Research Article

The Marketing–Logistics Intelligence Loop: A Machine Learning Framework for Connecting Customer Acquisition with Fulfillment Efficiency

Raiyan Haider[a], Sudarshan Sundararajan[b]

[a]Graduate, University of Lancashire, Preston, UK. ORCID: 0009-0002-1153-9987. E-mail: raiyanhaider6@gmail.com
[b]ORCID: 0009-0005-0246-8621 E-mail: Sudarshan.Sundararajan@gmail.com

Abstract

This paper proposes a marketing–logistics intelligence loop that connects customer acquisition, demand sensing, inventory planning, order execution, and post-purchase feedback. Existing studies demonstrate strong results within individual functions: customer purchase prediction supports marketing and inventory decisions, marketing-mix variables improve demand-driven forecasting, and machine learning can improve supply-chain performance when forecasts inform operational decisions. Yet the evidence remains distributed across separate analytical problems. The proposed architecture treats campaign exposure, customer propensity, promotion response, point-of-sale activity, channel inventory, order arrivals, fulfillment status, delivery performance, returns, complaints, and satisfaction as connected observations in a recurrent decision system. The framework has four layers. Customer acquisition intelligence estimates segment membership, purchase propensity, promotion response, and longer-term value. Demand and order-arrival intelligence converts those signals into SKU-, channel-, location-, and time-level forecasts. Fulfillment intelligence translates forecasts into replenishment, supplier allocation, workforce, routing, and delivery decisions. Feedback intelligence uses stockouts, delays, returns, complaints, and satisfaction to update future predictions and decision policies. The design combines supervised learning, time-series models, fuzzy classification, deep learning, optimization, and causal evaluation. It evaluates performance using forecast error, order-fulfillment cycle time, service level, inventory cost, delivery cost, profitability, complaint incidence, and satisfaction rather than accuracy alone. The architecture therefore frames acquisition as an operational input and fulfillment as a source of marketing learning.

Keywords: Customer acquisition; marketing-mix analytics; demand forecasting; inventory optimization; order fulfillment; customer satisfaction.

1. Introduction

Digital commerce has increased the importance of coordinating customer-facing decisions with the operational systems that fulfill resulting orders. Machine-learning applications in marketing support customer behavior prediction and targeted resource allocation, while supply-chain analytics supports demand forecasting and operational decision-making. Research has also examined how promotional activity and other marketing-mix variables can inform demand-driven forecasts. These streams indicate that acquisition signals may have value beyond campaign evaluation, because they can inform the timing and scale of downstream demand.

However, marketing and logistics analytics are often treated as distinct problems, despite the potential value of aligning marketing information with logistics operations. Demand forecasts alone may not fully capture the operational consequences of customer-level behavior, channel inventory, and order timing; research on retailer replenishment, for example, identifies inventory balances as relevant to predicting orders from point-of-sale and order data. This paper develops a machine-learning framework that connects acquisition and promotion signals to demand sensing, inventory and fulfillment decisions, and post-purchase feedback. Its scope is an integrated decision loop that evaluates both commercial and operational outcomes, including profitability, service, fulfillment efficiency, and customer experience.

1.1 The Operational Cost of Separating Customer Acquisition, Demand Planning, and Delivery Execution

Customer acquisition and fulfillment execution are often measured by different teams, data structures, and time horizons. Marketing systems record impressions, clicks, audiences, prices, offers, and conversions, whereas planning systems commonly emphasize historical sales, replenishment orders, inventory balances, and lead times. This separation can cause a firm to treat demand as an exogenous forecast even when campaigns, customer trends, price

changes, and channel decisions are actively shaping it. Marketing machine-learning research identifies supervised, unsupervised, reinforcement, text, image, voice, and video applications across the seven marketing-mix dimensions, but its principal unit of analysis remains the customer or campaign [1]. Supply-chain analytics research similarly separates descriptive visibility, predictive forecasting, and prescriptive optimization across plan, source, make, deliver, and return activities [2].

The operational consequence is demand information distortion. Upstream actors may observe retailer or distribution-center orders rather than underlying consumer purchases. POS data improve demand-planning accuracy in one consumer-products setting, but do not necessarily improve order-fulfillment planning [3]. A separate study shows that combining POS and retailer-order histories can improve retailer-order forecasting by as much as 125% relative to approaches using either source alone, because channel inventory creates a balance between sales and replenishment orders [4]. These findings reject a universal data hierarchy. The useful signal depends on the decision endpoint, the echelon, and the inventory process.

The cost extends beyond inventory. A campaign that increases orders without informing warehouse labor, supplier capacity, or delivery planning can increase cycle time and complaints. Conversely, a fulfillment delay can change future purchase probability, return volume, and customer satisfaction. Satisfaction models identify delivery time and order accuracy as major predictors of future satisfaction, with a random-forest model achieving 92% accuracy in one dataset of more than 100,000 online orders [5]. Returns also create a reverse-flow planning problem: online return rates can exceed 30% in some settings, and return volume can be predicted at retailer, product, and period levels using operational and retailer information [6]. A disconnected architecture therefore loses feedback precisely where acquisition and retention decisions depend on service quality.

1.2 Research Objectives and Questions for an Acquisition-to-Fulfillment Intelligence Loop

The study has four objectives. First, it specifies a shared data model that aligns customer, campaign, product, channel, inventory, order, fulfillment, transport, return, complaint, and satisfaction records. Second, it maps predictive models to decisions rather than treating algorithmic accuracy as the sole outcome. Third, it proposes an evaluation design that separates forecast quality from operational value. Fourth, it identifies governance and validation requirements for deploying customer-linked machine learning across enterprise platforms.

The framework addresses four questions. RQ1: Which acquisition signals, customer attributes, marketing-mix variables, and transaction histories improve forecasts at SKU, channel, location, and time levels? RQ2: How should customer-level predictions be aggregated into order-arrival, inventory, capacity, and delivery decisions? RQ3: Which combinations of forecasting and optimization improve service, cost, profitability, and satisfaction relative to disconnected baselines? RQ4: How can post-purchase outcomes identify model drift, campaign harm, stockout effects, and fulfillment failures without introducing leakage or privacy violations?

The questions distinguish prediction from intervention. A purchase-propensity model estimates who may buy, but a promotion policy determines whom to contact; a demand forecast estimates future volume, but an inventory policy determines quantities and timing. Promotion targeting research illustrates this distinction by combining customer-trend estimation with nonlinear mixed-integer targeting and reporting an 11% average reduction in weighted mean absolute percentage error and profit increases of 3% to 11% in its application [7]. The proposed loop extends that logic downstream, asking whether the selected promotion can be fulfilled at acceptable cost and service quality.

1.3 Contributions, Paper Organization, and Evidence Scope

The paper makes three contributions. It presents a layered architecture for connecting acquisition signals to fulfillment decisions; it defines a cross-platform feature and entity model; and it proposes comparative experiments for predict-then-optimize and end-to-end learning. The architecture draws on primary empirical studies of purchase prediction, demand forecasting, inventory replenishment, supplier allocation, order arrival, delivery positions, returns, complaints, and satisfaction. It also uses systematic reviews to identify model families and gaps in the broader literature [8], [9].

The evidence does not justify a claim that one algorithm dominates across supply chains. Results vary by demand regularity, product life cycle, data granularity, lead-time uncertainty, channel structure, and decision objective. For example, gradient tree boosting achieved 89% accuracy and an AUC of 0.95 for next-month purchase prediction in a

dataset exceeding 10,000 customers and 200,000 purchases [10], whereas a deep end-to-end replenishment model performed better than conventional predict-then-optimize approaches in an e-commerce field experiment [11]. These outcomes are informative but not directly comparable. The remainder of the paper therefore treats reported metrics as context-specific evidence and emphasizes validation across operational endpoints.

2. Literature Review: Marketing Analytics, Demand Forecasting, and Supply Chain Performance

2.1 Machine Learning for Customer Acquisition, Purchase Prediction, Promotion Targeting, and Satisfaction

The marketing literature supports a progression from descriptive customer profiles to dynamic prediction and sequential intervention. In non-contractual settings, recency, purchase timing, and monetary-value features can be updated monthly to estimate whether a customer will purchase within a future period. Gradient tree boosting performed best in one study, with 89% accuracy and an AUC of 0.95 on held-out data [10]. This result supports customer-level demand signals, but purchase probability is not equivalent to SKU demand, order quantity, or fulfillment workload. Aggregation requires product, channel, geography, timing, and substitution information.

Promotion targeting research adds customer-to-customer trend effects to base purchase probability. The model separates price and seasonality from relational trends inferred from transactional data, then feeds the estimates into a promotion optimization problem [7]. Its reported improvement in weighted mean absolute percentage error and profit indicates that marketing analytics can affect demand prediction and economic decisions simultaneously. The result also raises a measurement issue: targeting policies change observed demand, so post-campaign sales cannot be interpreted as an unaltered demand signal. The acquisition layer must retain treatment assignment, exposure, eligibility, and control information.

Longer-horizon acquisition policies can be formulated through reinforcement learning. A deep recurrent Q-network with attention was applied to sequential targeting in digital banking and was reported to improve long-term rewards while identifying influential marketing channels [12]. Such methods fit a loop architecture because decisions affect later observations. They also require stricter evaluation than one-period classification. A model may increase immediate conversion while attracting customers whose orders are costly to fulfill or whose service experience produces complaints. The relevant reward should therefore combine contribution margin, fulfillment cost, retention, returns, and service outcomes.

