Proceedings of the 5th Indian International Conference on Industrial Engineering and Operations Management, Vellore, Tamil Nadu, India, November 6-8, 2025

Proceedings of the International Conference on Industrial Engineering and Operations Management

Publisher: IEOM Society International, USA DOI: 10.46254/IN05.20250146
Published: November 6, 2025

Judgmental Forecasting: A Comprehensive Review of the Past Half-Century

Shivani Jaiswal and Sumit Kumar Yadav
Department of Management Studies
Indian Institute of Technology Roorkee
Roorkee, Uttarakhand, India
shivani_j@ms.iitr.ac.in, sumitky@ms.iitr.ac.in

Abstract

Forecasting is an integral part of supply chain management. Accurate forecasts lead to the effective utilization of inventory. The accuracy depends not only on how the forecasts are generated with the help of system statistical models, but also on how the human judgments incorporate the contextual information outside the system. The latter part is called judgmental forecasting, where the forecaster brings their experience, knowledge, and external information to the system to produce accurate forecasts. This paper presents a comprehensive review of the literature on judgmental forecasting from 1976 to 2025 (August). This is the first study to present fifty years of the review of judgmental forecasting to capture how this research field started and the status of its integration with the statistical system in the present. Drawing on 242 peer-reviewed journal papers from the Scopus database following the PRISMA framework, we conducted performance analysis and science mapping using tools like R Biblioshiny and VOSviewer. The future research directions are also discussed at the end of this article to help research scholars and industry practitioners delve into this niche research area to study and explore its practices.

Keywords
Judgmental forecasting, Judgmental adjustments, Human factors, Literature Review, Bibliometric analysis

1. Introduction
Forecast is the backbone of supply chain management. It is the prediction of future estimates using historical data. Under forecasts can result in the loss of sales, while over forecasts can result in inventory stocking, leading to holding costs. Forecasting is done using various statistical algorithms such as time series models, exponential smoothing, regression, ARIMA, etc. Accuracy is measured using the forecast error measurements. The smaller the error, the higher the accuracy. However, statistical algorithms alone cannot produce an accurate forecast. There are a lot of other contextual factors, such as price change, promotions, climate change, market intelligence, competitive environment, etc., which, when incorporated into the algorithms, can change the scenario of accuracy. These contextual factors cannot be incorporated automatically into the systems. It requires an expert with domain knowledge to bring these together and integrate them into the algorithm (Sroginis et al., 2023). Hence, the process of adjusting or revising any forecasts by the expert or forecaster is called judgmental forecasting. Forecast accuracy suffers from numerous judgmental adjustments, which take much time to monitor. Depending on the circumstances, these judgments may or may not increase forecast accuracy (Abolghasemi et al., 2020; Fildes et al., 2009). Forecasters often adjust the system generated forecasts if there is a lack of historical data as in case of new product, or they have various other external information which cannot be automatically integrated into the system, or to have a sense of ownership or mistaking noise for the signal leading them to make unnecessary modifications (Alvarado-Valencia et al., 2017; Petropoulos et al., 2016a). However, integrating external contextual information such as promotions, competitor activities, sudden weather changes, government policies, etc., is challenging. These factors are so dynamic that integrating them into the system introduces many variables, which become vulnerable to multicollinearity and dimensionality. Hence, high

prices, lack of resources, and organizational barriers make building and maintaining such models difficult (Trapero et al., 2015). Therefore, the task of experts' judgment becomes important.

The direction of adjustment also plays an important part in the forecast accuracy. A positive adjustment involves increasing the forecasted value, and a negative adjustment involves decreasing the predicted value to account for variables the original model did not consider. This can result from unanticipated favorable occurrences, modifications to the market environment, or fresh opportunities that could benefit the predicted outcome. The size of the adjustment also matters. Some studies show that small positive and large negative adjustments are valuable for improved accuracy (Fildes et al., 2009; Khosrowabadi et al., 2022). However, with human judgments comes the introduction of various biases and heuristics. According to Bendoly et al. (2009), heuristics primarily explain deviations in the decision-making process, whereas biases largely reflect variations in the outcome of decisions. There are a number of heuristics and biases studied in the literature, such as overconfidence bias, loss aversion bias, anchoring bias, recency bias, etc. (De Baets & Harvey, 2018; M. Lawrence & O'connor, 1995; Perera et al., 2019). When it comes to judgmental forecasting, Blattberg and Hoch (1990) recommend a 50-50 split between model and manager inputs because they discovered that the two regularly outperformed each other. Petropoulos et al. (2018) showed that taking the average of a forecast made by a model selected based on an algorithm and a forecast made by a model selected based on judgment can significantly exceed the statistical and judgmental selections separately. Judgmental forecasting is more about learning the best practices to produce an effective forecast. This comprehensive literature review will help us know how this field has evolved and how it has progressed over the past half-century. The remaining part of this paper is structured as follows: Research Questions in Section 2, Research Methodology in Section 3, Data Analysis and Findings in Section 4, Implications, Limitations, Future Research Directions in Section 5, followed by Conclusion in Section 6.

