but compliance dropped to 68% during peak demand periods. In parallel, the packing error rate increased from 4.5% off-peak to 11.8% during peak hours. These results indicate that the omission of the verification step under high workload conditions is a structural issue that directly affects order accuracy and service reliability.

# 5. Results and Discussion

## 5.1 Proposed Improvements

The proposed improvement model consisted of three coordinated interventions applied to the order routing, staff allocation, and packing verification processes of the dark kitchen. Each intervention addressed one of the dominant root causes identified during the diagnostic phase and was designed to improve service reliability, reduce variability during peak demand periods, and minimize operator-dependent errors.

The integrated model was validated through a pilot implementation conducted under real operating conditions during peak demand periods (Saturdays and Sundays, 3:00 p.m. to 6:00 p.m.). The pilot followed an As-Is / To-Be comparison approach, allowing a direct evaluation of the impact of each intervention on service-level performance.

**Component I: Machine Learning–Based Decision Support for Order Routing**

The pilot for Component I evaluated the effectiveness of a Machine Learning–based decision-support approach for order routing during peak demand periods. During the As-Is stage, routing decisions were performed manually based on operator judgment. During the To-Be stage, the decision-support model was used to prioritize and assign orders to delivery routes considering order volume and delivery zones.

The pilot was conducted under comparable demand conditions, and delivery performance indicators were monitored throughout the implementation period.

In operational terms, the decision-support system ranks incoming orders during peak demand based on predicted delivery time, allowing dispatchers to prioritize orders that are more likely to breach the promised delivery window (Figure 3).

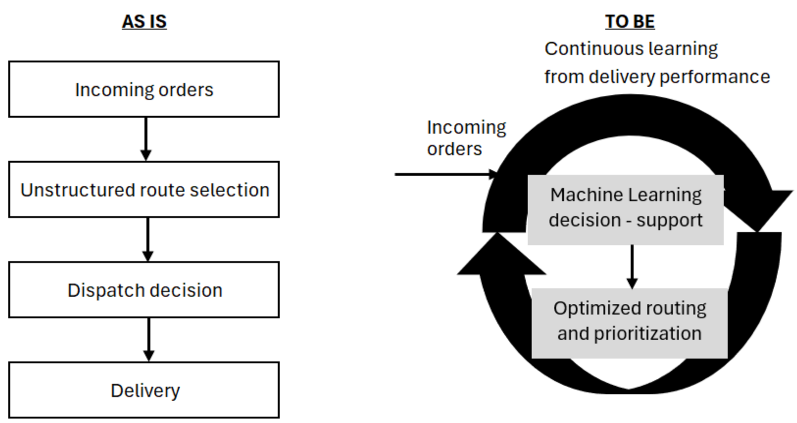



Figure 3. Order routing process before and after the implementation of the decision-support model.

**Component II: Linear Programming for Staff Allocation**

Component II evaluated the application of a Linear Programming model to determine the optimal allocation of packing staff during peak demand periods. Under the As-Is condition, staff allocation remained constant regardless of demand fluctuations. Under the To-Be condition, staff allocation was adjusted based on expected order volume during peak hours (Figure 4).

The pilot aimed to reduce workload imbalance and improve packing productivity while maintaining stable operational flow. Figure 4 illustrates the transition from a single-station packing layout to a flow-oriented packing cell, enabling better workload balance and reducing order backlog during peak demand periods.