Customer satisfaction provides an outcome bridge between logistics and marketing. A study using more than 100,000 online orders compared random forest, support-vector machines, and multilayer perceptrons; random forest predicted future satisfaction with 92% accuracy, and delivery time and order accuracy were identified as important drivers [5]. The model supports service-aware targeting, but satisfaction prediction remains observational unless changes in delivery performance are experimentally assigned. Complaint prediction offers a more operationally immediate target. Extreme gradient-boosted trees predicted infrequent complaints in semiconductor order fulfillment, and the analysis linked fulfillment quality with lower complaint tendency [13]. Threshold calibration is especially relevant because a complaint classifier may be used to prioritize scarce recovery resources.

2.2 Marketing-Mix Variables, Customer Information, and Demand-Driven Forecasting

Demand is shaped by marketing actions as well as historical sales. A fuzzy-classifier and back-propagation neural-network framework combined historical demand with advertising effectiveness, advertising expenditure, promotions, and marketing events [14]. Its central proposition is that manufacturers should sense and shape demand, rather than forecast an unchanging process. This distinction matters for acquisition-to-fulfillment coordination: campaign calendars, offer depth, media exposure, price, product, place, and customer response should enter forecasts with explicit time alignment.

Customer information extends the feature space beyond aggregate sales. Preferences, suggestions, post-consumption feedback, and customer histories have been used in demand forecasting, including intermittent-demand comparisons among extreme learning machines, K-nearest neighbors, decision trees, gradient boosting, and multilayer perceptrons [15]. Neural-network methods performed best in that application, while extreme learning machines were among the strongest predictors. These findings do not imply that customer data should always replace POS or time-series data.

Rather, they suggest a hierarchical design in which customer signals explain who may generate demand, marketing variables explain why demand changes, and operational data constrain what can be supplied.
Forecasting and supply-chain performance should be evaluated together. A hybrid ARIMAX and neural-network approach using time-series and explanatory variables produced statistically significant performance differences between traditional and machine-learning forecasting methods in a functional-product and steel-manufacturing context [16]. The study is useful because it links model choice to tangible supply-chain outcomes rather than reporting forecast error alone. Yet the abstract does not establish that the same model would perform across consumer promotions, slow-moving products, or highly seasonal fashion items.

Reviews of predictive supply-chain analytics classify applications into time-series forecasting, clustering, nearest-neighbor methods, neural networks, regression, support-vector methods, and support-vector regression [8]. Reviews of marketing machine learning similarly distinguish supervised, unsupervised, and reinforcement learning across marketing activities [1]. The common implication is architectural: different layers may require different models, and a single enterprise loop should not force one algorithm onto all endpoints.

2.3 Forecasting Evidence Across POS Data, Channel Inventories, Intermittent Demand, and New Products

The value of a data source depends on the forecast target. POS records approximate consumer withdrawals, but retailer orders reflect replenishment policies, inventory positions, batching, and lead times. POS improved demand-planning accuracy but not order-fulfillment planning in one large consumer-products supply chain study [3]. For slow-moving goods, a channel-inventory model using the last order, POS data, and expected lead-time demand improved order-timing prediction by 10 to 20 percentage points on average in a simulation and an empirical application involving a manufacturer, a central distribution center, and 53 hypermarkets [17]. The model also corrected shrinkage-induced bias. Thus, the loop should maintain distinct labels for consumer demand, retailer order timing, and supplier fulfillment demand.

Intermittent demand presents a second challenge. A Markov-combined method conditionally selected forecasting procedures according to demand and inventory states and outperformed single exponential smoothing, the Syntetos-Boylan Approximation, and Croston methods on two large Chinese e-commerce datasets [18]. Intermittency is particularly relevant to promotions and new products because zeros may represent low interest, unavailable inventory, delayed ordering, or an unobserved launch process. A model that treats all zeros as weak demand can systematically understate future orders.

New products lack long sales histories. The order-timing study indicates that POS-based channel-inventory nowcasting can help when order history is short or is a poor predictor of future demand [17]. New-product ordering research further incorporates stockouts into dynamic forecasting and optimization, reporting that a generalized cluster-while-regress method reduced prediction error by 6% to 9% relative to two clustering alternatives and increased average profit by 4.3% relative to current practice in a JD.com case [19]. Because that evidence concerns a specific platform and product setting, it supports a design principle rather than a universal performance claim: stockout exposure must be treated as censored demand.

The literature therefore favors a signal-specific forecasting layer. Customer and campaign variables explain prospective demand; POS and channel inventories support demand and order-timing inference; order histories describe replenishment behavior; and stockout flags identify missing sales. The architecture should preserve all these representations instead of collapsing them into one sales series.

2.4 Fulfillment, Inventory, Routing, Returns, and Customer-Service Outcomes

Fulfillment research increasingly links predictive models to prescriptive decisions. End-to-end deep learning can map historical features directly to replenishment quantities while learning the behavior of an optimal dynamic-programming solution under uncertain demand and vendor lead time. In a JD.com field experiment, the approach reduced holding cost, stockout cost, total inventory cost, and turnover rate relative to current practice [11]. A contrasting two-stage supplier study forecasts demand with a relational regressor chain and then solves a multi-objective supplier-selection and order-allocation problem; it reports that forecast-method choice changes supplier selection and allocated quantities [20]. These studies define the central comparison between modular predict-then-optimize and direct decision learning.

Service-capacity decisions also benefit from integrated models. A Seq2Seq architecture combining convolutional and long short-term memory components forecast logistics-service demand, quantified uncertainty through a dynamic distribution, and allocated third-party-forwarding capacity. It outperformed ARIMA and a two-stage particle-swarm-optimization and extreme-learning-machine benchmark in both point forecasting and capacity allocation [21]. Near-real-time order-arrival forecasting has likewise been addressed with an adaptive neuro-fuzzy methodology evaluated at three-hour intervals across three retailers [22]. These studies support a temporal bridge from campaign exposure to warehouse workload.

The execution layer includes workforce, order status, and routing. A five-working-day delivery-position model outperformed manual and automated forecasting approaches in one company and was used for workforce planning [23]. Deep-learning models for late-order classification achieved reported cross-validation accuracies as high as 100% for some architecture-classifier combinations, although the study's range across classifiers and architectures indicates that such values are setting dependent [24]. Last-mile research treats carrier booking as a stochastic packing problem and uses a machine-learning heuristic that performed better than progressive hedging in numerical experiments [25]. These outcomes concern different decision problems and should not be aggregated into one claim of general superiority.

Returns, complaints, and satisfaction complete the feedback chain. LASSO performed best for predicting future return volume in a study with more than 1,000 candidate interaction terms, partly because it selected informative interactions while remaining implementable [6]. Complaint prediction identifies customers at risk when confirmed fulfillment changes are likely to trigger dissatisfaction [13]. The evidence supports a closed-loop design in which service failures are labels for recovery and features for future acquisition decisions, rather than terminal records discarded after delivery.

3. Methods: Evidence Synthesis and Design of a Cross-Platform Machine Learning Framework

3.1 Corpus Mapping of Forecasting, Marketing, Inventory, Logistics, and Integration Studies

The evidence synthesis uses a function-by-decision matrix rather than a paper-by-paper narrative. Each study is coded by business function, prediction target, decision target, data granularity, model family, evaluation design, and reported limitation. The corpus includes primary studies of customer purchase, promotion targeting, demand, order timing, inventory, supplier allocation, service capacity, workforce, delivery, returns, complaints, and satisfaction. Review studies provide taxonomy and gap identification, including a review of 152 supply-chain forecasting papers spanning 1969 to 2023 [26], a review of 119 ML and DL demand-forecasting papers from 2015 to 2024 [9], and a review of 43 deep-learning supply-chain studies [27].

The synthesis distinguishes three evidence levels. Predictive evidence concerns out-of-sample classification or forecast accuracy. Decision evidence concerns inventory, allocation, routing, staffing, or profit after a model informs an action. Organizational evidence concerns integration capability, data management, and supply-chain outcomes. For example, the JD.com field experiment supplies decision evidence for end-to-end replenishment [11], while a survey of 520 mining executives links big-data analytics management capability with sustainable supply-chain outcomes through structural-equation modeling [28]. These evidence types cannot be treated as interchangeable.

The proposed synthesis records whether variables are observed before the prediction timestamp. This step is necessary because supply-chain datasets can contain leakage variables, such as post-delivery status or realized delay. A retail and logistics study reported that removing leakage variables reduced inflated performance and produced an AUC of 0.97 for delivery-delay classification and an R-squared of 0.99 for profit estimation [29]. The architecture therefore evaluates models under operational information availability, not retrospective completeness.

3.2 Data Model: CRM, Campaign, Transactional, POS, Inventory, OMS, WMS, TMS, and Delivery Records

The data model uses a shared event key linking customer, campaign, product, location, channel, order, inventory, shipment, and service outcome entities. CRM records contain customer identifier, consent state, segment, acquisition source, prior value, and contact history. Campaign records contain exposure time, channel, creative, offer, price, eligibility, treatment assignment, and response. Transactional records contain order lines, quantities, revenue, discount,

payment, cancellation, and customer linkage. POS records contain store, SKU, timestamp, sales, price, promotion, and availability.

