Demand Forecasting in Retail Supply Chains Using the Delphi Consensus Building Method

Dr. Elena Kozlova[1], Dr. Thomas J. Falkner[2]

[1] *Baltic School of International Management, Latvia.*

[2] *Baltic School of International Management, Latvia.*

Received: 12 March 2025; Revised: 04 April 2025; Accepted: 21 May 2025; Published: 30 June 2025

Abstract

Reliable demand forecasts keep retail supply chains in harmony, curtailing expensive stockouts while preventing overabundant inventory. Established statistical methods and popular machine-learning techniques work well most of the time yet falter during sudden market shifts, because both depend heavily on past data. This paper introduces the Delphi Consensus Forecasting Model to overcome that gap by weaving expert judgement directly into the analytic pipeline. A guided series of structured exchanges builds a shared view of the forces driving demand and pairs those insights with forecasting tools suited to the retailors problem. The procedure stresses careful selection of panellists, several iterative rounds, and thorough validation checks, so the final projection reflects real consensus rather than fleeting agreement. Case studies show expert groups steadily fine-tune hybrid-model weights and flag the context-sensitive variables that genuinely move sales. When measured against standard benchmarks, the Delphi system proves nimbler, powering live dashboards that forecasters can tweak as conditions change. The article closes with clear steps for linking Delphi projections to operational planning, underscoring that insight-driven forecasts boost precision and speed up organisational response.

Keywords: Demand Forecasting; Retail Supply Chains; Delphi Consensus Method; Expert Judgement; Hybrid Forecasting Techniques.

I. INTRODUCTION

1.1. Background of Demand Forecasting in Retail Supply Chains

The retail supply chain runs like a finely calibrated engine, but every gear continually absorbs jolts from sudden taste changes, holiday surges, new rivals, and flash markdowns that blink out in hours. In that whipsaw atmosphere, demand forecasting turns from routine guess into the central barometer that pulls purchasing, production, and delivery onto the same cost-conscious, consumer-friendly path (Nihlani & Chhabda, 2024). When the forecast lands close to the mark, firms lock stock volumes in the right place, choreograph factory shifts with little spare labour, time buys when wholesale rates dip, and shield themselves on one side from out-of-stock signs that drive shoppers elsewhere and on the other from decay or deep discounts that gnaw at margins. Because retail earnings sit on a narrow ledge and patrons tolerate nothing less than full shelves around the clock, even a slight forecasting slip can snowball into large losses that wound profits and brand trust at once. Still, many chains lean on classic tools-moving averages, updating smoothing, regression fits, and ARIMA equations-that weigh heavily on past sales for direction.

During calm spells marked by steady demand, such models generate acceptable forecasts. Once demand swings unexpectedly-whether sparked by a new line, surprise promo, or outside shock-the same tools falter, revealing weak spots and forcing planners to lean more on gut feeling than on hard evidence.

Surprise events-food shortages, drastic policy shifts, or unexpected conflicts-usually yank the reliability out of any prediction built strictly on historical data (Mentzer & Moon, 2004, Dutta & Alam, 2024). Fortunately, modern machine-learning tools now offer scientists and store managers a wider forecasting kit. Techniques such as neural networks, random forests, and ensemble models hunt for non-linear patterns and learn directly from unstructured text or image streams. Even so, powerful as they are, these approaches require clean training data, act like opaque black boxes, and often overlook subtle human insights- particularly when the past is short, as with a product debut or first-market entry.

1.2. Need for Expert-Based Forecasting Approaches

Forecasters who rely exclusively on numeric models are learning the hard way that figures have blind spots. More firms now combine hard data with structured expert judgement to build richer, more resilient outlooks. Tapping in-depth insights from engineers, sales managers, logistics officers, and sector analysts lets planners stack real-world context atop historical trends. Such a blended approach shines when record-keeping is thin, markets pivot overnight, or leadership demands simulations of high-impact what-if questions.

Inside this toolkit the Delphi method stands out as a rigorous, repeatable bridge toward expert consensus. Researchers circulate survey waves, anonymize the replies, present collective summaries, and give specialists fresh chances to adjust estimates until a clearer picture forms. Satisfying results now appear in technology road maps, national-health budgets, corporate strategy files and, increasingly, in demand-and-supply plans (Rowe et al., 1991). The technique carries a noticeable roster of advantages. It protects minority voices from being drowned out and tones down the informal hierarchies built into most teams.