2. Research Questions

This comprehensive review paper collates the findings from multiple studies in the field of judgmental adjustments in forecasting by examining the existing research. The following are the research questions that we will be addressing through this study: RQ1: How has research volume and focus in judgmental forecasting evolved over the past half-century? RQ2: What are the most influential journals in judgmental forecasting literature? RQ3: Who are the most prolific scholars in this field? RQ4: Which top countries are participating in this research area? RQ5: What are the key foundational themes in this research area? RQ6: What are the recent or present key themes and developments in this research area? RQ7: What are the future relationships among topics in this field?

3. Research Methodology

We adhered to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines for a thorough and unbiased integration of the existing literature on Judgmental Forecasting. The extensive literature search and selection process is depicted in Figure 1.

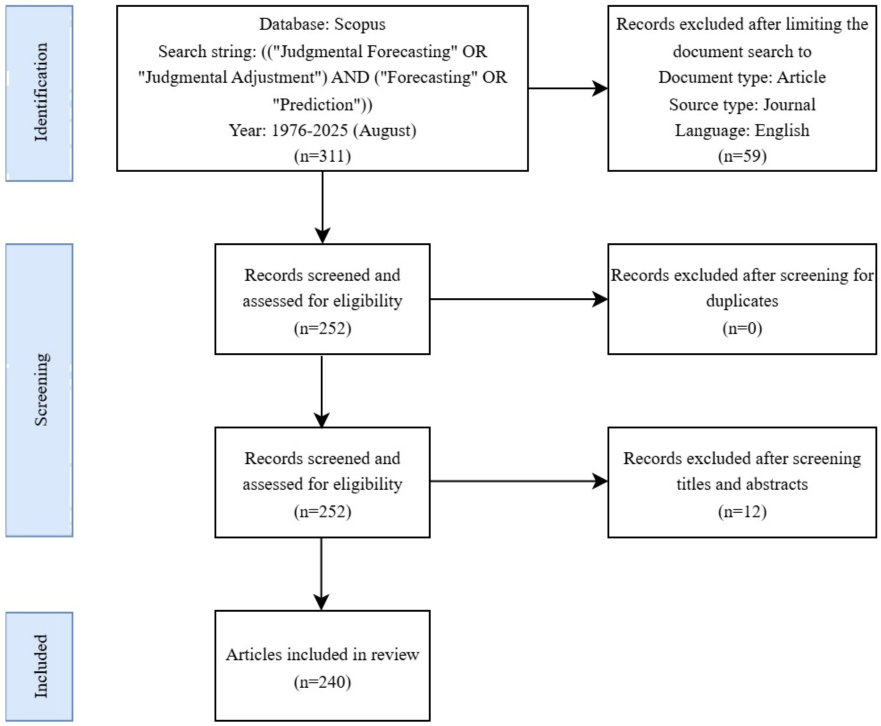

Figure 1. PRISMA flow chart

We have used the Scopus database for the literature search because it covers a large range of journals compared to other databases. We have collected studies published from 1976 to 2025 (August) timeframe. This time span was chosen to capture how this area has evolved over the past half-century. The literature search strategy employed a combination of keywords "Judgmental Forecasting", "Judgmental Adjustment", "Forecasting", and "Prediction" using the Boolean operators (AND/OR) to get the relevant papers. We obtained a total of 311 records from Scopus. To ensure the relevance of the collected studies, the screening and eligibility assessment were conducted. This followed a multi-step process, which was designed to ensure an impartial selection of high-quality relevant studies. After the Scopus database search that produced 311 records from 1976 to 2025 (August), a thorough screening approach was applied to identify the relevant documents. To obtain the most relevant articles for the scope of our study, we limited the results to Articles in Journal in English, which yielded 252 documents. We chose only the articles from journals to ensure the quality. The complete search query from the Scopus database was: TITLE-ABS-KEY (("Judgmental Forecasting" OR "Judgmental Adjustment") AND ("Forecasting" OR "Prediction")) AND (LIMIT-TO (SRCTYPE , "j")) AND (LIMIT-TO (DOCTYPE , "ar")) AND (LIMIT-TO (LANGUAGE , "English"))

The selected 252 records underwent a review to identify duplicate articles, and none were found. For the remaining 252 records, a secondary screening was performed that involved an in-depth review of titles and abstracts. While reviewing the titles and abstracts, studies were assessed according to established inclusion and exclusion criteria. Studies were excluded if the paper's focus was not on the judgmental adjustments to systematic forecasts, the paper was only about the application of different quantitative forecasting techniques, or the paper was a review paper. Ultimately, 12 more articles were dropped, and 240 articles were included for further analysis in our study.

4. Data Analysis and Findings

To answer our research questions for this bibliometric literature review, we used bibliometric tools like R Biblioshiny and VOSviewer, which are open-source tools for performance analysis and science mapping. Performance analysis helped us capture the publication trend, influential journals, authors, participating countries, keyword statistics, and citation statistics in this field. Science mapping helped us capture this research area's intellectual structure using co-citation analysis, bibliographic coupling, and co-word analysis. Our dataset contained 240 articles from 85 sources.