Operational records describe the state through which demand passes. OMS data capture order creation, promised date, allocation, split, cancellation, and status changes. WMS data capture receipt, pick, pack, exception, labor, and dispatch events. Inventory records distinguish on-hand, allocated, in-transit, safety, available-to-promise, and phantom or unavailable stock. TMS data capture carrier, route, service level, dispatch, tracking, delivery attempt, and arrival. Returns, complaints, refunds, reviews, and satisfaction scores are linked to the order line and fulfillment events.

A data contract should define timestamp semantics, identifier resolution, update frequency, missingness, permissible joins, and ownership. Customer identifiers require pseudonymization and consent controls. Product identifiers require stable SKU and substitution mappings. Event time and processing time should be retained separately so that late-arriving records do not rewrite historical forecasts without audit. The architecture also stores model version, feature snapshot, prediction horizon, decision, intervention, and realized outcome.

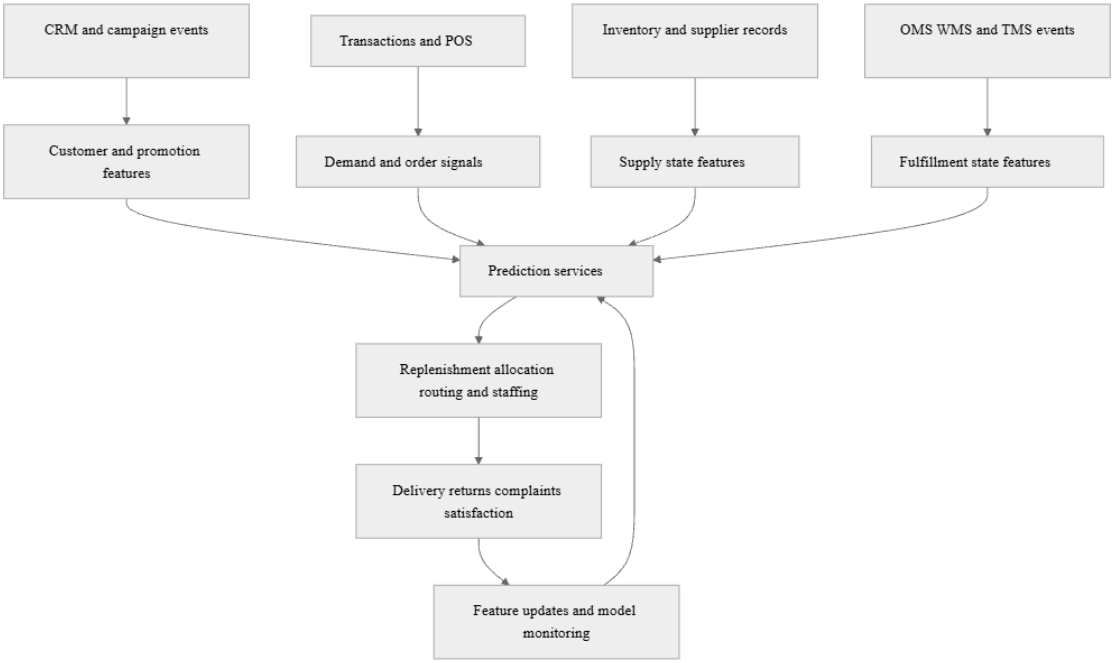

Fig. 1. Layered architecture of the marketing–logistics intelligence loop, from source events through prediction services and fulfillment decisions to outcome feedback.

The diagram expresses a relational rather than merely sequential architecture. Campaign data affect demand features; fulfillment events affect service predictions; and outcomes return to model monitoring. The design follows evidence that information-management integration supports visualization, automation, and intelligent coordination across supply-chain links [30], while cross- platform architecture studies emphasize shared feeds, forecasts, and operational triggers across enterprise systems [29].

3.3. Modeling Pipeline: Feature Engineering, Demand Prediction, Order Arrival Nowcasting, and Optimization

The modeling pipeline begins with time-aligned feature construction. Customer features include recency, frequency, monetary value, product affinity, prior response, channel engagement, and service history. Campaign features include exposure, offer depth, price, creative, audience, timing, and saturation. Product and market features include seasonality, lifecycle stage, substitutions, holidays when recorded, and channel availability. Operational features include on-hand inventory, stockout duration, lead time, open orders, supplier reliability, warehouse workload, carrier capacity, weather or traffic feeds where legally and operationally available, and current delivery status.

The first prediction task estimates customer purchase or response. Gradient boosting, calibrated logistic models, random forests, and sequential policies can be compared. The second task aggregates customer and campaign probabilities into SKU-channel-location forecasts, while preserving uncertainty and treatment status. The third task estimates near-real-time order arrivals at hourly or subdaily intervals. Adaptive neuro-fuzzy methods are appropriate comparators where arrival rates change rapidly [22]. For slow-moving products, channel-inventory nowcasting estimates the probability and timing of the next retailer order [17]. For intermittent demand, Croston-family, Markov-combined, tree, and neural models should be compared [18].

The prescriptive layer consumes predictive distributions rather than only point estimates. Replenishment optimization minimizes holding, stockout, ordering, and service penalties subject to capacity and lead-time constraints. Supplier allocation adds cost, quality, reliability, minimum quantities, and disruption risk. Routing assigns orders to facilities and carriers while respecting promised dates and capacity. Workforce planning converts predicted picks, packs, delivery positions, and exceptions into staffing decisions. End-to-end models may map features directly to decisions, as in the deep-learning replenishment framework [11], but modular models preserve interpretability and allow decision-specific constraints.

Model training uses rolling-origin validation, temporal holdouts, group splits by customer or SKU where required, and leakage audits. Calibration is evaluated for probabilities, prediction intervals for demand, and decision regret for prescriptions. Feature importance is interpreted with domain review rather than treated as causal evidence. For promotions, randomized holdouts or quasi-experimental designs estimate incremental response. For fulfillment, interventions should record the action that followed each prediction, because outcomes depend on the policy as well as the forecast.

3.4. Evaluation Design: Forecast Accuracy, Fulfillment Cycle Time, Service Levels, Cost, Profitability, and Satisfaction

Evaluation is organized around a scorecard with separate predictive and operational endpoints. Forecast metrics include mean absolute error, root mean squared error, weighted absolute percentage error where zeros permit it, bias, interval coverage, and order-timing accuracy. Classification metrics include AUC, precision, recall, calibration, and cost-weighted error. The selection of metrics must follow the decision. A rare complaint model may need recall at a feasible intervention budget, whereas an inventory forecast may require asymmetric penalties for underprediction.

Operational metrics include order-fulfillment cycle time, pick and pack duration, on-time shipment, on-time delivery, fill rate, stockout frequency, cancellation, return volume, complaint incidence, and workforce utilization. Financial metrics include holding cost, stockout cost, transport cost, promotion cost, contribution margin, and profit. Customer metrics include satisfaction, repeat purchase, recovery acceptance, and churn where measured. The literature supports this multidimensional approach because supply-chain performance includes integration, innovation, flexibility, efficiency, quality, and market performance, while customer satisfaction includes cost, flexibility, quality, and delivery [31].

The principal comparison is an incremental design: disconnected forecasting plus current planning versus an integrated loop with acquisition features and feedback. Secondary comparisons include marketing-only, logistics-only, integrated predict-then-optimize, and integrated end-to-end policies. Statistical analysis should report confidence intervals or bootstrap distributions where data permit, subgroup performance, and economic significance. Online experiments should randomize eligible customers or operational regions when feasible. Field validation is necessary because improvements in offline accuracy do not guarantee lower cost or better service.

4. Results: The Marketing–Logistics Intelligence Loop Architecture

4.1. Layer 1—Customer Acquisition Intelligence: Segmentation, Propensity, Promotion, and Long-Term Value Signals

The first layer converts customer and campaign events into decision-ready demand signals. Segmentation groups customers by observed purchase behavior, product affinity, channel, geography, service tolerance, and engagement. K-means clustering identified four customer segments in a retail and logistics dataset, but the reported segments should be treated as application-specific rather than universal [29]. Segmentation is useful only when it changes an action, such as audience selection, offer design, contact timing, service differentiation, or inventory reservation.

Propensity modeling estimates the probability of purchase, category adoption, response, cancellation, or return within a defined horizon. The non-contractual purchase framework demonstrates the value of dynamic purchase-timing and value features, with gradient tree boosting achieving an AUC of 0.95 in its test data [10]. The model output should be retained at customer-product-time level and then aggregated with campaign exposure, expected basket size, and available inventory. A high propensity without stock availability is not equivalent to fulfilled demand; it is a latent opportunity accompanied by a service risk.

Promotion intelligence estimates incremental response and expected contribution. Marketing-mix forecasting combines advertising effectiveness, expenditure, promotions, and events with historical demand [14]. Customer-trend targeting adds relational effects and optimizes multiple items and periods, reporting profit gains of 3% to 11% in its empirical application [7]. The loop extends the objective to include fulfillment cost, expected stockout, substitution, delivery promise, and return risk. A promotion should therefore be selected when its incremental value exceeds the acquisition, inventory, and service costs under the predicted supply state.

