AI-Driven Sales and Distribution Planning for Sustainable Logistics Operations in Saudi Arabia's FMCG Industry

Abstract

Saudi Arabia's fast-moving consumer goods (FMCG) sector depends on reliable replenishment across diverse outlets, dense urban markets and long intercity freight corridors. Conventional sales planning often treats demand forecasts, distributor allocation, inventory deployment and vehicle dispatch as separate activities. This review examines how artificial intelligence (AI) can connect these decisions while supporting the sustainability priorities of Saudi Vision 2030 and the National Transport and Logistics Strategy. A structured critical narrative approach integrates peer-reviewed studies on predictive demand planning, machine learning, operations research, logistics performance and responsible deployment with official Saudi policy sources. The proposed framework links validated data, hierarchical demand forecasting, stock-aware order planning, constrained vehicle-routing optimization and an auditable sustainability scorecard. It identifies important distinctions between predictive accuracy and operational value: a better sales forecast may fail to reduce freight emissions if minimum shipment quantities, delivery windows or incentives prevent execution. Illustrative scenarios demonstrate trade-offs rather than report observed Saudi company results. The review argues for phased, human-supervised deployment, explicit service and emissions constraints, transparent model monitoring and cross-functional governance. Further Saudi-specific longitudinal and intervention research is required to test economic, social and environmental effects.

Keywords: artificial intelligence; FMCG; sales planning; distribution networks; sustainable logistics; demand forecasting; Saudi Vision 2030; route-to-market; optimization

UNDER PEER REVIEW IN IJAR

1. Introduction

Distribution planning is often portrayed as a transport function, yet the operational choices that determine transport outcomes begin much earlier. Sales teams negotiate promotions and outlet coverage; demand planners estimate replenishment; warehouses determine pick waves; finance teams approve commercial conditions; and transport planners assign available vehicles. When those decisions are produced independently, a route can be mathematically efficient while carrying the wrong stock, serving the wrong customer priority or generating costly returns. This coordination problem is especially important for FMCG products, which combine low margins per unit, high delivery frequency, variable shelf availability, short product life for some categories and strong sensitivity to seasonal promotional activity. The consequence is not only financial inefficiency. Unnecessary vehicle kilometres, avoidable spoilage, packaging damage and emergency deliveries can also increase environmental burden.

Saudi Arabia provides a consequential setting because the Kingdom is expanding logistics capabilities as part of Vision 2030. The National Transport and Logistics Strategy seeks integrated transport services, stronger competitiveness and improved logistical connectivity; however, national infrastructure ambitions do not by themselves guarantee efficient last-mile commercial delivery (Ministry of Transport and Logistic Services, n.d.). FMCG firms must translate the national direction into granular decisions at the store, SKU, delivery, driver and depot levels. Large cities such as Riyadh, Jeddah and Dammam feature dense but congested delivery markets, whereas peripheral and intercity territories require coordination over longer distances. This paper considers those distinct operational environments without claiming one uniform Saudi routing pattern.

Artificial intelligence is useful here because the supply chain produces high-volume transactional and operational records that can illuminate relationships between demand, promotions, stock, location and transport disruption. Nonetheless, machine learning is not a substitute for constraint-aware operations research. A forecast estimates what may occur; optimization decides what should be shipped and when under available capacity, customer time windows and financial objectives. Successful integration requires causal caution, executable decision rights, human involvement and reliable measurement. This review therefore treats AI-driven planning as a socio-technical operating system rather than as a standalone algorithm.

2. Research Aim, Objectives and Questions

The aim is to critically assess how AI-supported sales and distribution planning can improve logistics efficiency and sustainability in Saudi FMCG networks, and to develop a defensible conceptual implementation framework. The objectives are to: (1) examine the data and forecasting needs of FMCG demand planning; (2) evaluate optimization approaches for allocation, replenishment and delivery routing; (3) identify measurable environmental, service and commercial impacts; (4) assess institutional and organizational conditions in Saudi Arabia; and (5) propose an implementation and evaluation agenda that avoids unsupported performance claims.

Five questions guide the review. RQ1: Which sales, inventory and transport data are needed for executable joint planning? RQ2: Where can machine learning add value beyond conventional forecasts and decision rules? RQ3: How should service reliability, logistics costs and environmental effects be balanced? RQ4: What governance and organizational constraints are likely to affect Saudi deployment? RQ5: Which research designs could establish causal rather than merely correlational effects of integrated AI planning?