The average citation per document was 34.86. We detected 397 authors among the 240 articles. The findings are as follows:

Publication Trends: To address our first research question regarding the volume and expansion of judgmental adjustments in forecasting, we examined the evolution of judgmental forecasting. Figure 2 displays the yearly distribution of journal articles, revealing a fluctuating growth pattern starting from 1976. In fact, most of this domain's publications were published between 2008 and 2025. The number of publications showed an erratic trend with notable highs and lows. As of the writing date (August 2025), the number of articles peaked in 2020 and 2025 with 13 papers each, and it is anticipated that the 2025 numbers will be exceeded by all the past years. This shows that though this area is very niche, the scope of the trend is still there.

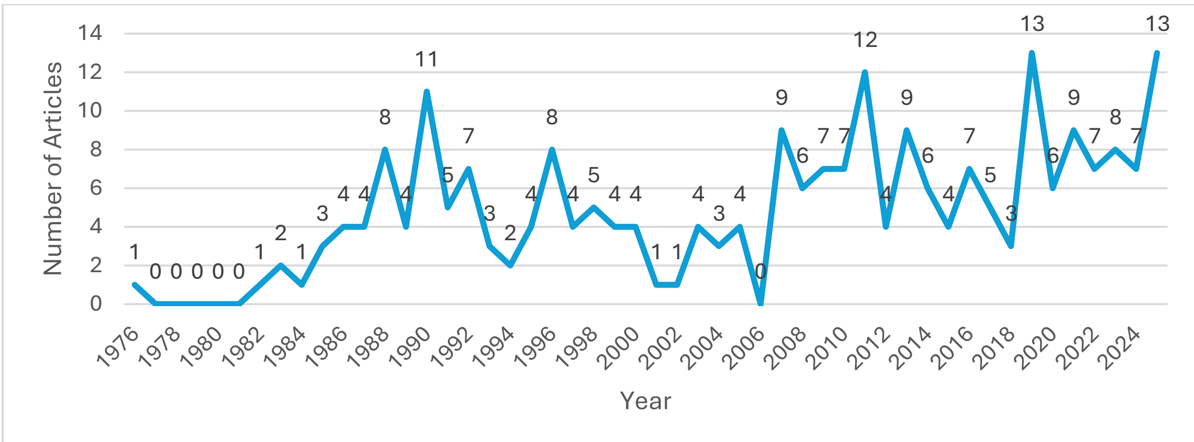

Figure 2. Publication Trend.

Most Influential Journals: To answer our research question 2, Figure 3 presents the top 10 prominent journals that have published documents in this area. Collectively, these journals published 152 articles, accounting for 63.33% of the 240 articles included in our study. The *International Journal of Forecasting* ranks first in the list with 84 papers. Next, the *Journal of Forecasting* has published 24 papers, while the *European Journal of Operational Research* has published 11 papers. The focus of these journals encompasses a variety of topics within judgmental forecasting, including seasonal adjustments, psychological aspects, organizational aspects of forecasting, and decision making under uncertainty. Thus, the articles published in these journals reveal the depth, variety, and interdisciplinary nature of research in the domain of judgmental forecasting.

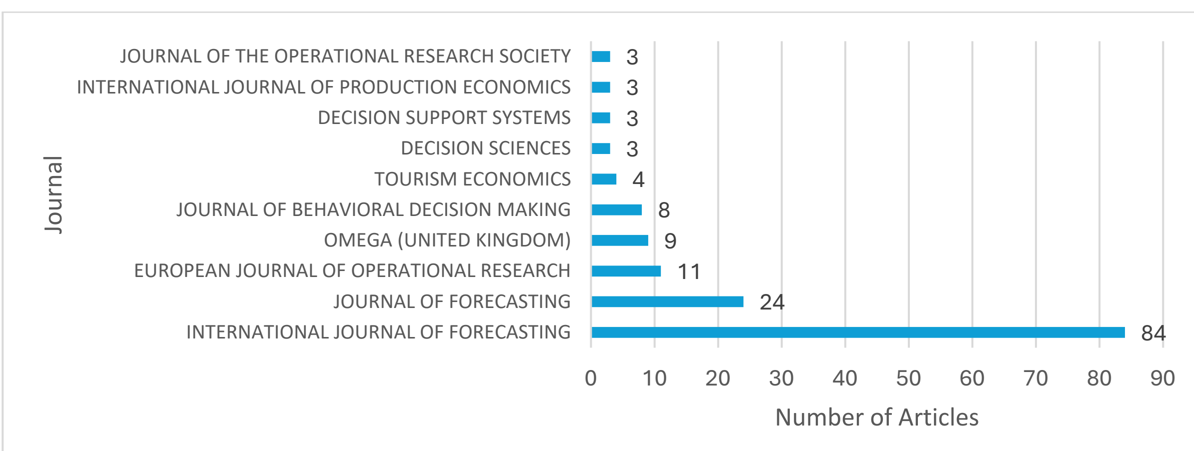

Figure 3. Influential Journals.

Most Influential Authors: To answer our research question 3, Table 1 presents the leading researchers in judgmental forecasting as identified in the Scopus database and ranked according to their publication output in prominent journals. The author with the most significant number of papers in this domain was Professor Paul Goodwin, affiliated with the University of Bath, Bath, United Kingdom. He is a featured author with 88 documents on the Scopus database and 4743 citations. The author has written about judgmental adjustments in forecasting in 23 articles, published between 1996 and 2025. Following closely, Professor Robert A. Fildes emerged as the second most prolific and influential author in this field and has written 17 articles related to judgmental forecasting, published between 1999 and 2025.