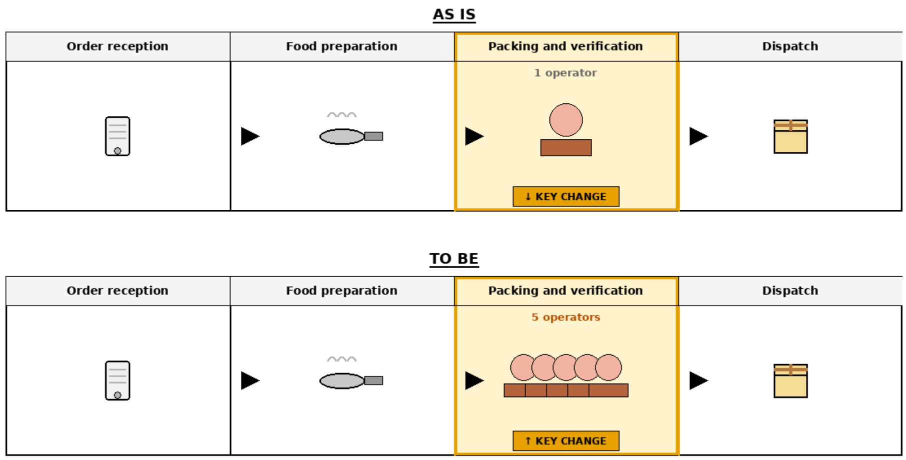
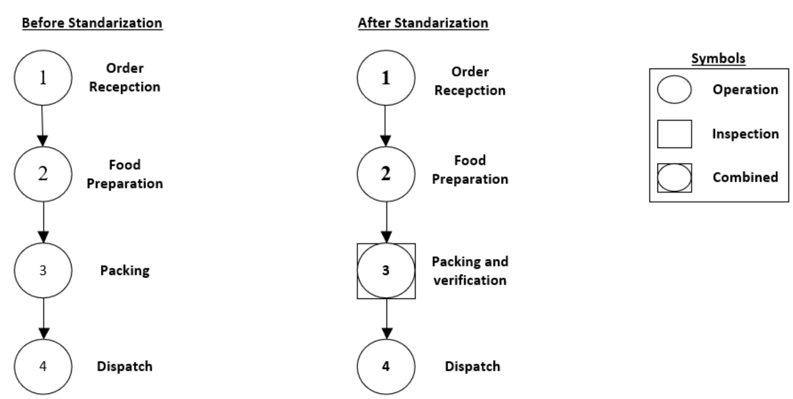



Figure 4. Packing cell layout before and after staff reallocation during peak demand periods

**Component III: Standard Work for Packing Verification**

The pilot for Component III implemented a standardized packing verification procedure supported by a digital checklist. During the As-Is stage, verification was performed inconsistently and frequently omitted during peak demand periods. During the To-Be stage, all orders were required to pass through the standardized verification step before dispatch.

This intervention aimed to reduce packing errors and improve order completeness (Figure 4).




| Summary Table | |
|---|---|
| **Activity** | **Quantity** |
| Operation | 4 |
| Inspection | 0 |
| Combined | 0 |
| **Total** | **4** |

| Summary Table | |
|---|---|
| **Activity** | **Quantity** |
| Operation | 3 |
| Inspection | 0 |
| Combined | 1 |
| **Total** | **4** |

Figure 5. Process Operation Diagram of the packing verification process before and after standardization

## 5.2 Validation

The pilot implementation was selected as the validation method for the proposed improvement model. Validation was conducted under real operating conditions during peak demand periods, defined as Saturdays and Sundays between 3:00 p.m. and 6:00 p.m., over a two-week observation window. Given the service-based and case-study nature of the research, data collection was performed using a census approach rather than statistical sampling. All orders processed during the selected peak demand periods were included in the analysis to ensure complete coverage of service-level failures and operational behavior. For Component I (order routing), delivery performance indicators were collected for a total of 305 orders, including delivery time and on-time delivery compliance. For Component II (staff allocation), packing productivity, backlog size, and staff utilization were monitored during all observed peak sessions. For Component III (packing verification), verification compliance and packing error rates were audited for 100% of observed orders during the pilot. This validation approach ensured consistent and reliable comparison between As-Is and To-Be conditions across all components, allowing the direct assessment of the impact of the proposed improvement model on service-level performance.

## 5.3 Numerical Results

The numerical results obtained from the pilot implementation provide a quantitative assessment of the impact achieved by each component of the proposed improvement model. The comparison between the As-Is and To-Be conditions was conducted using the key operational indicators defined in the methodology section. Table 2 summarizes the results observed during peak demand periods before and after the implementation of the proposed interventions.