3. Review Methodology

The study uses a structured critical narrative review, not a formal systematic review or original field experiment. Academic literature was identified through targeted searches of publisher pages, scholarly indexes and research-oriented search engines, emphasizing studies on demand planning, retail forecasting, AI supply chains, vehicle routing, sustainable distribution and implementation between 2020 and 2026. Earlier foundational research is included selectively when necessary to explain forecasting hierarchies or optimization. Saudi policy statements were checked against

UNDER PEER REVIEW IN IJAR

official transport and Vision 2030 sources. Because search access, database coverage and citation screening were not uniform, the review does not report a fabricated record count or claim exhaustive coverage.

The search logic combined terms such as “artificial intelligence” OR “machine learning” with “FMCG”, “retail demand forecasting”, “supply chain planning”, “sustainable distribution”, “emissions”, “vehicle routing” and “Saudi Arabia”. Inclusion prioritized peer-reviewed methodological and empirical work, authoritative surveys and official government documents directly relevant to the decision chain. Exclusion applied to untraceable performance claims, marketing-only case studies without methods and sources whose use of AI lacked a meaningful planning connection. Each selected source was examined for decision level, assumptions, outcome metrics, geographic setting and transferability. Evidence from non-Saudi datasets is treated as methodological guidance, not automatically as Saudi effectiveness evidence.

Thematic synthesis is organized around forecast quality, commercial decision integration, transport planning, sustainability measurement and governance. This approach enables comparison across distinct fields but has limits: publication bias may favor successful AI use cases, proprietary logistics data are rarely open, and causal validation is limited. Rather than report a pooled improvement percentage, the paper identifies decision mechanisms and proposes prospective tests with pre-defined baselines, uncertainty ranges and independent audit of measured results.

Table 1. Evidence-to-decision synthesis

Evidence stream	What it informs	Critical limitation
Demand forecasting	SKU-outlet replenishment and promotional planning	Improved error metrics may not increase availability
Human-guided analytics	Sales override and planner oversight	Expert interventions can introduce bias
Optimization research	Routes, fleet loads, time windows and service costs	Simplified assumptions may omit local practice
Saudi logistics strategy	Policy relevance and infrastructure context	National objectives are not company-level outcomes
Sustainability accounting	Fuel, emissions, spoilage and service trade-offs	Boundary choices can shift apparent savings

4. Saudi FMCG Logistics and Vision 2030

The Saudi market presents a heterogeneous network of hypermarkets, supermarkets, convenience stores, wholesalers, food-service customers and smaller neighborhood outlets. A distributor may simultaneously operate direct store delivery, third-party transport and wholesale arrangements, while sales planners serve differentiated territories with different order cadence and merchandising obligations. Demand spikes can be influenced by religious holidays, school calendars, tourism activity, advertising and local purchasing patterns. These factors matter because an aggregate monthly forecast may conceal the volatility that determines individual truck loads and shelf-level

UNDER PEER REVIEW IN IJAR

availability. For sustainable improvement, planning needs a hierarchy that connects national sales ambition to feasible delivery-day execution.

Vision 2030's logistics initiatives emphasize national connectivity, performance and integration across modes; its 2025 annual reporting also describes the expansion of logistics hubs and digital services (Saudi Vision 2030, 2026). For FMCG distribution, the practical interpretation is to strengthen data exchange between manufacturing or import gateways, central warehouses, regional nodes and store-level delivery. A retailer-facing AI platform could use these connections to reduce ad hoc transfers and schedule deliveries before service failures occur. Yet references to multi-modal networks should be proportionate to the actual business: many FMCG last-mile flows are road-based, and claiming rail or port optimization without operational data would be inappropriate.

Sustainability should not be reduced to fuel savings alone. Product waste from damaged or expired stock, labor conditions during extreme weather, refrigeration energy, loading efficiency, packaging returns and service access in less dense territories all matter. A narrow routing objective that minimizes kilometres might shift deliveries to inconvenient times, increase the retailer's storage burden or degrade product freshness. The Saudi context therefore requires an integrated scorecard that weights delivered service, fleet activity, inventory losses and worker feasibility. The government's strategic ambition provides the policy rationale, but company-level accountability depends on transparent baselines and operational evidence.

5. Conceptual Basis: Integrated Sales and Distribution Planning

Traditional planning often follows a linear sequence: sales teams provide projected volumes; supply planning allocates inventory; distribution receives an order file; and transport executes daily routes. This division can create a bullwhip effect when promotional signals are exaggerated or when shortage-driven lost sales are interpreted as weak demand. AI enables probabilistic forecasts and scenario testing, but the central design decision is whether predictions can be translated into shared constraints. A good system recognizes shelf capacity, service windows, transport utilization, stock freshness, local sales priorities and the economic value of unmet orders. It must also expose trade-offs to managers rather than hiding them inside a single proprietary score.