In their paper, Fildes et al. (2009) empirically investigated the strategies and directions in which judgmental adjustments could improve forecasting accuracy by collecting data on 60,000 forecasts and outcomes from four supply chain companies. Fildes and Goodwin (2007) discussed the principles of forecasting relating to the use of management judgment. The author highlights that justifications for making judgmental adjustments by documenting and assessing the results of judgmental interventions can improve the forecast accuracy of organizations. The author addresses that improving the current forecasting processes can add greater forecast value in their recent research work (Fildes et al., 2025).

Table 1. Number of articles, h-index, and citations of the top 10 authors.

S. No.	Authors	Articles	h-index	Citations
1	Paul Goodwin	23	16	1291
2	Robert A. Fildes	17	15	1331
3	Dilek Önkal	15	11	415
4	Marcus J. O' Connor	14	13	858
5	Michael J. Lawrence	13	13	1160
6	George Wright	10	8	263
7	Konstantinos Nikolopoulos	9	8	674
8	Ville A. Satopää	8	5	101
9	Peter Ayton	7	7	132
10	Shari De Baets	6	5	144

Top Countries: Table 2 illustrates the scientific output of the top 10 countries as found in Scopus, providing insights into our fourth research question. The UK and the USA lead in published papers, with Australia, Germany, Turkey, and the Netherlands following. This indicates a greater concentration of judgmental forecasting research outcomes in developed nations.

Table 2. Top 10 countries.

S. No.	Country	Number of Articles	Percentage
1	United Kingdom	175	28.88
2	United States of America	146	24.09
3	Australia	46	7.59
4	Germany	28	4.62
5	Turkey	25	4.13
6	Netherlands	20	3.30
7	France	18	2.97
8	Belgium	17	2.81
9	South Korea	16	2.64
10	China	15	2.48

Statistics of keywords: Figure 4 presents the word cloud, which shows the occurrence of the most frequent keywords. The larger the word in the word cloud, the more times it occurs in the document. Figure 4 shows that "judgmental forecasting" is the keyword that is used most often, highlighting its importance in forecasting, whether it is sales forecasting, energy forecasting, financial forecasting, or market forecasting. An expert always intervenes in the system to incorporate the information available outside of the system forecasts. Some other more frequently occurring keywords are "forecasting", "decision making", "judgment", and "judgmental adjustment". Judgmental adjustments to system-generated forecasts are one of the important parts of decision-making. Hence, this area needs more attention. In forecasting, positive and negative adjustments relate to the steps required to change a forecasted value in response to new information, changing conditions, or the performance of the forecasting model. These changes can increase or decrease the forecasted values' accuracy and reliability.

Figure 4. Word Cloud.

Citation analysis: There are many different metrics to evaluate a publication's impact, but the most common metric used is citations, which measures a publication's influence by counting the number of times it is referenced by other works (Donthu, Kumar, Mukherjee, et al., 2021). Table 3 shows the most cited articles in judgmental forecasting. Beiderbeck et al.'s (2021) article on preparing, conducting, and analyzing Delphi surveys is highly cited, with 355 citations. It is followed by Fildes et al. (2009), whose article on effective forecasting and judgmental adjustments has 330 citations, and by Makridakis et al.'s (1993) article on the real-time judgmentally based forecasting has 197 citations, respectively.

Beiderbeck et al.'s (2021) article provides a step-by-step guide on preparing, conducting, and analyzing a Delphi study by combining methodological advancements from other disciplines. This technique is widely used in judgmental forecasting to structure forecasters' opinions and predict technology adoption, promotion impacts, new product demand, or long-term demand where historical data is insufficient. Fildes et al.'s (2009) article discusses the impact of judgmental adjustments on forecast accuracy by comparing system-generated forecasts, judgmentally adjusted forecasts, and actual sales outcomes. The author found that large and negative adjustments often improved accuracy. They also developed a model to reduce the biases. The authors suggest a few improvement strategies, such as using the Blattberg-Hoch approach by restricting the weight adjustments to only 50% error bootstrapping, avoiding small adjustments, and reducing wrong-sided adjustments. Makridakis et al.'s (1993) article discusses the M2 competition, which included human judgmental forecasters, not just statistical models. The paper shows that judgment becomes most effective when combined with either more human forecasters or models.

Table 3. Top cited articles.

S. No.	Year	Document	Journal	Total Citations
1	2021	Beiderbeck et al. (2021)	MethodsX	355
2	2009	Fildes et al. (2009)	International Journal of Forecasting	330
3	1993	Makridakis et al. (1993)	International Journal of Forecasting	197
4	2007	Fildes and Goodwin (2007)	Interfaces	168
5	1991	Mitchell (1991)	Technology Analysis & Strategic Management	160
6	1993	Bunn and Salo (1993)	European Journal of Operational Research	156
7	2013	Davydenko and Fildes (2013)	International Journal of Forecasting	146
8	1995	Lim and O'Connor (1995)	Journal of Behavioral Decision Making	136
9	1985	M. J. Lawrence et al. (1985)	International Journal of Forecasting	123
10	1999	Goodwin and Fildes (1999)	Journal of Behavioral Decision Making	114

Co-citation analysis: It is a technique that assumes that publications cited together share thematic similarities. It discloses the intellectual structure of a research field, such as its underlying themes. This method is particularly adept at revealing seminal publications and knowledge foundations (Donthu, Kumar, Mukherjee, et al., 2021). To address our fifth research question, we conducted a document co-citation analysis of 1437 valid references that have been cited by 240 publications in our sample. With a significant number of references linked to judgmental forecasting, the aim was to enhance the network's interpretability and focus on the core publications. To achieve a manageable sample consisting of 50 valid references, we set a threshold that included cited references with a minimum of 5 citations. Figure 5 shows six major clusters we identified through document co-citation analysis. The closer the two publications are to each other, the more frequently they tend to be listed in the investigated bibliographies. The size of the node is directly proportional to the number of citations, underscoring the significance of each publication.