Table 2. Validation results

| Component | Key Indicator | As-Is Value | To-Be Value | Improvement Achieved |
|---|---|---|---|---|
| C1: ML-based order routing | Average delivery time (min) | 49 | 36.5 | −12.5 min |
| | On-time delivery rate (%) | 79% | 91% | +12 pp |
| C2: Staff allocation (LP) | Orders per staff-hour | 12.5 | 17.1 | +4.6 ord/staff-hr |
| | Average order backlog (orders) | 18 | 6 | −12 orders |
| C3: Standard Work – packing verification | Packing error rate (%) | 11.80% | 4.20% | −7.6 pp |
| | Verification compliance (%) | 68% | 95% | +27 pp |
| Overall performance | OTIF (%) | 78% | 91% | +13 pp |

The results indicate a consistent improvement across all evaluated indicators.
Component I achieved a substantial reduction in average delivery time (from 49 to 36.5 minutes) and improved the on-time delivery rate by 12 percentage points, from 79% to 91%, by introducing structured routing decisions during peak demand periods.

Component II improved operational capacity by balancing staff workload, resulting in a 36.8% increase in productivity (from 12.5 to 17.1 orders per staff-hour) and a reduction in order backlog from 18 to 6 pending orders. Component III strengthened order accuracy through process standardization, reducing the packing error rate by 64% (from 11.8% to 4.2%) and increasing verification compliance from 68% to 95%. Collectively, the integration of the three components resulted in an increase in the overall OTIF indicator from 78% to 91% (+13 percentage points), confirming the effectiveness of the proposed model in improving service level performance under real operating conditions.

## 5.4 Graphical Results

Graphical analysis was used to evaluate the evolution and stabilization of the main service-level indicators throughout the pilot implementation. The selected figures illustrate the progressive impact of the proposed interventions and support the numerical results presented in the previous section.

**Evolution of OTIF during the pilot implementation**

The OTIF indicator showed a sustained improvement throughout the pilot period. As illustrated in Figure 6, OTIF increased from 78.2% under the As-Is condition to 86.3% in the second week and stabilized at 91.4% in the final week of the pilot. This improvement reflects the combined effect of structured order routing, optimized staff allocation during peak demand periods, and the standardization of the packing verification process. The stabilization observed in the final week indicates that the improvements achieved were not temporary but reached a consistent operating level under real demand conditions.

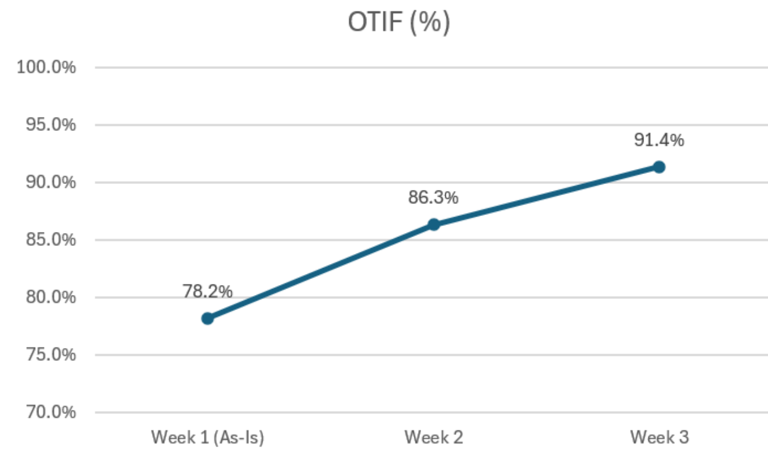



Figure 6. Weekly evolution of OTIF during the pilot implementation

Average delivery time during peak demand periods decreased notably after the implementation of the Machine Learning–based decision-support model for order routing. As shown in Figure 7, the average delivery time was reduced from 49.0 minutes in the As-Is condition to 36.5 minutes in the To-Be condition. This reduction demonstrates that structured routing decisions effectively mitigated delivery delays during high-order-volume periods, improving overall service responsiveness.