Consider three interacting decisions. Strategic planning determines depots, contracted carriers, territories and assortment architecture; tactical planning sets coverage frequency, replenishment policies, fleet capacity and seasonal promotions; operational planning assigns tomorrow's quantities, delivery sequences, load priorities and exception actions. Decisions should be linked but not recomputed with equal frequency. Depot placement is relatively stable, whereas demand signals and traffic conditions change daily or hourly. A layered model can protect stable commitments while allowing bounded operational adaptation. Without these boundaries, a continuously learning model may recommend infeasible changes that conflict with contracts, customer agreements or safety requirements.

The proposed structure comprises four interacting layers: descriptive visibility, predictive intelligence, constrained decision optimization, and monitored execution. Descriptive records provide historical sales, orders, stock and GPS trajectories; models estimate future demand and

UNDER PEER REVIEW IN IJAR

uncertainty; optimization chooses replenishment and routes; execution systems capture delivery confirmations, delays and returns. Feedback is evaluated for both predictive error and realized business results. Human override remains explicit because sales campaigns and customer commitments may contain information that has not yet entered a database.

6. AI for Demand and Sales Intelligence

Reliable demand forecasts are an important input, but observed sales are not always equal to customer demand. When a shelf is empty, recorded sales fall even if customers continue to want the product. Likewise, an exceptional discount can temporarily increase purchases while shifting future consumption. Training models on raw invoiced quantities without recording stockouts, returns or price changes can therefore encode misleading patterns. Analysts should distinguish sell-in to the distributor, sell-out to the shopper, dispatch quantity and true demand wherever possible. Forecast horizons should match decision lead times: a weekly warehouse plan has different requirements from the morning vehicle dispatch schedule.

The literature supports a broad model comparison rather than automatic selection of deep learning. Seasonal naïve methods, exponential smoothing and statistical regression offer interpretable baselines. Gradient-boosted trees accommodate heterogeneous covariates such as price, promotions, weather and local market features. Sequence models can represent complex temporal dependencies when enough quality data exist. Hybrid convolutional and recurrent methods have shown value on retail datasets, but comparative gains depend on features, horizons and dataset design (Castro Moraes et al., 2024). An approach that performs well in another country cannot be assumed to perform equally well across Saudi retail territories without local validation.

Forecast evaluation should use rolling-origin backtesting that mirrors real planning dates and excludes future information. Weighted absolute error, bias, forecast value added and probabilistic interval coverage answer different questions; a single mean percentage error can be distorted by low-volume SKUs. Disaggregation by city, product family, channel and promotion period reveals whether gains serve commercially important cohorts. If a high-accuracy model systematically underpredicts demand for remote stores, aggregate error may improve while service equity worsens. Consequently, the correct forecasting objective is not merely lower statistical loss but more dependable distribution decisions under explicit business priorities.

Human judgment remains valuable. Planners may know about an upcoming retail promotion or a store closure before that event becomes observable in historical data. Brau and colleagues (2023) examine how human information can be integrated with predictive analytics rather than appended as uncontrolled overrides. A Saudi FMCG system could require an override reason, time stamp, confidence level and after-action comparison. Such governance makes expert input measurable and enables learning, while preventing untraceable manual changes from defeating model evaluation.

7. Distribution Allocation, Replenishment and Stock Health

A distribution network can improve fleet utilization and still fail commercially if the wrong products are allocated to the wrong outlets. AI-driven sales planning should therefore connect

UNDER PEER REVIEW IN IJAR

demand forecasts to constraints on available-to-promise inventory, case-pack sizes, customer service levels, truck capacity and product freshness. In a multi-echelon system, stock may pass from a national distribution center to regional warehouses and then to retail outlets. Allocation rules must account for lead time, planned promotional visibility and the cost of transferring units between nodes. Algorithms may help estimate which demand is most likely to materialize, but policies must decide how scarce stock is distributed across customers.

Inventory objectives can conflict. High safety stock improves fill rate but increases storage cost and potentially waste, particularly for temperature-sensitive food. Lean inventories reduce holding costs but may require more emergency replenishments and small deliveries. An integrated system can evaluate the expected total cost of shortages, excess inventory, planned transport and exceptional trips. Uncertainty should be represented through prediction intervals or scenarios instead of treating a single-point forecast as certainty. This becomes important when religious holidays, travel movements or new product introductions make the demand process less stable.

Returns and reverse logistics deserve special attention. Unsold promotional stock, damaged packaging and expired items can generate backhaul movements that are invisible in forward-route optimization. A coordinated plan may pair returns with outbound deliveries where hygiene and product-handling conditions allow, avoiding separate travel. Nevertheless, the sustainability benefit must be measured after accounting for additional handling, possible contamination restrictions and the true destination of returned material. More efficient transportation cannot compensate for poor demand governance if it merely moves an unnecessarily produced item more cheaply.