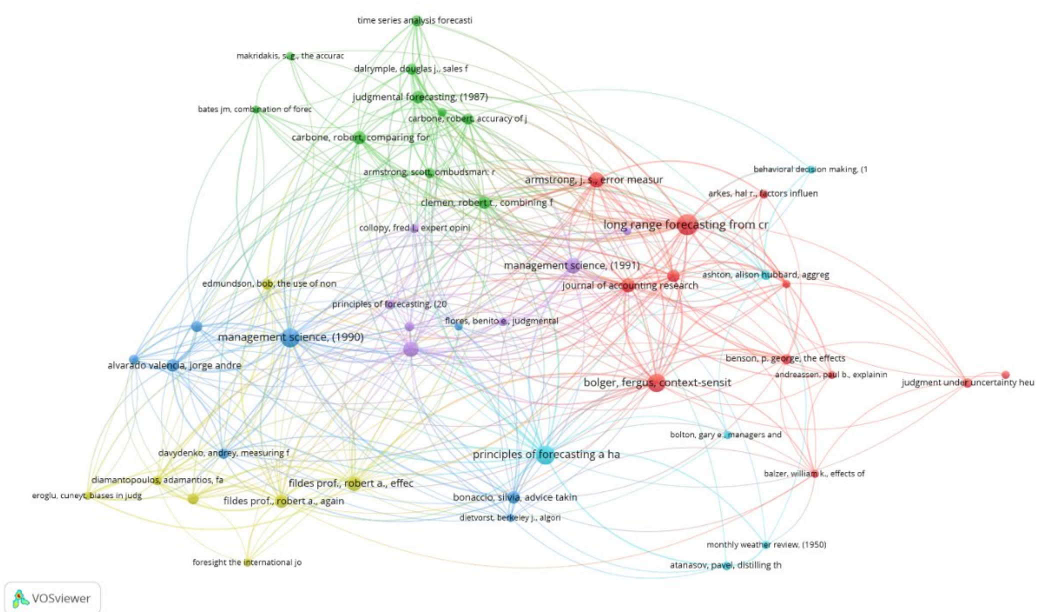

Figure 5. Co-citation analysis.

Cluster 1: This cluster, in red color, encompasses twelve articles primarily focusing on the cognitive processes, feedback performance, and accuracy measures in judgmental forecasting. In the study by Andreassen (1988), the subjects were asked to predict the price movements, and they made trading decisions based on heuristics, because an experiment supported the hypothesis of heavy trading due to large price changes. Their performance was dependent on how the information was presented to them. The author identified how cognitive biases and salience effects influence people's expectations about future prices and their trading behavior. Similarly, Andreassen and Kraus (1990) connected cognitive behavior with forecasting accuracy and bias in the context of time series. Armstrong and Collopy's (1992) article compared the common error measures for evaluating forecasting methods across different time series. Balzer et al. (1989) discussed cognitive feedback (CFB). They showed with the help of an experiment that task information is an essential component of (CFB). The CFB display format and CFB elaboration had very little effect on the judgmental performance. Benson and Önkal (1992) experimented and revealed that calibration feedback was more effective than outcome, resolution, and covariance feedback for effective forecasting by the forecasters.

Cluster 2: This cluster, in green color, encompasses ten articles primarily focusing on the gains of combining forecasts, comparison of judgmental and statistical methods, and the role of manual adjustments by humans. Bates and Granger's (1969) article introduced the idea of combining forecasts. The authors developed the mathematical framework for finding optimal weights for combining forecasts. They revealed that combined forecasts were more accurate than individual forecasts if they were not perfectly correlated. Clemen (1989) suggests and encourages that the simple combination of forecasts improves accuracy. An experiment by Carbone et al. (1983) showed that technical expertise, judgmental adjustments, and the time spent on analysis do not have an impact on improving forecast accuracy. Similar results were shown by Carbone and Gorr (1985), where judgmental adjustments improved the accuracy of only some objective forecasts. Angus-Leppan and Fatseas (1986) showed that an incautious combination of judgment and statistical methods can exacerbate accuracy.

Cluster 3: This cluster, in blue color, contains eight articles primarily focusing on how human judgment is integrated into the forecast support system (FSS). The key difference between combining and integrating forecasts is that we aggregate different types of forecasting methods; in integrating forecasts, we try to incorporate judgments into FSSs. A field study was done by Alvarado-Valencia et al. (2017), where they compared three types of judgmental integration methods. They found that judgmental adjustment often improves accuracy when the expert has high knowledge and low relative credibility of system forecasts. Arvan et al. (2019) showed the importance of integrating human judgment with quantitative models. Davydenko and Fildes (2013) recommend using the geometric mean of Mean Absolute Error ratios between adjusted and baseline forecasts to improve the accuracy of judgments.