Long-term value signals include repeat purchase, margin, retention, complaint history, and service recovery. Sequential reinforcement learning can optimize long-term rewards rather than one-period conversions [12]. In the proposed architecture, its reward function is constrained by fulfillment feasibility and fairness rules. The model should record why an audience was selected, what offer was delivered, and which operational conditions were known at decision time. This record permits later separation of targeting quality from fulfillment quality.

4.2.Layer 2—Demand and Order-Arrival Intelligence: From Campaign Exposure to SKU-, Channel-, and Time-Level Forecasts

The second layer translates customer-level signals into operational demand. The aggregation is hierarchical: customer-product probabilities become expected baskets, baskets become SKU-channel demand, channel demand becomes location and time forecasts, and those forecasts become order arrivals at distribution centers or suppliers. Each level retains uncertainty, because aggregation can hide concentration in a few customers or campaigns. Campaign exposure should enter the feature window before demand realization, while post-order status and final delivery should not.

The layer maintains separate targets for sell-through, retailer orders, distribution-center orders, and warehouse arrivals. POS data can improve demand planning but may not improve fulfillment planning [3]. Combining POS and retailer orders can better predict supplier orders because inventory balance links the two signals [4]. For slow-moving products, a dynamic channel-inventory estimate supports nowcasting of the next order and improved order-timing accuracy [17]. The architecture therefore uses a multi-target data model rather than treating all observed transactions as identical demand.

Order-arrival nowcasting updates at the operational cadence. The near-real-time e-commerce methodology forecasts individual retailer arrivals at three-hour intervals and emphasizes retailer-specific prediction for daily decisions [22]. Campaign spikes, flash promotions, and social-commerce events should produce temporary features with decay, saturation, and uncertainty. A forecast service publishes quantiles or scenarios to OMS, WMS, and TMS systems, allowing capacity decisions before confirmed orders arrive. The service should also report signal provenance, such as campaign exposure, POS movement, channel inventory, or order history.

New products and stockouts require special treatment. When a product has limited order history, POS-based channel inventory and customer-level affinity can supplement historical orders [17]. When stockouts suppress sales, observed demand is censored. New-product research that explicitly incorporates shortages into forecasting and ordering reports lower prediction error and higher profit than comparison methods in a JD.com case [19]. The layer should therefore include availability-adjusted demand, lost-sales indicators, substitution, and replenishment constraints.

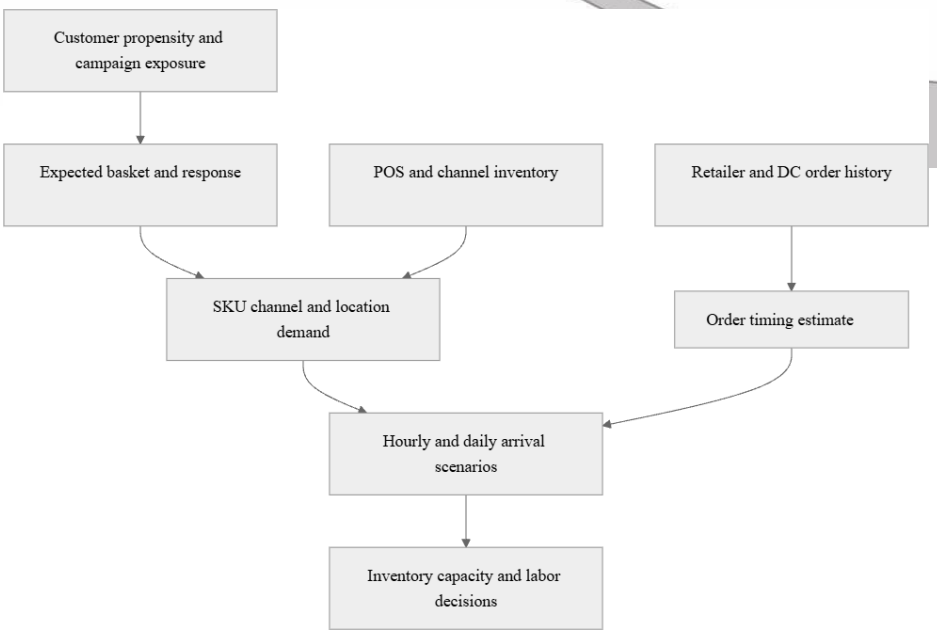

Fig. 2. Demand and order-arrival intelligence flow, from campaign exposure and channel signals to hourly and daily arrival scenarios and capacity decisions.

This flow distinguishes demand generation from order timing. The distinction is supported by evidence that POS, retailer orders, and distribution-center orders have different forecasting value depending on the planning endpoint [3].

4.3. Layer 3—Fulfillment Intelligence: Inventory Replenishment, Supplier Allocation, Order Routing, Workforce Planning, and Delivery

The third layer converts forecasts into coordinated operational decisions. Inventory replenishment uses demand scenarios, vendor lead time, current inventory, open orders, service targets, and cost. End-to-end deep learning directly predicts replenishment quantities and learned the behavior of a dynamic-programming solution without assuming demand or lead-time distributions [11]. A modular alternative forecasts demand and then optimizes orders. The architecture supports both, with feasibility constraints and human override for exceptional products.

Supplier allocation uses forecast demand alongside supplier price, capacity, quality, lead time, reliability, and disruption risk. A relational regressor chain followed by multi-objective programming demonstrates that forecast selection can change supplier selection and allocated quantities [20]. Supplier-disruption prediction can use engineered agility features, with one OEM case reporting 80% accuracy for late-order prediction [32]. These models should not be used to exclude suppliers solely on a predictive score; they should generate scenarios and escalation triggers within procurement governance.

Order routing assigns each order to a facility, inventory pool, carrier, or service option. Decisions should consider customer promise, distance, available-to-promise stock, margin, split-order cost, and delivery risk. CRM-derived preferences may inform promised service or facility selection, but sensitive attributes should not determine service quality without governance. Third-party logistics routing can be formulated as stochastic packing with machine-learning heuristics that produce good solutions within short computational times [25]. Multi-channel fulfillment profitability also depends on the cost and margin of different fulfillment methods, as shown by a nonlinear mixed-integer model using empirical retailer data [33].

Workforce planning converts predicted positions, picks, packs, exceptions, and delivery workload into staffing. Delivery-position forecasting over a five-working-day horizon outperformed manual and automated alternatives in one company and informed optimized workforce planning [23]. Execution monitoring classifies late orders, updates promised dates, and detects exceptions. Deep architectures can represent temporal dependencies in complex supply chains, but cross-validation accuracy must be tested against leakage and operational baselines [24]. For on-demand grocery, process-informed deep learning models represent warehouse workload, dispatch capacity, courier interactions, and sequential fulfillment stages when predicting cycle time [34].

4.4. Layer 4—Feedback Intelligence: Returns, Complaints, Stockouts, Delays, and Satisfaction as Learning Signals

The fourth layer treats operational outcomes as future features and decision labels. Returns identify product, customer, channel, fit, quality, and delivery patterns. LASSO predicted return volume at retailer, product-type, and period levels and remained consistent across robustness tests [6]. Return forecasts should influence acquisition targeting, product recommendations, reverse-logistics capacity, and inventory availability. A campaign that produces high gross sales but high returns may have lower contribution than its acquisition dashboard suggests.

Complaints are sparse but operationally actionable. In semiconductor fulfillment, extreme gradient-boosted trees predicted whether fulfillment changes would trigger complaints, and threshold tuning

was identified as a practical calibration mechanism [13]. The loop can prioritize proactive notification, alternative allocation, expedited delivery, or account-manager intervention. The intervention and its outcome must be recorded so that complaint prevention is not confused with complaint suppression.

Stockouts and delays are both service failures and measurement events. Stockouts reduce observed sales and may create substitution, cancellation, or delayed demand. Delivery delays alter satisfaction and future purchase behavior. The architecture therefore creates event labels for stockout onset, lost sale, substitution, promise breach, late delivery, return, complaint, and recovery. These labels enter future propensity and demand models only after their temporal position is respected.

Satisfaction closes the customer loop. Delivery time and order accuracy predict satisfaction in e-commerce data [5], while broader supply-chain research links integration, flexibility, efficiency, quality, and delivery with customer satisfaction [31]. Satisfaction should be segmented by customer, product, route, facility, and campaign, but not used as a simple proxy for all service quality. A high score may reflect low expectations, and a low score may reflect product attributes outside logistics control. The feedback layer therefore combines ratings with objective delivery, return, complaint, and repeat-purchase records.

4.5. Proposed Data Designs: System Architecture Diagram, Process Flowchart, Entity-Relationship Model, Feature Matrix, and KPI Dashboard

The proposed system architecture has an event-ingestion layer, a governed analytical store, feature services, model services, optimization services, execution interfaces, and monitoring. CRM, campaign, POS, OMS, WMS, TMS, ERP, carrier, and customer-service data enter through batch or streaming pipelines. A canonical event model resolves identifiers and timestamps. The feature store creates point-in-time snapshots for training and real-time scoring. Model services publish predictions with uncertainty, explanation, version, and expiration time. Optimization services return feasible decisions, constraints, and escalation states.