8. AI-Supported Transport and Route Planning

Daily route planning converts orders into executable journeys. A useful model must consider vehicle capacities, driver shifts, retailer receiving hours, loading sequences, traffic-dependent travel times, delivery durations, product segregation and return collections. Classical vehicle-routing optimization provides a disciplined framework for these constraints. Machine learning can improve the estimates supplied to the optimizer: expected travel time, service duration, failed-delivery probability and order uncertainty. The combined architecture is generally more defensible than relying on a black-box prediction model to generate unrestricted routes.

Sales territory design should not be confused with dispatch sequencing. A salesperson’s territory may be shaped by relationship coverage, merchandising intensity and commercial potential, whereas delivery territories seek feasible loading and travel patterns. Integration can reduce duplicative travel, but forcing identical territories onto both sales and transport organizations can degrade service. A better planning layer coordinates scheduled visits, delivery cycles and strategic customers while allowing role-specific execution. This distinction connects the paper directly to senior sales and distribution management rather than representing the research as pure transport engineering.

Sustainable routing is multi-objective. Minimizing kilometres can conflict with reducing emissions when routes have different congestion profiles, payloads or vehicle types. Minimizing cost may incentivize full but late trucks, while minimizing delivery lateness may create nearly empty

UNDER PEER REVIEW IN IJAR

emergency vehicles. The solution can use weighted objectives or constrained thresholds, such as a minimum on-time delivery rate and a maximum emissions budget. Decision-makers should inspect Pareto-efficient choices rather than accepting a single "optimal" route that obscures assumptions. In pilot studies, actual fuel usage and route adherence should be captured before claims of emissions savings are made.

9. Sustainability Metrics and Measurement Boundaries

Environmental metrics require explicit accounting boundaries. Tailpipe CO2-equivalent estimates can be computed from verified fuel consumption and appropriate factors; broader well-to-wheel calculations also account for fuel production, and product-level life-cycle assessments include upstream manufacture and packaging. The selected boundary must be stated before comparing scenarios. Using straight-line distance instead of logged vehicle kilometres, or assuming uniform fuel consumption despite varying load and congestion, creates fragile results. The most feasible first-stage metric for many fleets is recorded litres per delivered case, accompanied by total fuel and service quality.

Operational measures should include on-time-in-full delivery, cost per successfully delivered case, kilometres per stop, truck cube utilization, failed deliveries, freshness-related returns and product waste. Measuring only the average can conceal deterioration for peripheral areas or small customers. Sustainability dashboards should display absolute quantities as well as intensity ratios: lower emissions per case can coexist with higher total emissions if volumes grow rapidly. The company also needs worker-centered checks such as safe driving hours, safe unloading requirements and predictable service schedules.

Attribution is difficult when promotions, fuel costs, retail expansion, road conditions or customer mix change during a pilot. A before-and-after comparison without a control group may wrongly credit AI for unrelated operational changes. Where feasible, comparable route clusters can be phased into deployment, with pre-specified outcomes, matched baselines and seasonality controls. Independent review of route logs, invoices and fuel records would provide a stronger basis for Saudi-specific evidence than promotional case studies reporting only percentage improvements.

Table 2. Proposed FMCG logistics sustainability scorecard

Dimension	Suggested measure	Interpretation / safeguard
Commercial	Service-adjusted contribution margin	Do not hide missed outlet demand
Customer service	On-time-in-full delivery (%)	Report by region and outlet class
Fleet	Fuel litres / 1,000 delivered cases	Use actual fuel records when possible
Network	Empty kilometres / total kilometres	Account for backhauls and deadheading
Inventory	Expired and damaged cases (%)	Segment by product shelf life

UNDER PEER REVIEW IN IJAR

Workforce	Driving hours and safety exceptions	Respect legal and employer requirements
Model quality	Forecast bias and interval coverage	Audit drift by region and SKU

10. Proposed Integrated AI Planning Architecture

A practical architecture begins with a governed data layer joining ERP invoices, order management, warehouse stocks, customer attributes, promotions, vehicle telemetry, proof of delivery and returns. Standard identifiers are critical because outlet, distributor and SKU codes may differ across platforms. Data preparation should flag missing delivery times, duplicate orders and outlier quantities. Privacy and commercial confidentiality controls must be applied to retailer transactions, employee location records and account-level performance. Data are then organized in a feature store with versioned definitions and a reproducible training pipeline.

The predictive layer produces a hierarchy of forecasts and uncertainty estimates at relevant planning levels. A separate service model estimates expected stop time and delivery failure risk, while a travel-time model learns traffic and territory effects. These models feed an optimization layer that balances inventory position, allocation priority, vehicle capacities, delivery windows and sustainability thresholds. A decision interface exposes feasible alternatives and their expected implications; commercial managers can approve constrained overrides with reasons that remain auditable. Execution data return to the system for quality monitoring, exception analysis and incremental retraining.