Cluster 4: This cluster, in yellow color, contains seven articles primarily focusing on biases, product knowledge, information, and organizational practices in judgmental forecasting. Edmundson et al. (1988) and Diamantopoulos and Mathews (1989) revealed that having specific product knowledge improves forecast accuracy. The cluster also contains one of the highly cited papers in this domain by the authors Fildes and Goodwin (2007). The authors conducted a survey to find out if organizations follow the forecasting principles and found that organizations lack these principles, which guide the best practices in forecasting.

Cluster 5: This cluster, in purple color, contains seven articles focusing on the importance of experts' knowledge, their experience, and how they interact with the decision support systems (DSS). The study by Collopy and Armstrong (1992) shows that experts value seasonality, trends, aggregation, and sudden changes while selecting extrapolation methods. However, the treatment of discontinuities was still a neglected area in this field, even though experts saw it as an important thing to consider. Flores et al. (1992) introduced the centroid method for adjusting forecasts. Goodwin and Wright (1993) suggest the need for incorporating cognitive processes in DSS.

Cluster 6: This cluster, in sky blue color, contains six articles focusing on differential weighting, guidelines of forecasting, and statistical aggregation techniques. Ashton and Ashton (1985) highlight that adding more experts' forecasts to the aggregate can significantly improve the accuracy. In a study, (Bolton et al., 2012) found that the forecasters' knowledge level and their profile in the organization impacted the improved accuracy.

Bibliographic coupling: It is a technique that assumes two publications sharing common references are similar in content. It is especially designed to help understand the development of themes in the research domain by generating thematic clusters based on citing publications. This technique increases the visibility of recent and niche publications, making it a valuable tool for researchers. To answer our sixth research question, we selected articles for bibliographic coupling from 2015 to 2025 to identify recent research in judgmental forecasting (Donthu, Kumar, Pandey, et al., 2021). Thus, to achieve a sample consisting of 43 documents, we set the minimum number of citations of a document to 5. Figure 6 shows five major clusters we obtained with the help of document bibliographic coupling.

Cluster 1: This cluster, in red color, contains twelve articles that focus on aggregation methods, bias, information, noise, judgmental forecasting methods, cognitive biases, and time-series support. In a study by Satopää et al. (2021), the authors proposed a BIN (Bias, Information, Noise) Bayesian model, which helped improve the performance of forecasters and the forecasting methods. Similarly, Satopää et al. (2023) later used the BIN model on forecast aggregation techniques to improve accuracy. An experiment by Legaki et al. (2021) found that gamification can improve the learning outcomes regarding the biases and heuristics in judgmental forecasting by 15 percent. Satopää et al. (2016) proposed a model that can aggregate the probabilities provided by the forecasters for an event that will happen or not. De Baets and Harvey (2018) highlight statistical support that can take contextual information, such as promotions, into account, reducing biases, removing random error, and improving the accuracy of forecasts.

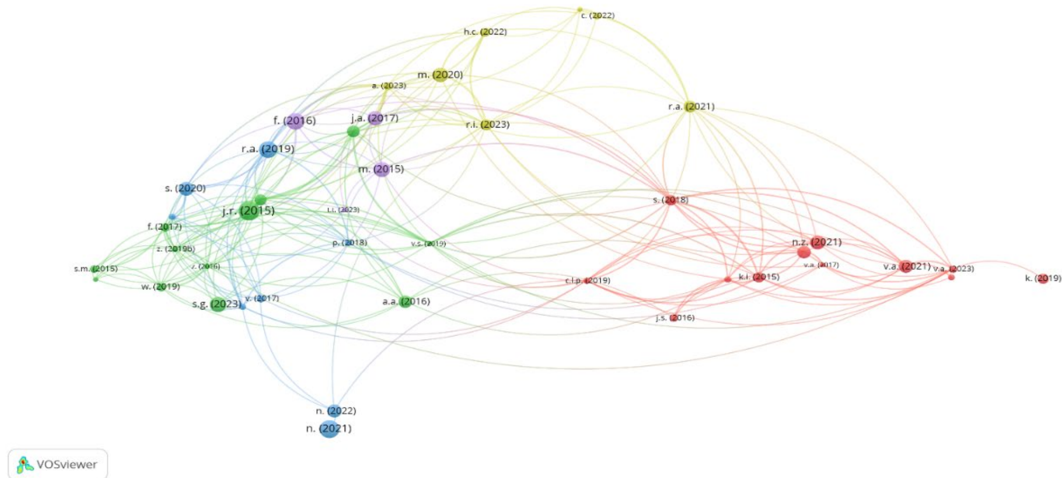

Figure 6. Bibliographic coupling

Cluster 2: This cluster, in green color, encompasses twelve articles focusing on the role of humans in forecasting, judgmental adjustments, data presentation effects, Large Language Models (LLMs)/ Artificial Intelligence (AI) in forecasting, brain imaging studies, and debiasing techniques. Promotions are found to be one of the main reasons for judgmental forecasting. Trapero et al. (2015) proposed a model that integrates Principal Component Analysis to overcome the issues of dimensionality and multicollinearity, which can handle the promotional variable to produce the forecasts. Makridakis et al. (2023) highlight using LLMs to support decision-making in judgmental forecasting with precautions. Syntetos et al. (2016) aim to study the judgmental adjustments to ERP-generated Order-Up-To (OUT) levels using data from 1800 SKUs and found that forecasting effectiveness depends on the size and direction of adjustment and justification of that adjustment in the form of documentation. In a laboratory experiment by Han et al. (2019), the participants' brain activity using electroencephalography (EEG) found that adjustments based on pattern identification performed better than forecast selection in the case of model selection.