Fig. 3. System architecture, from event ingestion and the governed event store through feature, prediction, and optimization services to operational outcomes.

The entity-relationship design uses customer and order-line keys as central links, with product, campaign, location, inventory, supplier, shipment, return, complaint, and satisfaction entities attached through valid temporal relationships. A campaign exposure links to a customer and creative; a transaction links to an order and product; an order line links to inventory allocation and shipment events; a shipment links to delivery outcome; and a return or complaint links back to the order line and service history. The model must permit one order to contain multiple products, one order to split across shipments, and one customer to have multiple campaigns and outcomes.

Table 1. Feature families, examples, decision endpoints, and leakage controls.

Feature family	Examples	Decision endpoint	Leakage control
Customer	recency, frequency, value, affinity, prior service	acquisition and retention	snapshot before decision
Campaign	exposure, offer, price, channel, treatment	response and demand	exclude post-response variables
Demand	POS, orders, seasonality, substitutions	SKU and channel forecast	distinguish sell-through from orders
Inventory	on-hand, allocated, in-transit, stockout	replenishment and routing	point-in-time availability
Fulfillment	workload, picks, pack time, carrier state	cycle time and staffing	exclude final outcome
Feedback	returns, complaints, ratings, delays	recovery and model update	timestamp outcome after service

The KPI dashboard should display forecast bias and error by hierarchy; purchase and promotion lift; arrival-nowcast error; fill rate; stockout and cancellation; cycle time; on-time delivery; holding, stockout, and transport cost; contribution margin; return and complaint rate; satisfaction; and subgroup fairness. Dashboard users should see both aggregate and exception views. The design follows the principle that forecasting metrics should be selected with supply-chain KPIs and post-process operational outcomes [26], rather than allowing a single accuracy statistic to govern the loop.

5 Results: Predictive and Prescriptive Decision Modules Across the Fulfillment Network

5.1 Demand Forecasting Module: Comparative ML, Deep Learning, Fuzzy Classification, and Intermittent-Demand Models

The demand module compares model families by data condition and decision horizon. Gradient boosting is a strong baseline for customer purchase prediction and structured tabular demand because it handles nonlinear interactions and heterogeneous features [10]. ARIMAX provides a transparent time-series comparator when explanatory variables such as price, promotion, and economic conditions are available. Neural networks and fuzzy-neural systems represent nonlinear marketing effects, while convolutional and recurrent architectures capture temporal or cross-series structure. The fuzzy back-propagation framework combining demand, advertising, expenditure, promotions, and events outperformed benchmark forecasting methods in its experiments [14].

Deep learning should be assigned where sequence length, cross-series dependence, or high-dimensional data justify its cost. A CNN-LSTM model integrated demand forecasting with logistics-service capacity allocation and outperformed ARIMA and a two-stage optimization benchmark [21]. A CNN sales model reported a correlation of 0.98 and a highest mean error of 1.78% in its e-commerce application [35]. These figures are endpoint-specific and should not be compared directly with AUC, inventory cost, or order-timing accuracy. A model comparison must preserve the unit, horizon, dataset, and target.

Fuzzy classification is useful when demand states, inventory conditions, or managerial categories are uncertain. Adaptive neuro-fuzzy inference predicted near-real-time order arrivals at three-hour intervals across three retailers [22]. Intermittent-demand methods require a separate track. The Markov-combined method used demand and inventory states and outperformed SES, SBA, and Croston on two large e-commerce datasets [18]. Customer-information models show that neural and extreme-learning methods can improve intermittent-demand prediction, although computation time and statistical testing remain relevant [15].

Table 2. Decision modules, primary targets, comparators, and reported evidence.

Module	Primary target	Main comparator	Reported evidence
Customer propensity	next-period purchase	gradient boosting and other ML	89% accuracy; AUC 0.95 [10]
Promotion demand	SKU and period demand	benchmark forecasts	fuzzy neural method outperformed benchmarks [14]
Near-real-time arrivals	three-hour order arrivals	time-series and neuro-fuzzy	high accuracy across three retailers [22]
Intermittent demand	demand with many zeros	SES, SBA, Croston	MCM outperformed classical comparators [18]
Sales forecasting	e-commerce consumption	data mining and CNN	correlation 0.98; highest mean error 1.78% [35]

The module reports forecasts at customer, SKU, channel, location, and network levels. Hierarchical reconciliation prevents marketing forecasts from exceeding inventory or capacity assumptions without an explicit scenario. The model registry stores separate performance for normal, promotion, stockout, launch, and recovery periods. This structure is preferable to selecting one global champion because the literature reports different strengths for structured ML, fuzzy systems, deep learning, and intermittent-demand procedures

5.2. Inventory and Allocation Module: End-to-End Replenishment, Supplier Selection, and Stockout-Aware Ordering

The inventory module translates demand distributions into replenishment and allocation decisions. The predict-then-optimize design first estimates demand and lead time, then minimizes expected holding, stockout, procurement, transport, and service costs. Its advantage is transparency: planners can inspect forecast drivers and optimization constraints. Supplier-selection research demonstrates this integration by feeding a relational regressor-chain forecast into multi-objective programming and showing that forecast choice changes supplier and quantity decisions [20].

The end-to-end design maps input features directly to replenishment quantities. The deep-learning inventory study trained on historical observations to approximate an optimal dynamic-programming policy without prespecified demand or lead-time distributions. Its JD.com field experiment reported lower holding, stockout, total inventory, and turnover costs than current practice [11]. End-to-end learning is attractive when decisions are frequent and the action space is stable. It is less transparent when constraints change, when the training policy embeds historical inefficiencies, or when a new supplier or product has no analog in the data.

Stockout-aware ordering is necessary for acquisition-linked planning. Campaign exposure can raise latent demand while inventory limits observed sales. New-product research explicitly models shortages and reports a 6% to 9% reduction in prediction error and a 4.3% increase in average profit over current practice in its case study [19]. The module therefore separates observed sales, estimated lost sales, substitution, and backorders. It also allocates inventory across channels according to margin, promise, substitution, customer priority rules, and fairness constraints.

Supplier risk and quality enter as scenario variables. Predictive disruption models can identify late orders using engineered supply-chain features, including agility, with one case reporting 80% accuracy [32]. The system should use that probability to adjust safety stock, expedite decisions, or alternative sourcing, not to make an irreversible supplier judgment. Dual-channel studies further show that data-driven marketing can increase inventory in both channels while not always increasing total expected profit [35]. This finding cautions against treating demand growth as operational success without evaluating margin, returns, and coordination.

5.3. Execution Module: Fulfillment-Cycle-Time Prediction, Order Status Detection, Delivery Positions, and Last-Mile Routing

The execution module begins when an order is created and ends when delivery, cancellation, or return is recorded. Fulfillment-cycle-time prediction estimates elapsed time from order creation through picking, packing, dispatch, and delivery. Process-informed deep learning represents workload, processing capacity, delivery capacity, dispatching, and courier interactions in on-demand grocery operations [34]. The prediction should be stage-specific so that a delay in picking is not attributed to the carrier. Stage predictions support dynamic promise updates and targeted intervention.

Order-status detection classifies whether an order is likely to be late or at risk. Deep architectures, including LSTM, one-dimensional CNN, and temporal convolutional designs, were evaluated with several classifiers and reported high cross-validation accuracy in a complex supply chain [24]. The module requires strict temporal splits and leakage tests because final status fields can make retrospective classification appear more accurate than real-time prediction. A risk score should be accompanied by calibrated probability, expected delay, cause category, and recommended action.

Delivery-position forecasting predicts the number of positions or deliveries required over a short horizon. A practical model with a five-working-day horizon outperformed manual and automated alternatives and informed workforce planning [23]. The output connects demand forecasts with labor scheduling and vehicle preparation. It should be recalculated after major campaign changes, order-arrival deviations, carrier disruptions, and warehouse exceptions.

Last-mile routing combines order priority, promised time, vehicle or carrier capacity, distance, and stochastic parcel size. Machine-learning heuristics for third-party delivery addressed a variable-cost and size-bin-packing problem and performed better than progressive hedging in numerical experiments [25]. Multi-channel fulfillment profitability models show why routing cannot be optimized solely for distance: fulfillment method, margin, and fixed or variable cost affect online profitability [33]. The execution module therefore reports service and contribution outcomes together.

5.4. Decision Comparison Designs: Baseline Versus Integrated Loop, Predict-Then-Optimize Versus End-to-End Learning

The primary experiment compares a disconnected baseline with an integrated loop. The baseline uses historical sales for demand, current replenishment rules, manual campaign planning, static workforce assumptions, and reactive exception management. The integrated condition adds campaign exposure, customer propensity, channel inventory, stockout correction, order-arrival nowcasting, service-risk scores, and feedback outcomes. Both conditions must receive the same inventory, supplier, and carrier constraints. The comparison evaluates forecast quality, cycle time, service level, total cost, profit, complaint rate, return rate, and satisfaction.

The second experiment compares predict-then-optimize with end-to-end learning. Predict-then-optimize produces a forecast distribution and passes it to an explicit optimization model. End-to-end learning maps features to actions by imitating or approximating an optimal policy. The former supports explanations and constraint changes; the latter may reduce decision error when the operational objective is difficult to express and historical data are rich. The JD.com evidence favors end-to-end replenishment in its field setting [11], whereas supplier-selection evidence demonstrates the practical interpretability of sequential forecasting and optimization [20].