It is useful to state a conceptual objective. Let C represent distribution and inventory cost; E estimated emissions; W preventable waste; and L unmet service commitments. The network seeks to minimize a weighted function $J = \alpha C + \beta E + \gamma W + \delta L$ subject to demand, stock, fleet, route and labor feasibility. The weights represent management priorities and cannot be assumed universal. A more transparent implementation may set minimum service standards and emissions ceilings before minimizing avoidable expense. This formulation is illustrative, not a solved optimization model or evidence of implemented company software.

UNDER PEER REVIEW IN IJAR

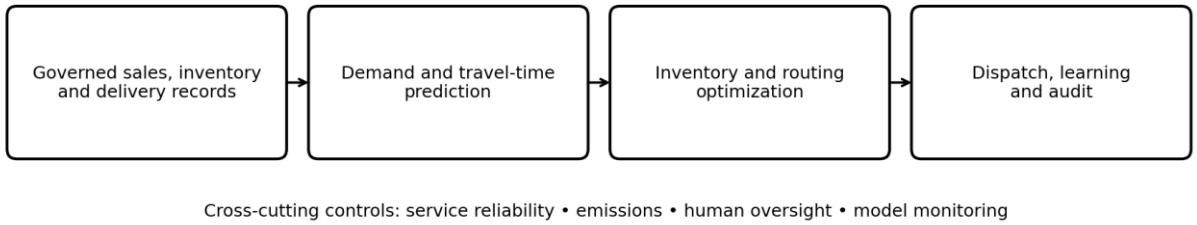


Figure 1. Proposed architecture for integrated AI-driven sales and sustainable distribution planning (author-developed conceptual figure).

11. Illustrative Scenario Analysis: Interpreting Trade-offs

To make the decision problem tangible, Table 3 and Figure 2 present hypothetical planning outcomes for a distribution region. They are designed for methodological discussion only and are not estimates of results for Pladis or any other Saudi firm. The scenario compares a baseline with a forecast-led plan, an integrated sales-and-routing plan, and an integrated plan subject to an explicit service safeguard. Values are normalized to baseline 100 so that the reader can inspect relative changes without confusing invented quantities with actual fuel or revenue savings.

The example demonstrates why the "best" approach depends on what is protected. Forecast-led replenishment alone can reduce avoidable inventory but may increase delivery frequency if the transport plan is unchanged. Integrated planning can reduce travel and handling, yet a cost-only optimization may compromise on-time service. Adding a service threshold may require a modest operational sacrifice while protecting customer reliability. These directional patterns are plausible hypotheses for field testing, not general numerical effect sizes. A real evaluation would require observed vehicle activity, promotions, demand and returns over enough weeks to measure variability.

Table 3. Hypothetical normalized scenario values (baseline = 100)

Scenario	Logistics cost	Fuel use	Product waste	Service score
Current rule-based plan	100	100	100	100
Forecast-led only	97	101	93	101
Integrated cost-led AI plan	88	91	89	96
Integrated AI with service constraint	91	93	90	103

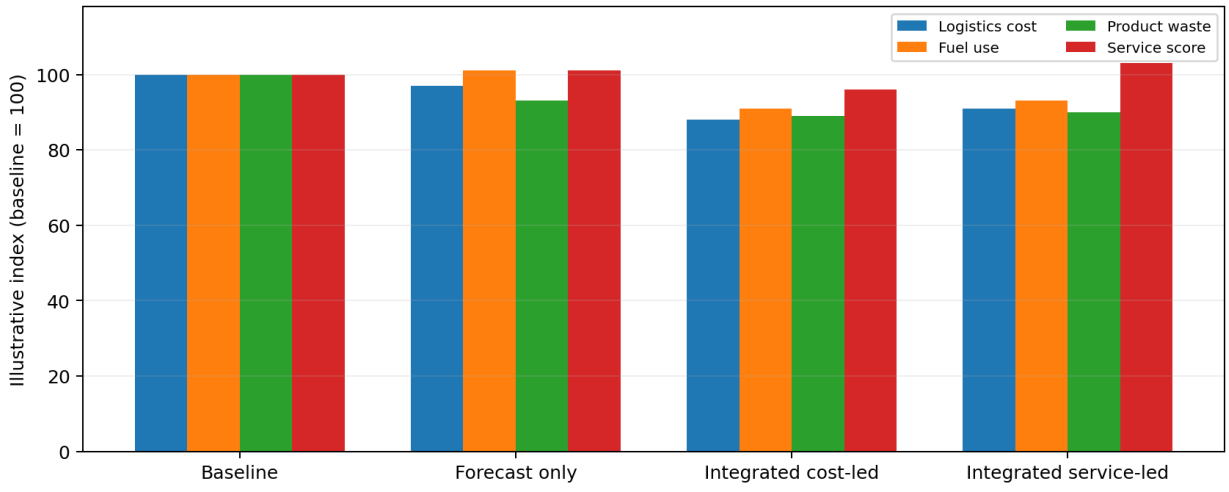


Figure 2. Hypothetical normalized planning trade-offs; values are illustrative only, not observed or experimentally validated results.