Cluster 3: This cluster, in blue color, contains eight articles focusing on the interaction between human forecasters and AI and statistical models, mishandling of information, scenario-based biases, and anchoring effects. Kourentzes et al. (2021) identified that judgmental scenario planning paired with simple hierarchical forecasting is more reliable during unprecedented shocks like COVID-19. De Baets and Harvey (2020), with the help of an experiment, found that forecasters can select and adjust statistical model-based forecasts more effectively by learning from visual information of past performance. Khosrowabadi et al. (2022) identified that product characteristics were an important factor for manually adjusting AI-generated demands using ML algorithms such as random forests. They also investigated the quality of adjustments using the decision tree technique.

Cluster 4: This cluster, in yellow color, contains seven articles focusing on the interaction of human judgment with predictive analytics in supply chains, managerial advice, adjusting promotions, forecasting issues, personality traits, and systematic events. The article by Abolghasemi et al. (2020) introduced a regime-switching approach to quantify and incorporate external events such as sudden climate change, promotions, market dynamics, etc., into baseline forecasts. Fildes and Goodwin (2021) reveal that inefficient forecasting practices, such as judgmental overrides, have been continued for years because of model misfit, cognitive biases, and organizational incentives. The authors suggest that effective forecasting requires top-down standardization and changes in behavior and managerial practices. Brau et al. (2023) compared different types of forecast integration methods with their Human-Guided Learning model and found that their model was more effective in producing reliable forecasts.

Cluster 5: This cluster, in purple color, contains four articles focusing on losses in adjustments, forecasting for fashion and new products, and factors for making effective human judgments. Petropoulos et al. (2016) examined the effect of a significant decrease in forecast accuracy, often called 'big losses' because of the wrong direction of adjustments or judgmental overshoots. Niranjan et al. (2023) revealed in their experiment of the newsvendor task that people were dependent too much on recent demand, which is known as mean-reverting bias. Interventions such as training or hiding past demand can reduce this bias, suggesting that multiple cognitive biases interact and influence judgment.

Co-word analysis: It is a technique that inspects the actual content of the publication itself. The unit of analysis for the co-word analysis is "words". Similar to co-citation analysis, it assumes that words that frequently appear together have a thematic relationship. It offers a preview of the future of the research area (Donthu, Kumar, Mukherjee, et al., 2021). We kept the "Author keywords" as the unit of analysis. To answer our seventh research question, we obtained a sample of 37 keywords that frequently appeared together by keeping the minimum number of occurrences of a keyword to 5. Figure 7 shows the five major clusters obtained with the help of co-word analysis.

Cluster 1 in red contains twelve keywords. They are forecasting (77 occurrences), judgmental adjustment (25 occurrences), forecasting support systems (15 occurrences), demand forecasting (13 occurrences), forecast accuracy (10 occurrences), decision support systems (8 occurrences), sales (8 occurrences), statistical forecasting (7 occurrences), supply chain management (6 occurrences), demand planning (5 occurrences), human judgments (5 occurrences) and supply chains (5 occurrences). This cluster highlights how human judgment and statistical forecasting techniques are integrated with decision and forecast support systems in supply chain management and their effect on forecast accuracy. Cluster 2, in green, contains nine keywords. They are judgmental forecasting (141 occurrences), time series (17 occurrences), forecasting accuracy (10 occurrences), adjusting forecasts (8 occurrences), sales forecasting (8 occurrences), extrapolation (7 occurrences), time series analysis (6 occurrences), and heuristics (5 occurrences). This cluster contains one of the highly occurring keywords in this field, judgmental forecasting. Hence, this is the core cluster of our study. The cluster shows the relevance of judgmental forecasting in time series and the heuristics used by the expert. Cluster 3 in blue contains eight keywords, and they are decision making (25 occurrences), judgment (21 occurrences), article (10 occurrences), human (8 occurrences), prediction (7 occurrences), uncertainty (7 occurrences), decomposition (5 occurrences), and Delphi (5 occurrences). This cluster discusses the behavioral aspects of decision-making during judgmental forecasting. It focuses on the decomposition techniques for combining forecasts and the Delphi technique for structuring the adjustment process. Cluster 4 in yellow contains five keywords, and they are calibration (9 occurrences), evaluating forecasts (6 occurrences), probability forecasting (6 occurrences), subjective probability (6 occurrences), and combining forecasts (5 occurrences). This cluster emphasizes forecast evaluation and probabilistic methods. The cluster is about combining forecasts to maximize accuracy gains. Cluster 5 in purple contains three keywords: forecasting method (11 occurrences), accuracy (8 occurrences), and bias (7 occurrences). This cluster focuses on the bias that happens because of the judgments of the forecasters in different forecasting methods and how it impacts the accuracy.