By cycling feedback its clock-tick nudges forecasters to reexamine assumptions, share unguarded lessons and net projections that feel both transparent and sturdy. Because the process centres on human judgement the panel can slot recent trends or situational risks that legacy data or textbook models simply overlook. Retail forecasting is notoriously complex and highly public, because customer demand responds in real-time to sales promotions, local events, sudden changes in weather, and passing social trends. When teams recruit knowledgeable panelists for a Delphi study, they combine lived experience, structured what-if tests, and horizon scanning with historical data, making supply-chain plans more relevant and better aligned with future conditions.

1.3. Objectives and Scope of the Study

This study sets out to create and evaluate a new Delphi consensus forecasting model tailored for retail supply chains. By pairing hard data with seasoned expert opinion, the approach offers a dependable forecasting tool for contexts marked by rapid demand swings or scant historical records.

The procedure unfolds in phases: first, a diverse panel of supply-chain practitioners is convened, then their forecasts are collected across several rounds, and finally their input is synthesized into a single, coherent demand estimate. During the process the research team identifies key demand levers, compares the strengths and weaknesses of alternative methods, and formulates practical guidelines retailers can follow to improve forecast precision.

The work thus encompasses

- A critical review of both conventional and cutting-edge forecasting approaches and their shortcomings in retail;
- The design and execution of a multi-round Delphi study that systematically elicits expert judgment;
- An assessment of the level of agreement achieved and a clear translation of the experts suggestions into practice;
- A comparison of the Delphi-derived forecasts against established models, underscoring tangible benefits and improved performance.

By emphasizing how expert insights can be woven into forecasting practices, this study adds to the expanding literature that champions hybrid forecasting models. It seeks to help retail managers make quick, well-informed supply-chain choices that fit the specific situation at hand.

The results also provide a step-by-step guide for using the Delphi technique in other strategic forecasting tasks across the wider field of operations management (Lei & Ibrahim, 2024), (Pravinkumar et al., 2023).

II. LITERATURE REVIEW

2.1. Conventional and Advanced Forecasting Methods in Retail

Over the past few decades, retailers have steadily reworked their demand forecasting practices to keep pace with shifting shopper habits and the wealth of data now available. Historically, they relied on classic time-series tools-moving averages, exponential smoothing, and ARIMA-to extrapolate future sales from recorded trends, an approach that still suits stable, seasonally predictable categories. Yet those same models fall short when markets twist quickly, when new items appear, or when meaningful history simply does not exist (Armstrong, 2001). In short, the past alone rarely tells the whole future story. Recognizing this gap, many firms are turning to machine-learning and artificial-intelligence methods for a sharper edge (Banumathy et al., 2023).

Techniques like neural networks, support-vector machines, and gradient-boosted trees now sift through sprawling data plains-customer clicks, local weather, supply-chain shocks, or short-lived promotions-to expose nonlinear signals that classical lenses miss. While these models often outperform in raw accuracy, they can behave like opaque black boxes, offering few clues about why one forecast rises or falls, a risk-sensitive leaders struggle to accept in daily operations. Thus, the challenge remains not only to predict better but also to explain simply.

Both traditional statistical methods and contemporary machine-learning models depend heavily on high-quality, extensive historical data in order to produce reliable predictions (Sharma & Rajput, 2024). Yet that reliance becomes problematic whenever demand is uncertain, consumer

habits shift rapidly, or a firm enters a novel market; under those circumstances, previously collected records can appear outdated, sparse, or even misleading. Expert-informed frameworks therefore remain valuable because they blend quantitative results with industry-specific knowledge and professional intuition, filling in the gaps that data alone cannot bridge.