PEER REVIEW COPY

12. Implementation Roadmap for Saudi FMCG Organizations

A credible implementation starts with a diagnostic phase rather than immediately purchasing an advanced AI platform. The organization should map order-to-delivery processes, identify duplicated data sources, establish a reference definition for on-time-in-full service and audit the quality of location, timestamp and stock records. A steering group representing sales, supply chain, finance, transport, IT and frontline dispatch should approve baseline metrics. At this stage, simple process interventions—consistent customer identifiers, reliable proof of delivery and standardized promotion records—may yield benefits independently of AI and should not be misattributed to later algorithms.

The second phase should pilot a small network with identifiable comparators: for example, matched delivery clusters or sequential weekly treatment periods. Model development begins with transparent forecasting and heuristic baselines before moving to more complex techniques where justified. Candidate models are evaluated out of sample, and optimization is tested in simulation against actual route constraints. Dispatchers must be able to reject infeasible recommendations, and each override should be recorded. Commercial management should confirm that retailer relationships and promotion commitments remain represented in the planned service levels.

In the third phase, operational deployment should use a controlled release with fallback procedures. The system can initially recommend orders and routes for human approval rather than automatically dispatching trucks. Managers monitor the forecast bias, late-delivery reasons, exceptional stock transfers and data outages. A clear escalation path matters: transport supervisors may resolve a road disruption; commercial managers may reprioritize a national account; data stewards may correct mislabeled SKU records. Governance should be proportionate, avoiding bureaucratic approvals for routine recommendations while preserving traceability for decisions that materially affect customers or employees.

Scaling should follow demonstrated net value, not the number of models deployed. Evidence should cover both cost and service outcomes, include the expense of software integration and staff training, and evaluate whether improvements persist after peak-season effects. Regional transfer requires retraining and separate validation because the delivery ecology in Riyadh may differ from Jeddah, the Eastern Province or remote corridors. Long-term ownership of the platform should include documented model refresh cycles, cybersecurity provisions, vendor-exit arrangements, data-retention rules and a mechanism to investigate anomalous operational recommendations.

13. Managerial, Policy and Social Implications

For senior sales management, AI planning changes the discussion from a monthly target-versus-actual report to an accountable conversation about feasible demand, shelf availability and delivery economics. Sales incentives that reward gross shipments without considering returns or route cost can undermine sustainable logistics. Performance frameworks should therefore combine net sales quality, customer service and avoidable distribution burden. This requires cross-functional negotiation rather than imposing a logistics metric on commercial teams. A well-designed

UNDER PEER REVIEW IN IJAP

dashboard can surface uncertainty and show why a seemingly attractive promotion could require exceptional vehicle capacity.

For transport operators and distributors, the main opportunity lies in more predictable capacity use, fewer emergency trips and improved coordination across regional hubs. These benefits will be uneven if smaller carriers lack digital records or if warehouse practices remain manual. Policymakers and industry associations can support interoperability, digital capability development and evaluation standards without mandating a particular AI vendor. National Vision 2030 strategy documents provide direction for logistics connectivity and modernization, while individual firm-level initiatives should be assessed through real efficiency, worker-safety and service outcomes rather than claimed technological sophistication.

Social consequences require explicit treatment. Automated territory targets may influence sales staff workloads and income; live vehicle monitoring may create privacy and trust concerns; and route schedules must recognize heat, rest periods and safe unloading. Human oversight should be meaningful, with employees able to question obviously erroneous forecasts or dispatch instructions. Furthermore, optimizing purely for high-volume customers could weaken service to lower-density regions. A public-interest perspective may therefore require minimum coverage and reliability thresholds even when these are not the least-cost option.

14. Critical Discussion

The reviewed evidence suggests that AI has the clearest operational rationale when it improves an interconnected planning workflow. Technical research often reports predictive improvements, yet a forecast can be statistically superior without altering purchase orders or routes. Conversely, a modest predictive improvement may deliver substantial operational value if planners act on it and capacity constraints permit better consolidation. This creates an evaluation problem: academic papers frequently measure algorithm outputs, whereas businesses need reliable, auditable results on cost, service and sustainability. The appropriate unit of analysis is the decision process, not the model in isolation.