A third design tests information ablation. Models are trained with customer and campaign variables, POS and channel inventory, order histories, fulfillment state, and feedback variables in successive combinations. This design identifies whether acquisition signals add value after operational data are included. It also tests the finding that POS may improve demand planning but not fulfillment planning

[3]. A fourth design tests robustness across normal periods, promotions, launches, intermittent demand, stockouts, and disruption periods. Statistical inference should report uncertainty across time windows and business units rather than a single pooled metric.

The decision comparison must avoid rewarding a model for selecting easier orders or withholding demand. Constraints should enforce campaign eligibility, inventory conservation, promise feasibility, labor limits, and service obligations. Results should report both realized outcomes and counterfactual opportunity. For example, a lower stockout rate may arise from carrying more inventory, reducing promotions, improving forecasts, or reallocating scarce units. The final comparison therefore uses cost, profit, service, customer outcomes, and policy compliance.

6. Discussion: Integration Benefits, Tradeoffs, and Evidence-Based Design Choices

6.1. How Marketing Signals Reduce Demand Information Distortion and Improve Fulfillment Responsiveness

Marketing signals can reduce uncertainty before orders appear in operational systems. Campaign calendars, audience counts, exposure, offer depth, price, and customer propensity provide leading indicators that POS and replenishment histories reveal only after demand has materialized. Marketing-mix forecasting explicitly combines advertising effectiveness, expenditure, promotions, events, and historical demand [14]. Customer purchase prediction adds dynamic recency, frequency, and value features [10]. Together, these signals can indicate when demand will change, which products will be affected, and which channels are likely to receive the response.

The value is greatest when the operational target is correctly defined. POS captures consumer withdrawal, retailer orders encode replenishment policy, and distribution-center orders encode another echelon's inventory decision. Evidence that POS improves demand planning but not order-fulfillment planning [3] shows why acquisition data should not simply replace order data. Evidence that POS and retailer orders together can improve supplier-order forecasting by up to 125% [4] supports a fusion model that respects inventory balance. Marketing signals should therefore be treated as one layer in a multi-echelon signal system.

The loop also improves responsiveness by connecting forecasts to capacity. Near-real-time order-arrival prediction supports hour-to-hour logistics decisions [22]. Campaign exposure can be translated into expected picks, packs, delivery positions, and carrier demand before all orders are confirmed. This anticipatory capacity response may reduce late deliveries and customer complaints, but the evidence base has not yet established a general causal effect from campaign-logistics integration to these outcomes. The proposed architecture makes that effect testable.

6.2. The Central Debate: Forecast Accuracy Versus Operational and Customer-Level Value

Forecast accuracy is necessary but insufficient. A lower mean absolute error may have little value if errors occur in low-margin products, while a modest forecast improvement in a constrained SKU can prevent stockouts and protect customers. Inventory research demonstrates this distinction by evaluating holding, stockout, total inventory, and turnover costs after an end-to-end policy [11]. Supplier-allocation research similarly shows that forecast selection changes suppliers and quantities, meaning that forecast error propagates into procurement decisions [20].

Customer outcomes add another layer. Satisfaction prediction identifies delivery time and order accuracy as important variables [5]. Complaint prediction shows that fulfillment changes can trigger infrequent but consequential customer responses [13]. A campaign with strong conversion but poor delivery may reduce repeat purchase or increase returns. Conversely, allocating scarce stock to customers with high predicted value may improve short-term margin while creating fairness or channel-conflict concerns. The objective should therefore be a constrained value function combining margin, service, cost, and customer outcomes.

The literature also contains evidence that stronger demand stimulation does not guarantee higher supply-chain profit. In dual-channel settings, improved data-driven marketing can increase inventory in both channels while total expected profit does not always increase [35]. This result supports a portfolio of KPIs rather than a single acquisition or forecast target. Model selection should be based on downstream regret, stockout cost, service levels, contribution, and satisfaction, with forecast metrics used to diagnose rather than dictate decisions.

6.3. Tradeoffs Among Model Complexity, Interpretability, Real-Time Responsiveness, and Implementation Cost

Model complexity creates capacity as well as burden. Deep learning can represent temporal dependencies and integrate multiple operational stages, as shown in logistics-service capacity and fulfillment-cycle-time studies [21], [34]. End-to-end replenishment can shorten the decision process and avoid explicit distributional assumptions [11]. These advantages matter when decisions are frequent, data are abundant, and the operational policy is stable.

Simpler models remain valuable. LASSO predicted return volume accurately while selecting interactions and remaining easy to implement [6]. Tree ensembles perform well on structured customer and satisfaction data [5], [10]. Markov-combined intermittent-demand forecasting improves on classical methods while remaining computationally accessible [18]. A modular architecture can use simple models for stable, interpretable endpoints and reserve deep learning for high-volume sequential problems.

Interpretability is operational rather than cosmetic. Planners need to know whether a forecast changed because of a promotion, stockout, channel inventory, supplier delay, or data defect. Procurement teams need reasons for supplier-risk alerts; customer-service teams need reasons for complaint prioritization. Explainable AI was explicitly used to interpret an integrated inventory prediction model in an ERP setting, where LightGBM achieved an R-squared of 0.90 and a mean absolute error of 39.25 in the reported application [36]. These values are context-specific, but the design illustrates how explanation can accompany prediction.

Real-time responsiveness also has costs. Streaming integration, feature computation, model monitoring, and optimization increase infrastructure and governance requirements. A batch forecast may be adequate for slow-moving products, while a three-hour arrival forecast is appropriate for e-commerce operations [22]. The architecture should therefore assign latency according to decision value, not impose real-time processing everywhere. Human review remains appropriate for low-frequency, high-consequence decisions and for model states outside training support.

6.4. Cross-Platform Integration Requirements for CRM, Marketing Automation, OMS, WMS, TMS, and Analytics Systems

Integration requires shared identifiers, event-time semantics, service interfaces, and ownership. CRM and marketing automation must expose campaign eligibility, treatment, exposure, response, and suppression rules. OMS must receive forecast scenarios, allocation recommendations, promise-risk scores, and exception states. WMS must expose inventory, workload, pick, pack, and dispatch events. TMS must expose carrier capacity, route status, tracking, delivery attempt, and exception data. The analytics layer must preserve point-in-time feature snapshots and model versions.

The technical architecture should support both batch and event-driven exchange. A common event schema reduces reconciliation between customer, order, inventory, and shipment records. Cross-platform integration studies propose unified architectures connecting heterogeneous feeds to shared forecasts and prescriptive triggers across OMS, WMS, and TMS [29]. ERP integration research similarly emphasizes shared data across stakeholders for material forecasting and inventory sufficiency [36]. These proposals require operational testing because conceptual integration does not establish data quality or realized savings.

Organizational alignment is equally necessary. Research on logistics-marketing and logistics-production integration finds that performance effects depend on internal and external integration, and that logistics-marketing integration alone did not consistently reduce costs, stockouts, or lead times [37]. This finding cautions against assuming that an API or data lake will resolve conflicting incentives. Marketing may optimize conversion, logistics may optimize utilization, and customer service may optimize complaint closure. The loop needs shared KPIs, decision rights, escalation rules, and post-decision accountability.

A practical deployment sequence begins with identity and event governance, then demand-signal fusion, then inventory and capacity decisions, and finally adaptive acquisition and feedback policies. Each stage should retain a fallback process and measure incremental value. Integration is successful when information changes a feasible decision and improves a defined operational or customer endpoint, not when more systems are connected.

7. Discussion: Governance, Generalizability, and Research Gaps in the Marketing–Logistics Loop

7.1. Data Quality, Privacy, Explainability, and Bias in Customer- and Operations-Level Machine Learning

Data governance must be treated as an analytical condition rather than an administrative afterthought. Acquisition models combine identifiers, exposures, transactions, prices, and responses, whereas logistics models add inventory states, fulfillment events, carrier records, and returns. Misaligned timestamps, duplicated orders, missing channel events, and post-outcome variables can create leakage and distort operational predictions. A DataCo study illustrates the problem: removing leakage variables reduced inflated performance claims while preserving strong delivery-delay classification and profit estimation results [38]. Point-of-sale data also improved demand-planning accuracy without improving order-fulfillment planning in every context, showing that data granularity does not guarantee decision value [3].

Privacy and consent constrain the use of customer-level signals. Promotion-targeting research shows that transactional data can support customer-trend estimation when social data are unavailable because of cost and privacy constraints [7]. A deployable loop should therefore minimize collected attributes, separate personally identifiable information from analytical keys, record consent and suppression status, and restrict features to the decision horizon. These controls must cover model training, campaign activation, order routing, and feedback ingestion because a customer may consent to personalization without consenting to unrestricted operational profiling.

Explainability is required when a prediction changes price exposure, inventory priority, delivery promise, or complaint handling. A supply-chain integration study used explainable artificial intelligence to interpret inventory predictions, with LightGBM achieving an R^2 of 0.90 and a mean absolute error of 39.25 in its case setting [36]. Explanation methods should identify influential features, uncertainty, data vintage, and applicable operating conditions rather than present a generic feature ranking. Human review is especially appropriate for rare complaints, supplier disruptions, and out-of-distribution demand, where automated confidence may be poorly calibrated.