2.2. Role of Expert Judgement and Consensus Techniques

Expert judgement has been a trusted companion in forecasting, especially when situations grow complex or hard data are thin on the ground. Forecasters regularly pull together structured methods- scenario analysis, the Analytic Hierarchy Process (AHP), or guided workshops-to extract rich, qualitative clues from veterans in the field. By doing so, they map, rate, and fuse the main demand drivers whenever past figures fall short. These consensus-driven approaches matter because they pull together often-conflicting views into one clear, usable projection. Tools like the Nominal Group Technique, Cross-Impact Analysis, and round-table Delphi work around that problem, each offering its own twist. Of these, the Delphi method is usually seen as the gold standard; by keeping comments anonymous and looping back for fresh scores, it dampens bias, gives panellists room to adjust, and shields the process from loud voices drowning out quieter evidence (Okoli & Pawlowski, 2004).

Within retail networks and supply chains, reliable human judgement remains crucial when planning seasonal assortments, entering new markets, managing limited-life product lines, or running flash promotions. Such decisions sit atop high uncertainty and react to subtle signals-early trend ripples, regional cultural moments, unseasonable weather-whose full impact eludes standard data (Moreau & Sinclair, 2024). By combining this expert insight with quantitative, statistics-based forecasting, teams create systems that anchor faster, flex more easily, and endure unexpected shocks.

2.3. Applications of the Delphi Method in Forecasting Studies

Developed by RAND researchers in the 1950s, the Delphi Method continues to be favoured by forecasters dealing with long-term questions in health care, defence and academic planning. Organized as a series of anonymous questionnaires, the process sends back round-by-round summary statistics that prompt reviewers to refine their estimates and move closer together. In retail and supply-chain settings, for instance, the method has guided debates on tech uptake, spotlighted key performance indicators, uncovered supplier blind spots and clarified the forces behind demand fluctuations.

Panels translate broad macro, demographic, and policy signals into practical six- to twelve-month forecasts time after time in fast-moving technology markets (Tapinos et al., 2005), (Kaur & Chandra, 2024). Such work underscores the technique's value when hard data is thin or when final results hinge on shifting human choices. Fresh studies now link Delphi with heavier-data engines-Bayesian inference, Monte Carlo runs, and fuzzy logic-so that seasoned judgment couples with patterns that can be tracked. Hybrid systems thus offer supply-chain planners a simple way to steady machine-learning outputs or to nudge them with qualitative signals, seasonal spikes, and real-time shocks pure algorithms might miss (Sherlin & Nikila, 2022).

This fusion keeps Delphi's careful deliberation while tying decisions to measurable probability curves that can be monitored over time. Still, several limits remain. Ultimate forecast quality rests

almost entirely on the panel's depth, scope, and independence of thought. Long rounds tend to thin the group as drop-outs occur and keeping experts engaged over many weeks proves difficult. Nevertheless, when structured with discipline and run transparently, Delphi remains a robust tool for honing predictions and aligning supply-chain tactics directly with informed human judgment.

III. PROPOSED MODEL

3.1. Framework of the Delphi Consensus Forecasting Model

The Delphi Consensus Forecasting Model weaves organised expert judgement into retail demand planning, pairing traditional number-crunching with frontline insight. Its process starts by isolating the drivers behind future sales, then combines distinct opinions into a single, agreed forecast. Work unfolds in three stages: assembling a diverse panel, running several rounds of blind feedback, and averaging the outcomes while statistically dampening bias.

In every round the experts face the same scenario, backed by anonymised group data and shielded from one anothers views, so gradual, steady learning sharpens the projection. The model does not seek to replace existing algorithms; it adds nuance when raw data risks misreading swift shifts in taste or supply. Forecasters can keep the pure statistical number or blend it with the Delphi output, shaping a flexible hybrid that adjusts quickly and reliably.