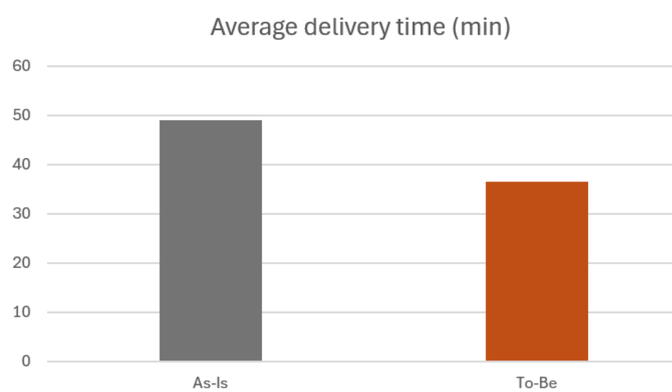



Figure 7. Comparison of average delivery time during peak demand periods

Packing errors exhibited a significant reduction following the implementation of the standardized packing verification procedure. As illustrated in Figure 8, the packing error rate decreased from 11.8% under the As-Is condition to 4.2% during the To-Be stage. The progressive reduction and stabilization of this indicator confirms that the introduction of Standard Work improved compliance with verification procedures and reduced order inaccuracies during peak demand periods.

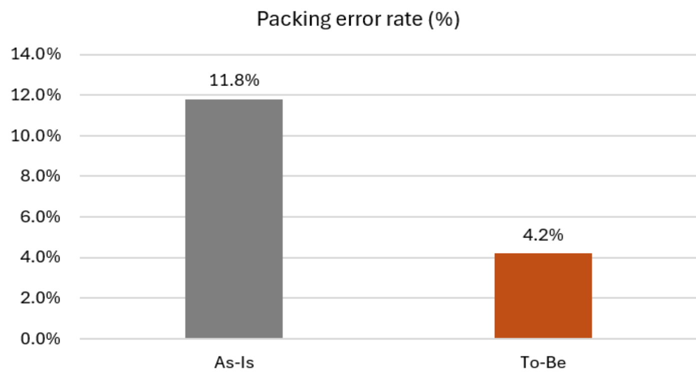



Figure 8. Packing error rate before and after standardization

### 5.5 Discussion

The improvement in on-time delivery rate from 79% to 91% and the reduction in average delivery time from 49 to 36.5 minutes (Component I) are consistent with findings reported by Huq et al. (2025) and Jorge et al. (2024), who demonstrated that ML-based demand prediction and structured sequencing can substantially reduce delays in food delivery contexts with high demand variability. The use of a Decision Tree regressor in this study provided an operationally interpretable model suited for dispatcher use, a practical trade-off relative to more accurate but less transparent ensemble methods, and one that is relevant to the dispatch-oriented application described by Cosmi et al. (2024) and Kancharla et al. (2024). Regarding Component II, the 36.8% productivity gain and backlog reduction from 18 to 6 orders corroborate the applicability of LP-based staffing models in variable-demand environments, as demonstrated by Bhuiyan and Mazumder (2024), and reflect a structural improvement in capacity management rather than a marginal efficiency adjustment.

The 64% reduction in packing error rate and the rise in verification compliance from 68% to 95% (Component III) are consistent with Van Assen's (2023) argument that the sustained impact of lean service interventions depends on how well standardized procedures are embedded in the operational context rather than offered as optional tools. The mandatory QR-based digital checklist satisfied this condition by reducing compliance friction precisely under the time pressure of peak demand periods, where operator-dependent omissions had been identified as the dominant root cause of order inaccuracies. This finding reinforces the view that the effectiveness of Standard Work is contingent not only on its technical design, but on its integration into the routine workflow.

At the aggregate level, the OTIF improvement from 78% to 91% demonstrates that the simultaneous optimization of routing, staffing, and verification yields synergistic effects beyond what any isolated intervention could produce independently. The progressive stabilization across pilot weeks (78.2% → 86.3% → 91.4%) suggests operational learning and adaptation rather than a one-time effect, and the service level achieved surpasses the 90% industry benchmark which the literature associates with meaningful gains in customer satisfaction and loyalty in food delivery contexts (Arli et al., 2024; Martinez et al., 2023). The modular architecture of the framework, with each component independently targeting a distinct root cause, allows practitioners to adopt the interventions selectively according to their specific operational gaps, increasing its replicability across diverse dark kitchen configurations.