Bias can enter through stockouts, selective campaign exposure, unequal delivery access, and feedback generated only by customers who complain. Stockouts censor observed demand, and recent work on new products explicitly models shortage effects because sales records understate unconstrained demand [19]. Complaint prediction also concerns an infrequent outcome, requiring threshold calibration before operational use [13]. Governance should therefore include subgroup error checks, missingness audits, counterfactual demand adjustments, drift monitoring, and appeal procedures. The following control path connects these safeguards to decisions.

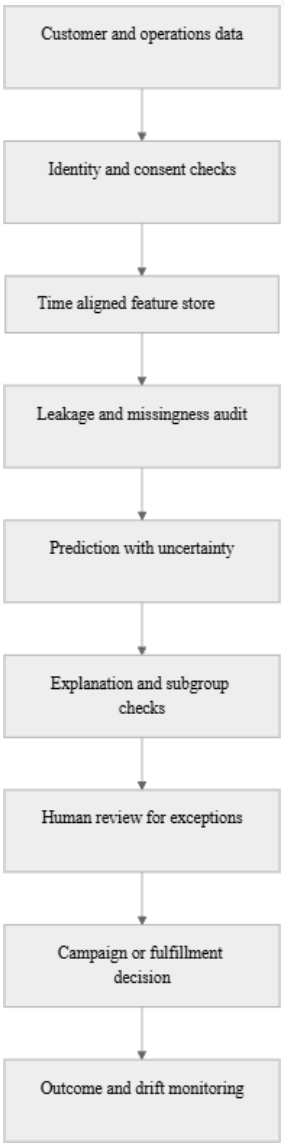

Fig. 4. Governance control path, from identity and consent checks through leakage audit, prediction, explanation, and human review to decision and drift monitoring.

7.2. Evidence Gaps in Causal Measurement of Acquisition Decisions on Inventory, Delivery, and Profitability

The main evidence gap concerns causal linkage. Purchase prediction can be accurate without showing that an acquisition intervention caused additional demand, and demand prediction can improve without proving that inventory or delivery outcomes improved. Customer-trend targeting offers stronger evidence than ordinary prediction: its authors combine instrumental-variable estimation with optimization and report an 11% reduction in weighted mean absolute percentage error and a 3% to 11% profit increase in a retail application [7]. Yet the result concerns promotion policy and profit, not the full downstream effects on stockouts, replenishment, fulfillment time, or returns.

Most integration evidence remains observational, simulated, or case specific. A field experiment at JD.com reported lower holding, stockout, and total inventory costs for an end-to-end replenishment model relative to current practice [11], while a Spanish FMCG study using structural equation modeling found that logistics-marketing integration did not consistently reduce costs, stockouts, or lead times [37]. These findings are not directly contradictory: one evaluates an operational algorithm in a field setting, whereas the other estimates organizational relationships across firms. Neither identifies how a particular acquisition treatment changes

downstream capacity, service promises, customer satisfaction, and contribution margin.

Future studies should define a causal estimand before model selection. Candidate estimands include the incremental orders caused by a campaign under capacity constraints, the change in stockout probability per acquired customer, the effect of promised delivery speed on conversion and return behavior, and net profit after media, inventory, transport, and service costs. Randomized campaign holdouts should be linked to SKU-location-period demand and fulfillment records. Mediation analysis can then distinguish direct conversion effects from indirect effects operating through inventory availability and delivery reliability.

7.3. Limitations of the Current Corpus: Fragmented Domains, Uneven Empirical Validation, and Limited Longitudinal Integration

The corpus is divided among marketing prediction, demand forecasting, inventory control, routing, workforce planning, returns, and satisfaction. Reviews classify these applications separately, and closed-loop supply chains remain underrepresented in big-data forecasting research [8]. Many studies use one firm, one product category, one country, or a single operational process. Delivery-position forecasting, for example, was validated in one company and identified the derivation of pick counts as an unresolved scheduling issue [23].

Validation practices also vary. Some studies report field experiments, while others use numerical experiments, simulations, cross-validation, surveys, or conceptual architectures. High predictive accuracy does not establish portability: a delivery-delay model may perform well in one dataset, while a near-real-time order-arrival model is demonstrated across three retailers but remains tied to a particular e-commerce operating context [22]. The literature rarely follows an acquisition signal through repeated campaign cycles, inventory reactions, delivery outcomes, returns, complaints, and customer lifetime value. Longitudinal integration is therefore the necessary bridge between predictive performance and closed-loop claims

7.4. Future Validation Designs: A/B Experiments, Rolling-Origin Forecast Tests, Digital Twins, and Multi-Echelon Field Studies

A credible program should combine complementary designs. A/B experiments can randomize campaign exposure, incentive level, message timing, or delivery promise while preserving operational constraints. Rolling-origin forecast tests should retrain models only with information available at each prediction date and evaluate point forecasts, intervals, stockouts, service levels, cycle time, returns, and profit. Digital twins can stress-test policies under promotion spikes, lead-time variation, channel substitution, and supplier disruption before live deployment; simulation is already used to compare intermittent-demand and inventory-state forecasting approaches [17]. Multi-echelon field studies should randomize or stagger adoption across stores, distribution centers, suppliers, and carriers, permitting estimates of network spillovers rather than isolated node effects.

Each design should compare statistical accuracy with decision outcomes. Predict-then-optimize and end-to-end learning require common baselines, identical information sets, cost accounting, calibration tests, and ablation of marketing features. Supplier allocation research demonstrates why this matters: changing the forecasting method altered both supplier selection and allocated order quantities [20]. The strongest evidence will therefore be a preregistered, staged evaluation in which acquisition treatments, forecasts, operational policies, and customer outcomes remain traceable across the network.

8. Conclusion

8.1. Theoretical and Managerial Implications of Connecting Acquisition Signals with Logistics Decisions

The marketing-logistics loop extends demand-driven forecasting from prediction toward coordinated decision-making. Customer acquisition signals describe intended demand, while inventory, capacity, and delivery states determine whether that demand can be converted into profitable service. Research on CRM-based routing, customer satisfaction, and demand-driven forecasting supports treating delivery reliability, order accuracy, and customer response as connected outcomes rather than separate departmental metrics. The theoretical implication is a shift from isolated functional models to a constrained feedback system in which acquisition, fulfillment, and post-purchase behavior update one another.

For managers, the loop changes the unit of analysis from campaign performance or forecast error to customer-level contribution under operational constraints. A promotion should be assessed by incremental margin after inventory, delivery, return, and service costs. A delivery intervention should be assessed by conversion, satisfaction, repeat purchase, and capacity consumption. This framing also clarifies why integration requires organizational agreements, shared identifiers, and common outcome definitions rather than data connection alone.

8.2. Implementation Priorities and Directions for Empirical Testing

Implementation should begin with governed identity resolution, event-time data, point-in-time feature snapshots, and leakage controls. Firms can then establish a baseline forecast and fulfillment policy before adding campaign, customer, and feedback features. Deployment should proceed through shadow scoring, human-reviewed exceptions, controlled experiments, and monitored expansion across products and locations. Evaluation must report forecast accuracy alongside service level, stockout rate, fulfillment cycle time, delivery reliability, returns, complaints, contribution margin, and customer satisfaction.

Empirical work should test heterogeneous effects across customer segments, products, channels, and echelons. It should also measure model drift after promotions, assortment changes, supply disruption, and policy adoption. The resulting evidence can determine when integrated intelligence improves customer-centered profitability, when simpler models are preferable, and which governance controls preserve reliable decisions as the loop operates over time.

References

1. E. W. T. Ngai and Y. Wu, “Machine learning in marketing: A literature review, conceptual framework, and research agenda,” Journal of Business Research, vol. 145, pp. 35–48, Mar. 2022, doi: 10.1016/j.jbusres.2022.02.049.
2. G. C. Souza, “Supply chain analytics,” Business Horizons, vol. 57, no. 5, pp. 595–605, Jul. 2014, doi: 10.1016/j.bushor.2014.06.004.
3. A. Narayanan, F. Sahin, and E. P. Robinson, “Demand and order-fulfillment planning: The impact of point-of-sale data, retailer orders and distribution center orders on forecast accuracy,” Journal of Operations Management, vol. 65, no. 5, pp. 468–486, Apr. 2019, doi: 10.1002/joom.1026.
4. B. D. Williams, M. A. Waller, S. L. Ahire, and G. D. Ferrier, “Predicting retailer orders with POS and order data: The inventory balance effect,” European Journal of Operational Research, vol. 232, no. 3, pp. 593–600, Jul. 2013, doi: 10.1016/j.ejor.2013.07.016.
5. M. Zaghloul, S. Barakat, and A. Rezk, “Predicting e-commerce customer satisfaction: Traditional machine learning vs. deep learning approaches,” Journal of Retailing and Consumer Services, vol. 79, p. 103865, Apr. 2024, doi: 10.1016/j.jretconser.2024.103865.
6. H. Cui, S. Rajagopalan, and A. R. Ward, “Predicting product return volume using machine learning methods,” European Journal of Operational Research, vol. 281, no. 3, pp. 612–627, Jun. 2019, doi: 10.1016/j.ejor.2019.05.046.
7. L. Baardman, S. B. Boroujeni, T. Cohen-Hillel, K. Panchamgam, and G. Perakis, “Detecting customer trends for optimal promotion targeting,” Manufacturing & Service Operations Management, vol. 25, no. 2, pp. 448–467, Oct. 2020, doi: 10.1287/msom.2020.0893.
8. M. Seyedan and F. Mafakheri, “Predictive big data analytics for supply chain demand forecasting: methods, applications, and research opportunities,” Journal Of Big Data, vol. 7, no. 1, Jul. 2020, doi: 10.1186/s40537-020-00329-2.
9. K. Douaioui, R. Oucheikh, O. Benmoussa, and C. Mabrouki, “Machine learning and deep learning models for demand forecasting in supply chain management: A critical review,” Applied System Innovation, vol. 7, no. 5,
10. p. 93, Sep. 2024, doi: 10.3390/asi7050093.
11. A. Martínez, C. Schmuck, S. Pereverzyev, C. Pirker, and M. Haltmeier, “A machine learning framework for customer purchase prediction in the non-contractual setting,” European Journal of Operational Research, vol. 281, no. 3, pp. 588–596, May 2018, doi: 10.1016/j.ejor.2018.04.034.
12. M. Qi et al., “A practical end-to-end inventory management model with deep learning,” Management Science, vol. 69, no. 2, pp. 759–773, Dec. 2022, doi: 10.1287/mnsc.2022.4564.