There is also a risk of conceptual inflation. Terms such as “AI logistics platform”, “intelligent distribution” and “sustainable optimization” may describe very different implementations. A basic dashboard that aggregates sales does not demonstrate machine learning, and a shortest-path map does not establish an integrated supply chain optimizer. Researchers should specify model inputs, training procedure, objective function, binding constraints, baseline comparator and data governance. Credible scholarship is strengthened by admitting where a proposed architecture remains conceptual, rather than attributing imagined achievements to a real Saudi company.

The sustainability claim needs similar discipline. Fuel intensity can improve because vehicles are more fully loaded, but absolute emissions may grow with network expansion. Improved shelf availability may reduce stockout-driven extra visits but increase total refrigeration or waste if quantities are excessive. Consequently, conclusions should differentiate efficiency, environmental impact and commercial growth. Neither predictive accuracy nor route length is a complete sustainability indicator; robust assessment spans logistics, inventory, service and worker well-

UNDER PEER REVIEW IN IJAR

being. The evidence most needed in Saudi Arabia is operational longitudinal data linking these measures.

15. Research Gaps and Future Research

First, the literature needs more field studies of integrated sales–distribution optimization in Saudi FMCG networks, where sales negotiations, distributor arrangements and warehouse practices interact with routing constraints. Research should compare genuine joint planning with forecasting-only and transport-only interventions. Second, datasets should distinguish lost sales from true demand and track promotional effects, stockouts and returns. Third, sustainability assessment should use measured fuel consumption and defensible waste records rather than simplified mileage proxies wherever feasible.

Fourth, researchers should investigate organizational adoption and human-AI coordination. It remains unclear how differences in sales incentives, trust, skill levels or cross-functional decision rights moderate AI impact. Fifth, models need out-of-distribution tests during sudden demand shocks, extreme weather, supply interruptions and regional market expansion. Sixth, transparent multi-objective optimization should be studied with equity constraints so small outlets and remote communities do not become residual priorities. Finally, future work should include implementation costs, carbon-accounting boundaries and vendor dependence when estimating long-run economic and environmental value.

A feasible future mixed-methods design could pair a twelve-month retrospective data audit with a prospective cluster-level pilot. Pre-registration would define primary outcomes such as on-time-in-full rate and verified fuel per case, plus guardrails for returns and driver safety. Comparable delivery territories would receive forecast-only or integrated AI-assisted planning, followed by interviews with sales managers, drivers and warehouse planners. Seasonal controls and blinded validation of operational records would make findings more credible than single-case before-and-after promotional metrics. Any use of proprietary enterprise data would require employer authorization and appropriate anonymization.

15.1. Proposed Evaluation Protocol and Data Dictionary

A rigorous pilot requires a defensible unit of assignment. Randomizing individual deliveries can create contamination because the same truck serves several outlets, and a dispatcher may respond to one optimized route by changing adjacent unoptimized routes. Cluster assignment by depot, stable delivery region or route group is therefore preferable when operationally feasible. The evaluation design should specify a stabilization period, exclude one-off commissioning disruptions from steady-state estimates and retain an intention-to-treat analysis of all assigned clusters. Statistical methods should adjust for baseline volume, outlet mix and promotional intensity, and confidence intervals should accompany estimated changes. Management should also maintain a log of simultaneous process changes, since replacing a vehicle fleet or opening a warehouse can alter emissions independently of the proposed AI system.

UNDER PEER REVIEW IN IJARC

A minimum data dictionary should define unique outlet ID, delivery point coordinates, SKU identifier, ordering channel, order timestamp, forecast issue time, customer delivery window, shipment quantity, documented stockout, promotion start and end, vehicle type, capacity, fuel reading, departed and arrived times, proof-of-delivery outcome, returns cause and vehicle kilometres. Records should be assessed for completeness, timestamp consistency, geography quality and duplicate handling before model training. Missingness is not always random: underserved outlets may have less reliable telemetry, making them vulnerable to systematically poorer model recommendations. Audits should therefore segment missing records by geographic territory, retailer category and carrier. Sensitive account information should be minimized, access controlled and retained according to applicable organizational policy.

Evaluation should separate prediction metrics from intervention outcomes. Weighted forecast error, bias and calibration test model behavior, whereas actual litres per delivered case, incremental truck trips, stock availability and returns test operational value. A forecast reduction with no operational change can be a legitimate technical result but should not be described as a logistics sustainability impact. Analysts should compare point estimates with measurement uncertainty, since fuel receipts and GPS tracks can disagree. Where the organization uses outsourced transport, incentives and reporting obligations must be agreed with the carrier in advance so that apparent savings are not achieved simply by transferring unmeasured cost or emissions to a subcontractor.