## 6. Conclusion

This study demonstrated that the objectives of the proposed service-level improvement model were successfully achieved within the context of a make-to-order dark kitchen. The integrated application of analytical decision-support tools and standardized operational practices allowed the main sources of service-level failures to be effectively addressed during peak demand periods.

The implementation of a Machine Learning–based decision-support approach improved order routing efficiency, reducing average delivery time from 49 to 36.5 minutes and increasing the on-time delivery rate from 79% to 91%. The use of Linear Programming for staff allocation enhanced workload balance during peak hours, resulting in a 36.8% productivity increase and a reduction in order backlog from 18 to 6 pending orders. Additionally, the introduction of a standardized packing verification process reduced the packing error rate by 64% (from 11.8% to 4.2%) by minimizing operator-dependent omissions.

Collectively, these interventions increased the overall OTIF indicator from 78% to 91% (+13 percentage points), exceeding the industry benchmark of 90% and validating the effectiveness of the proposed integrated model under real operating conditions.
The primary contribution of this research lies in the development and validation of an integrated service-improvement framework tailored to the operational dynamics of dark kitchens. Unlike previous studies that address routing, staffing, or error reduction in isolation, this model simultaneously integrates these elements within a single operational structure, providing a practical and replicable pathway for similar service-based food operations facing high demand variability.

To further strengthen the model and achieve higher service-level performance, future iterations should incorporate long-term sustainability mechanisms. These include continuous monitoring of routing performance, periodic recalibration of staff allocation parameters, and audit routines to ensure sustained compliance with standardized verification procedures. Incorporating such control mechanisms would help preserve operational stability and enable further improvements in service reliability over time.

Regarding study limitations, the pilot was conducted within a single dark kitchen operation over a two-week observation window, which restricts the generalizability of the findings to operations with substantially different demand profiles, kitchen configurations, or staffing structures. Furthermore, the Decision Tree regression model was trained on historical data specific to this operational context, and its predictive accuracy may vary under different routing environments. Future research should evaluate the scalability of this integrated framework across multiple dark kitchen units, incorporate real-time dynamic routing updates, and assess the long-term sustainability of the achieved improvements through extended monitoring periods.

## References

Ahmed, S., Al Asheq, A., Ahmed, E., Chowdhury, U. Y., Sufi, T., and Mostofa, M. G., "The intricate relationships of consumers' loyalty and their perceptions of service quality, price and satisfaction in restaurant service," *The TQM Journal*, vol. 35, no. 2, pp. 519–539, 2023. https://doi.org/10.1108/TQM-06-2021-0158.

Arli, D., Van Esch, P., and Weaben, S., "The Impact of SERVQUAL on Consumers' Satisfaction, Loyalty, and Intention to Use Online Food Delivery Services," *Journal of Promotion Management*, pp. 1159–1188, 2024. https://doi.org/10.1080/10496491.2024.2372858.

Bhuiyan, J. A., and Mazumder, R., "Application of Linear Programming in Employee Allocation: A Case Study in Emerging Economy," *International Journal of Science and Research Archive*, vol. 11, no. 1, pp. 253–259, 2024. https://ijsra.net/content/application-linear-programming-employee-allocation-case-study-emerging-economy.

Cosmi, M., Oriolo, G., Piccialli, V., and Ventura, P., "Courier assignment in meal delivery via integer programming: A case study in Rome," *Omega (United Kingdom)*, 2024. https://doi.org/10.1016/j.omega.2024.103237.

Huq, F., Sultana, N., Roy, P., Razzaque, M., Huda, S., and Hassan, M., "Cross regional online food delivery: Service quality optimization and real-time order assignment," *Computers and Operations Research*, 2025. https://doi.org/10.1016/j.cor.2024.106877.