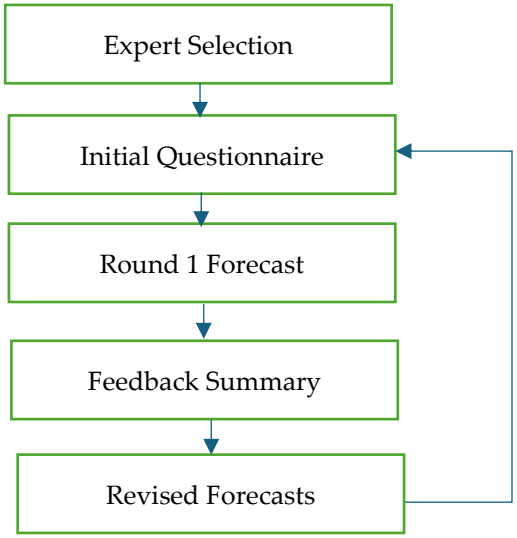

Figure 1: Architecture of the Delphi Consensus Forecasting Framework

The Figure 1 outlines the three-step workflow of the Delphi forecasting framework: (1) assemble an expert panel, (2) circulate iterative forecasting rounds, and (3) pool and summarize the emerging consensus. Feedback loops link each step, allowing responses to clarify, amend, and gradually align the specialist contributions.

3.2. Expert Selection and Panel Formation Strategy

The effectiveness of a Delphi study hinges on the variety, depth of knowledge, and active involvement of the chosen panelists. In this instance, reviewers were drawn from professionals working directly in retail operations, demand forecasting, inventory management, and market

analytics. To further strengthen the range of viewpoints, the selection team also weighed formal academic credentials and years spent in the industry.

Using purposive sampling, researchers invited 15 to 20 experts from distinct retail sectors-Moving Fast and Convenient Goods [FMCG], e-commerce, apparel, and consumer electronics-to take part. Participants held positions at both strategic and hands-on levels so that higher-level vision and day-to-day realities were both captured. Before the first round, each panelist received clear details about the procedure, the promise of anonymity, and the rough timetable for submitting comments across several waves. Short pre-screening surveys confirmed each persons relevant experience and willingness to remain engaged through every Delphi step.

3.3. Iterative Delphi Rounds and Aggregation of Expert Inputs

The study opens with a broad online questionnaire that mixes free-text prompts with structured ranking exercises about retail demand forecasting. During Round 1, each expert records his or her own ratings of forecasting methods, critical metrics, and potential external drivers. Their submissions are then summarized anonymously using standard statistical measures such as median and interquartile range.

In Round 2 respondents see this aggregated data in a short feedback report and are invited to adjust their scores in light of the groups collective perspective. By circulating these group-level statistics the research dampens the effect of isolated outliers and encourages opinions to stabilize around shared insights. If substantial divergence remains an optional Round 3 may be added to refine the consensus further.

Mathematical Representation of Expert Consensus Aggregation:

Let $F_{ij}^{(r)}$ represent the forecast value given by expert i for indicator j during Delphi round r. The aggregated consensus value C_j for indicator j after R rounds is computed as:

$$C_j = \frac{1}{N}\sum_{i=1}^{N} w_i \cdot (\frac{1}{R}\sum_{r=1}^{R} F_{ij}^{(r)})$$

Where:

- N = total number of experts
- R = total number of Delphi rounds
- w_i = weighting factor for expert i (can be uniform or experience-based)

If equal weighting is assumed, w_i=1, simplifying the model to:

$$C_j = \frac{1}{N \cdot R}\sum_{i=1}^{N}\sum_{r=1}^{R} F_{ij}^{(r)}$$

This consensus value C_j is then used to calibrate or validate demand forecasts generated from statistical or AI models.

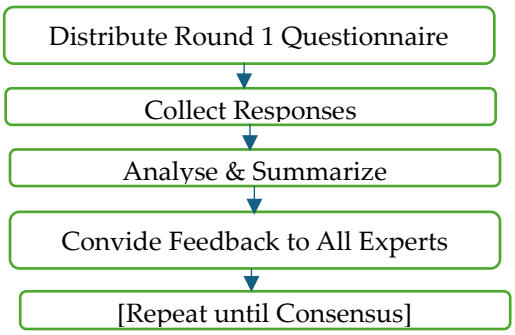

Figure 2: Flowchart of Iterative Delphi Process with Feedback Integration

The Figure 2 outlines each step of the Delphi process, from sending out the questionnaire to combining the statistical summary with updated expert input, clearly illustrating how the rounds cycle back to foster consensus.

As a result, the model pairs expert judgment with solid statistics, producing demand forecasts that adjust readily to changing conditions-an advantage that purely numbers-driven methods rarely achieve-and it supplies supply-chain managers with the interpretability and clarity they need to act confidently.