13. Y. Song, W. Wang, and S. Yao, “Customer acquisition via explainable deep reinforcement learning,” Information Systems Research, vol. 36, no. 1, pp. 534–551, May 2024, doi: 10.1287/isre.2022.0529.
14. P. Moder and K. Hoberg, “Predicting complaints in semiconductor order fulfilment with machine learning,” International Journal of Production Research, vol. 63, no. 13, pp. 4776–4799, Dec. 2024, doi: 10.1080/00207543.2024.2442548.
15. A. Kumar, R. Shankar, and N. R. Aljohani, “A big data driven framework for demand-driven forecasting with effects of marketing-mix variables,” Industrial Marketing Management, vol. 90, pp. 493–507, Jun. 2019, doi: 10.1016/j.indmarman.2019.05.003.
16. M. Zohdi, M. Rafiee, V. Kayvanfar, and A. Salamiraad, “Demand forecasting based machine learning algorithms on customer information: an applied approach,” International Journal of Information Technology, vol. 14, no. 4, pp. 1937–1947, Feb. 2022, doi: 10.1007/s41870-022-00875-3.
17. J. Feizabadi, “Machine learning demand forecasting and supply chain performance,” International Journal of Logistics Research and Applications, vol. 25, no. 2, pp. 119–142, Aug. 2020, doi: 10.1080/13675567.2020.1803246.
18. T. Schlaich and K. Hoberg, “When is the next order? nowcasting channel inventories with point-of-sales data to predict the timing of retail orders,” European Journal of Operational Research, vol. 315, no. 1, pp. 35–49, Nov. 2023, doi: 10.1016/j.ejor.2023.10.038.
19. X. Tian, H. Wang, and E. Erjiang, “Forecasting intermittent demand for inventory management by retailers: A new approach,” Journal of Retailing and Consumer Services, vol. 62, p. 102662, Jul. 2021, doi: 10.1016/j.jretconser.2021.102662.
20. Y. Yu, H. Wang, P. Cao, and S. Han, “Dynamic demand forecasting and ordering for new products considering the influence of inventory shortages,” International Journal of Production Research, pp. 1–28, Apr. 2026, doi: 10.1080/00207543.2026.2662977.
21. S. Islam, S. H. Amin, and L. J. Wardley, “Machine learning and optimization models for supplier selection and order allocation planning,” International Journal of Production Economics, vol. 242, p. 108315, Sep. 2021, doi: 10.1016/j.ijpe.2021.108315.
22. S. Ren, T. Choi, K.-M. Lee, and L. Lin, “Intelligent service capacity allocation for cross-border-e-commerce related third-party-forwarding logistics operations: A deep learning approach,” Transportation Research Part E Logistics and Transportation Review, vol. 134, p. 101834, Jan. 2020, doi: 10.1016/j.tre.2019.101834.
23.
24. K. H. Leung, D. Y. Mo, G. T. S. Ho, C. H. Wu, and G. Q. Huang, “Modelling near-real-time order arrival demand in e-commerce context: a machine learning predictive methodology,” Industrial Management & Data Systems, vol. 120, no. 6, pp. 1149–1174, May 2020, doi: 10.1108/imds-12-2019-0646.
25. P. Eichenseer, L. Hans, and H. Winkler, “A data-driven machine learning model for forecasting delivery positions in logistics for workforce planning,” Supply Chain Analytics, vol. 9, p. 100099, Dec. 2024, doi: 10.1016/j.sca.2024.100099.
26. M. M. Bassiouni, R. K. Chakrabortty, K. M. Sallam, and O. K. Hussain, “Deep learning approaches to identify order status in a complex supply chain,” Expert Systems with Applications, vol. 250, p. 123947, Apr. 2024, doi: 10.1016/j.eswa.2024.123947.
27. M. E. Bruni, E. Fadda, S. Fedorov, and G. Perboli, “A machine learning optimization approach for last-mile delivery and third-party logistics,” Computers & Operations Research, vol. 157, p. 106262, Apr. 2023, doi: 10.1016/j.cor.2023.106262.
28. M. A. Jahin, M. S. H. Shovon, J. Shin, I. A. Ridoy, and M. F. Mridha, “Big data—Supply chain management framework for forecasting: Data preprocessing and machine learning techniques,” Archives of Computational Methods in Engineering, vol. 31, no. 6, pp. 3619–3645, Mar. 2024, doi: 10.1007/s11831-024-10092-9.
29. F. H. Shavaki and A. E. Ghahnavieh, “Applications of deep learning into supply chain management: a systematic literature review and a framework for future research,” Artificial Intelligence Review, vol. 56, no. 5,
30. pp. 4447–4489, Sep. 2022, doi: 10.1007/s10462-022-10289-z.
31. S. Bag, L. C. Wood, L. Xu, P. Dhamija, and Y. Kayıkçı, “Big data analytics as an operational excellence approach to enhance sustainable supply chain performance,” Resources Conservation and Recycling, vol. 153,
32. p. 104559, Nov. 2019, doi: 10.1016/j.resconrec.2019.104559.
33. S. FNU and S. FNU, “CROSS-PLATFORM INTEGRATION OF PREDICTIVE ANALYTICS WITH
34. FULFILLMENT AND LOGISTICS OPERATIONS,” International Journal of Research In Commerce and Management Studies, vol. 08, no. 01, pp. 1057–1070, 2026, doi: 10.38193/ijrcms.2026.8186.

35. H. Lin, J. Lin, and F. Wang, “An innovative machine learning model for supply chain management,” Journal of Innovation & Knowledge, vol. 7, no. 4, p. 100276, Oct. 2022, doi: 10.1016/j.jik.2022.100276.
36. R. Thekkoote, “Understanding big data-driven supply chain and performance measures for customer satisfaction,” Benchmarking An International Journal, vol. 29, no. 8, pp. 2359–2377, Oct. 2021, doi: 10.1108/bij-01-2021-0034.
37. A. Brintrup et al., “Supply chain data analytics for predicting supplier disruptions: a case study in complex asset manufacturing,” International Journal of Production Research, vol. 58, no. 11, pp. 3330–3341, Nov. 2019, doi: 10.1080/00207543.2019.1685705.
38. R. Ishfaq and N. Bajwa, “Profitability of online order fulfillment in multi-channel retailing,” European Journal of Operational Research, vol. 272, no. 3, pp. 1028–1040, Aug. 2018, doi: 10.1016/j.ejor.2018.07.047.
39. J. Wei, Z. Ye, C. Yang, C. Chen, and G. Ma, “Process-Informed deep learning for enhanced order fulfillment cycle time prediction in on-demand grocery retailing,” Oct. 2024. doi: 10.1145/3627673.3680056.
40. S. Xu, H. Tang, and Y. Huang, “Inventory competition and quality improvement decisions in dual-channel supply chains with data-driven marketing,” Computers & Industrial Engineering, vol. 183, p. 109452, Jul. 2023, doi: 10.1016/j.cie.2023.109452.
41. F. Soygazi, H. AKDAĞ, and K. Öztürk, “ENHANCING SUPPLY CHAIN WITH INTEGRATION
42. ARCHITECTURE AND MACHINE LEARNING,” DergiPark (Istanbul University), 2025, [Online].
43. Available: http://jes.ksu.edu.tr/en/download/article-file/5194233
44. C. M. Giménez and E. Ventura, “Logistics-production, logistics-marketing and external integration,” International Journal of Operations & Production Management, vol. 25, no. 1, pp. 20–38, 2005, doi: 10.1108/01443570510572222.
45. A. Katangoori, “An empirical analysis of data-driven supply chain optimization in retail and logistics,” Supply Chain Analytics, vol. 14, p. 100206, Mar. 2026, doi: 10.1016/j.sca.2026.100206.