Jorge, D., Rocha, T., and Ramos, T., "A time-driven simulation–optimization framework for the dynamic heterogeneous order-courier assignment problem for instant deliveries," *Transportation Research Part E: Logistics and Transportation Review*, 2024. https://doi.org/10.1016/j.tre.2024.103783.

Kancharla, S. R., Van Woensel, T., Waller, S. T., and Ukkusuri, S. V., "Meal delivery routing problem with stochastic meal preparation times and customer locations," *Networks and Spatial Economics*, vol. 24, pp. 997–1020, 2024. https://doi.org/10.1007/s11067-024-09643-1.

Kasoju, A., Vishwakarma, T., and Kasoju, A., "The role of AI-enhanced fast delivery services in strengthening customer retention and loyalty in competitive markets," *Frontiers in Artificial Intelligence*, 2025. https://doi.org/10.3389/frai.2025.1612772.

Manhas, P., Sharma, P., and Quintela, J., "Product Innovation and Customer Experience: Catalysts for Enhancing Satisfaction in Quick Service Restaurants," *Tourism and Hospitality*, pp. 559–576, 2024. https://doi.org/10.3390/tourhosp5030034.

Martinez, M., Ruiz, P., Izquierdo, A., and Pick, D., "The influence of food values on satisfaction and loyalty: Evidence obtained across restaurant types," *International Journal of Gastronomy and Food Science*, 2023. https://doi.org/10.1016/j.ijgfs.2023.100770.

Rezende, G., Mariano, A., Santos, M., and Coelho, A., "Modeling Customer Satisfaction in the Food Industry: Insights from a Structural Equation Approach," *Procedia Computer Science*, pp. 130–137, 2024. https://doi.org/10.1016/j.procs.2024.08.250.

Van Assen, M. F., "The service concept – a missing link in lean for service operations," *Managing Service Quality*, vol. 33, no. 6, pp. 899–916, 2023. https://doi.org/10.1080/14783363.2023.2222064.

Yang, Y., Umboh, S., and Ramezani, M., "Freelance drivers with a decline choice: Dispatch menus in on-demand mobility services for assortment optimization," *Transportation Research Part B: Methodological*, 2024. https://doi.org/10.1016/j.trb.2024.103082.

Zosu, S. J., Amaghionyeodiwe, C., Oyetunji, E., and Yussouf, A., "Last-mile delivery optimization: Balancing cost efficiency and environmental sustainability," *International Journal of Emerging Trends in Engineering Research*, vol. 12, no. 11, pp. 153–166, 2024. https://doi.org/10.30534/ijeter/2024/0112112024.

## Biographies

**Cielo Luis Barreto** is a student of Industrial Engineering at University of Lima with experience in the financial area.

**Esteban Puerta Roman** is a student of Industrial Engineering at University of Lima with experience in the commercial and financial area.

**Gino Viacava-Campos** is currently a professor in industrial engineering career at Universidad de Lima. He holds a Master in Operations and Logistics from Universidad Peruana de Ciencias Aplicadas (UPC) (Perú) and is Industrial Engineer from the Universidad de Lima (Perú). He has experience in industrial and service companies to improve productivity, development of business models and structure compensation systems. He has worked in human resources area, operations and logistics areas of a transnational company, locally and overseas. As consultant he has experience with construction, manufacturing, services, mining, commercial, finance and education sector. His main research interests are lean manufacturing, master production schedules, logistics, compensation systems for small and medium-sized enterprises.

**Jorge Antonio Corzo-Chavez** is currently a professor in industrial engineering career at Universidad de Lima, chief executive officer at Conquista Lab, and associate consultant at Manpower Group for multinational companies. He holds a Master in Business Administration from Universidad ESAN (Perú), a Master in Marketing Science from ESIC Business & Marketing School (Spain) and is Industrial Engineer from the Universidad de Lima (Perú). He has experience in advising industrial and service companies to improve productivity and development of business models. He has worked in areas of entrepreneurship and innovation, marketing and commercial for the construction, metalworking, mining, services, and education sectors. His main research interests are lean manufacturing, master production schedules, logistics, machine learning, and innovation for small and medium-sized enterprises.