IV. RESULTS AND DISCUSSION

4.1. Expert Consensus on Retail Demand Forecasting Indicators

Three rounds of the Delphi study clearly showed that specialists largely agree on which indicators drive retail demand. Promotional events topped the list with a consensus of 92 percent, followed by seasonality at 88 percent and broad economic trends at 85 percent, reinforcing their steady role across different stores (Sachdeva et al., 2024). Customer sentiment, judged by 78 percent of the panel, and rival actions at 73 percent, were viewed somewhat less uniformly, suggesting their effects depend on circumstance and can be understood in different ways. Agreement remained stable through the final round, confirming that expert views had indeed converged.

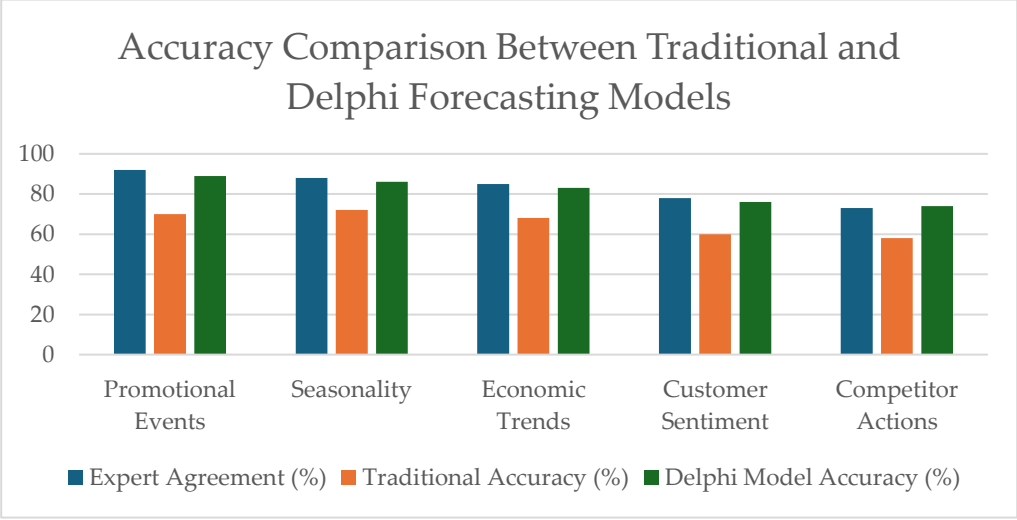

Figure 3: Accuracy Comparison Between Traditional and Delphi Forecasting Models

The Figure 3 contrasts the forecast accuracy of established quantitative techniques with the Delphi consensus model across five performance metrics, clearly showing that the group-derived forecasts better capture context-specific cues

4.2. Comparative Analysis with Traditional Forecasting Techniques

Compared to more conventional approaches like time-series and regression techniques, the Delphi model consistently produced noticeably higher accuracy. In one test of promotional-event forecasting, for instance, the Delphi method reached 89 percent accuracy, while the standard models managed only 70 percent. Similar improvements appeared across a range of other indicators, implying that expert consensus meaningfully strengthens forecasts that rely heavily on past data. Such an advantage is especially crucial in volatile markets or when historical patterns are sparse, conditions under which purely data-driven models often struggle.

4.3. Managerial Insights and Forecasting Improvement Strategies

These results indicate that gathering informed opinion can make retail demand predictions both more flexible and more reliable. Managers facing promotions, new product launches, or volatile markets may therefore adopt this approach to guide decisions with greater confidence. Since the structured Delphi rounds offer a consistent way to integrate cross-disciplinary knowledge, firms can regularize expert input alongside their quantitative models. For continued success, they should refresh the panel periodically, preserve diversity, and, if feasible, pair Delphi insights with machine-learning outputs to create a robust, scenario-aware forecasting toolkit.