15.2. Transferability and Research Ethics

Transferability across Saudi regions should be demonstrated rather than presumed. Outlet density, local store receiving rules, delivery restrictions and distance between hubs can vary substantially. A model trained in a dense metropolitan market may show good aggregate accuracy but unacceptable service failures on long routes with limited alternatives. Rolling regional holdout tests and geographically stratified pilot reporting can help distinguish generalizable decision improvements from local fit. Researchers should also specify whether performance results apply to ambient grocery, chilled goods or mixed fleets. Refrigerated transport brings additional energy considerations, and reverse logistics adds vehicle handling complexity that cannot be represented by a simple outbound-distance metric.

Finally, research ethics should shape system design from the outset. Sales representatives, drivers and dispatchers should receive an explanation of which performance indicators are monitored, who can view them and how automated recommendations may affect schedules or incentives. A formal process should permit correction of false data and challenge of inappropriate algorithmic decisions. Employer permission is essential before publishing any internal operational records. If data are anonymized for academic publication, researchers must assess re-identification risk from rare delivery patterns and commercially sensitive customer combinations. These safeguards improve not only compliance but also the reliability of field evidence by making human participation and reporting more credible.

UNDER PEER REVIEW IN IJAR

16. Limitations

This is a structured critical narrative review supported by a conceptual framework. It does not present proprietary company data, original Saudi route observations, trained AI models or causal proof of business improvements. The illustrative chart is a hypothetical reasoning device and must not be cited as empirical evidence. Search coverage is not exhaustive, evidence sources vary in context and methodological rigor, and results from foreign retail markets may not transfer to Saudi customer practices. Government reports express strategic policy direction but do not verify organizational performance. Consequently, the proposed framework is a research agenda and decision aid rather than a validated operational product.

17. Conclusion

AI-driven sales and distribution planning offers a plausible pathway toward more sustainable Saudi FMCG logistics when it integrates demand forecasting, inventory allocation, feasible dispatch and accountable performance measurement. The strongest model is hybrid: machine learning anticipates demand and operating conditions, optimization translates those estimates into constrained decisions, and human managers preserve knowledge of customers and commercial commitments. Saudi Vision 2030 and the National Transport and Logistics Strategy provide a relevant development context, but performance must be established at the organizational level through verified records and rigorous pilots. Future research should measure service, logistics cost, emissions, returns and worker impact together. This evidence-led approach can help distinguish meaningful transformation from algorithmic branding and support practical distribution improvement in the Kingdom.

References

Brau, R.I., Aloysius, J., Siemsen, E., Waller, M.A. and related co-authors (2023). Demand planning for the digital supply chain: How to integrate human judgment and predictive analytics. Journal of Operations Management, 69(6), 965–982. https://doi.org/10.1002/joom.1257

Castro Moraes, P. et al. (2024). Hybrid convolutional long short-term memory models for sales forecasting in retail. Journal of Forecasting. https://doi.org/10.1002/for.3073

Application of artificial intelligence in demand planning for supply chains: a systematic literature review (2025). The International Journal of Logistics Management. https://doi.org/10.1108/IJLM-02-2024-0120

Demand Forecasting Using Machine Learning to Manage Product Inventory for Multi-channel Retailing Store (2023). IEEE International Conference on Omni-layer Intelligent Systems. https://doi.org/10.1109/COINS57856.2023.10189241

Ministry of Transport and Logistic Services (n.d.). National Transport and Logistics Strategy. https://mot.gov.sa/en/ntls

UNDER PEER REVIEW IN IJAR

Ministry of Transport and Logistic Services (n.d.). About the Ministry: Vision and Mission. https://api.mot.gov.sa/en/about-motls

Saudi Vision 2030 (2026). Vision 2030 Annual Report 2025. https://www.vision2030.gov.sa/media/ecdjfopq/vision2030_annual_report_2025_en.pdf

Saudi Vision 2030 (n.d.). Vision 2030 Overview. https://www.vision2030.gov.sa/en/overview

Saudi Press Agency (2026). NIDLP Reports SAR1,045 Billion Contribution to Non-Oil GDP in 2025. https://www.spa.gov.sa/en/N2615760

Saudi Press Agency (2025). NIDLP Drives 39% Growth in Non-Oil GDP to SAR986 Billion in 2024. https://spa.gov.sa/en/N2363341

Saudi Logistics Services (2025). Annual Report 2025: Market Overview. https://sal.sa/report2025

Saudi Logistics Services (2024). Annual Report 2024: Operating Environment. https://sal.sa/report2024/business-drivers/operating-environment

Saudi Industrial Development Fund (2023). Integrated Growth Report: Robust Logistics Networks. https://web.sidf.gov.sa/en/MediaCenter/ESGIntelligence/SIDF_Integrated%20Growth%20report.pdf

Kingdom of Saudi Arabia National Portal (n.d.). Transportation and Logistics Ecosystem. https://my.gov.sa/en/content/transport

UNDER PEER REVIEW IN IJAR