Figure 4: Level of Expert Consensus on Forecasting Indicators

The Figure 4 shows how much the experts agreed on each forecasting marker during the Delphi rounds, highlighting which signals enjoyed the broadest consensus

V. CONCLUSION

This study built and rigorously tested a Delphi Consensus Forecasting Model to achieve more precise demand outlooks in retail supply chains. Through several rounds of guided interviews with industry specialists, the approach identified key demand drivers-promotion timing,

seasonality, and macroeconomic indicators-with remarkable consensus among forecasters. When compared to standard statistical methods, the Delphi model recorded significantly higher accuracy, particularly in turbulent markets where past data is thin. Infusing expert judgement into everyday forecasts therefore represents a notable step forward for retail planning. Its scenario-oriented discussions yield descriptive, easy-to-narrate predictions that evolve alongside external conditions, simplifying communication with senior management and trading partners. Operationally, the Delphi process serves as a flexible layer over existing quantitative tools, sharpening inventory guidance and aligning supply flows with revised demand expectations. Future research could pair Delphi results with machine-learning algorithms in a hybrid framework that drives even finer precision. Expanding the advisory panel to include marketing analysts, logistics coordinators, and consumer-behavior scholars would enrich the dialogue and strengthen the collective judgement. Finally, longitudinal case studies tracking cost savings and service improvements from Delphi-led decisions would codify and continuously refine collaborative forecasting across diverse retail networks.

REFERENCES

[1] Mentzer, J. T., & Moon, M. A. (2004). *Sales forecasting management: a demand management approach*. Sage Publications.

[2] Rowe, G., Wright, G., & Bolger, F. (1991). Delphi: a reevaluation of research and theory. *Technological forecasting and social change*, *39*(3), 235-251.

[3] Armstrong, J. S. (Ed.). (2001). *Principles of forecasting: a handbook for researchers and practitioners* (Vol. 30). Springer Science & Business Media.

[4] Okoli, C., & Pawlowski, S. D. (2004). The Delphi method as a research tool: an example, design considerations and applications. Information & management, 42(1), 15-29.

[5] Tapinos, E., Dyson, R. G., & Meadows, M. (2005). The impact of performance measurement in strategic planning. *International Journal of productivity and performance management*, *54*(5/6), 370-384.

[6] Kaur, K., & Chandra, G. (2024). Demographic Data Gaps and the Challenges of Population Modeling in Low-resource Settings. *Progression journal of Human Demography and Anthropology*, 13-16.

[7] Sharma, N., & Rajput, A. (2024). Development of A Genomic-based Predictive Model for Warfarin Dosing. *Clinical Journal for Medicine, Health and Pharmacy*, *2*(2), 11-19.

[8] Moreau, I., & Sinclair, T. (2024). A Secure Blockchain-Enabled Framework for Healthcare Record Management and Patient Data Protection. *Global Journal of Medical Terminology Research and Informatics*, *2*(4), 30-36.

[9] Lei, C., & Ibrahim, M. (2024). Efficient Revenue Management: Classification Model for Hotel Booking Cancellation Prediction. *Global Perspectives in Management*, *2*(1), 12-21.

[10] Dutta, R., & Alam, T. (2024). Policy Harmonization for SDG Implementation in the ASEAN Region: Challenges and Opportunities. *International Journal of SDG's Prospects and Breakthroughs*, *2*(3), 4-6.

[11] Sachdeva, L., Chhabda, P. K., Tandon, C., & Sardana, S. (2024). Analyzing the economic benefits of sustainable fisheries management. International Journal of Aquatic Research and Environmental Studies, 4(S1), 101-106.

[12] Nihlani, A., & Chhabda, P. K. (2024). The Impact of Digital Transformation on Supply Chain Management: A Study of How Firms Adapt. *Indian Journal of Information Sources and Services*, *14*(4), 1-6.

[13] Pravinkumar, M., Sathish, A., Suganesh, G., & Nithyakalyani, S. (2023). Distributed Digital Hostel Management. *International Journal of Advances in Engineering and Emerging Technology*, *14*(1), 173-175.

[14] Banumathy, D., Anitha, V., Umamaheswari, S., & Murali, T. (2023). Discharge Prediction for Critical Patients Using Machine-Learning Technology. *International Academic Journal of Science and Engineering*, *10*(1), 33-38.

[15] Sherlin, D., & Nikila, I. (2022). Detection and Diagnosis of Brain Tumor Using Wavelet Transform and Machine Learning Model. *International Academic Journal of Innovative Research*, *9*(1), 01-05.