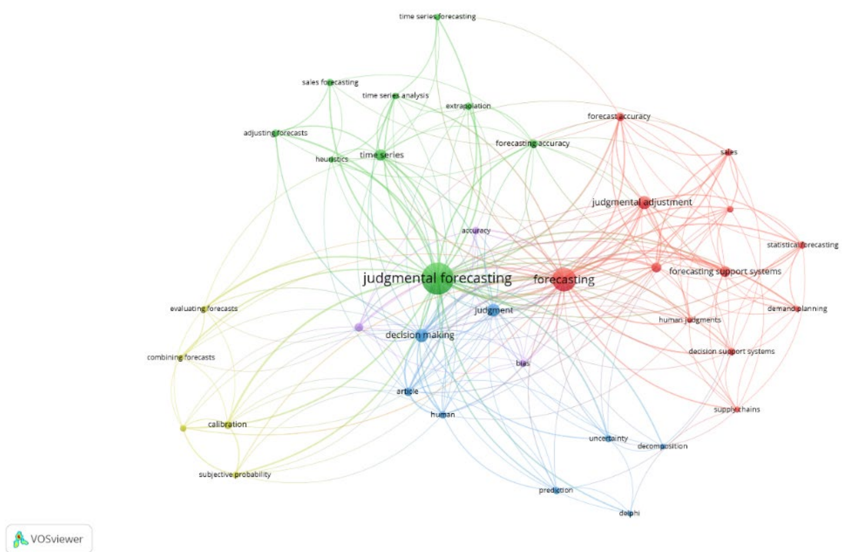

Figure 7. Co-word analysis

5. Implications, Limitations, and Future Research Directions

The study provided a comprehensive review of the literature on judgmental forecasting. It presented how the experts were performing judgmental forecasting in the very early stages and how they are performing now. It also explained the rising demand for judgments in forecasting, even when the technology has progressed over the years. This will help businesses, forecasters, demand planners, and experts to identify algorithms and systems that can integrate the valuable information provided by the experts, and which can document the reasons for making any adjustments, as adjustment documentation helps improve forecast accuracy. There are different fields of forecasting, such as sales, healthcare, energy, weather, and finance. Our study provided a review of all such fields of forecasting since we haven't used any specific keyword related to any particular field. Future studies can focus on a review of a particular field, such as sales forecasting. Our database was limited only to Scopus, though Scopus is considered to have a wide range of journals. Still, future studies can collect documents from other databases such as Web of Science, Google Scholar, etc., and investigate what other themes can emerge from the review. We have used tools like R Biblioshiny and VOSviewer. The future studies can utilize different tools and observe what different themes are coming out. Future research can also explore how the emerging technologies using AI-ML can benefit decision-making and manual adjustments in forecasting.

6. Conclusion

The revision of forecasts in the form of manual adjustments is very frequent today in industries. The performance analysis in our study helped us understand how this niche research area has grown over the past half-century. It helped us to know which author is most prolific in this area and which journals and countries capture the most research on this topic. Science mapping helped us delve into the intellectual structure of the research constituents. This helped us know about the past, present, and future of this area. The study helped us know how the size and direction of manual adjustments have affected the accuracy in various research and how the different forecast combination techniques have evolved. The study revealed various articles that emphasize the expertise and domain knowledge of the forecasters. In contrast, many studies discussed integrating judgments and statistical forecasts in the forecast support systems. Various studies emerged in our review that also talked about the biases introduced in judgmental forecasting.

References

Abolghasemi, M., Hurley, J., Eshragh, A., and Fahimnia, B., Demand forecasting in the presence of systematic events: Cases in capturing sales promotions. *International Journal of Production Economics*, vol. 230, no. 107892, 2020.

Alvarado-Valencia, J., Barrero, L. H., Önkal, D., and Dennerlein, J. T., Expertise, credibility of system forecasts and integration methods in judgmental demand forecasting. *International Journal of Forecasting*, vol. 33, no. 1, pp. 298–313, 2017.

Andreassen, P. B., Explaining the price-volume relationship: The difference between price changes and changing prices. *Organizational Behavior and Human Decision Processes*, vol. 41, no. 3, pp. 371–389, 1988.

Andreassen, P. B., and Kraus, S. J., Judgmental extrapolation and the salience of change. *Journal of Forecasting*, vol. 9, no. 4, pp. 347–372, 1990.

Angus-Leppan, P., and Fatseas, V., The Forecasting Accuracy of Trainee Accountants Using Judgemental and Statistical Techniques. *Accounting and Business Research*, vol. 16, no. 63, pp. 179–188, 1986.

Armstrong, J. S., and Collopy, F., Error measures for generalizing about forecasting methods: Empirical comparisons. *International Journal of Forecasting*, vol. 8, no. 1, pp. 69–80, 1992.

Arvan, M., Fahimnia, B., Reisi, M., and Siemsen, E., Integrating human judgement into quantitative forecasting methods: A review. *Omega*, vol. 86, pp. 237–252, 2019.

Ashton, A. H., and Ashton, R. H., Aggregating Subjective Forecasts: Some Empirical Results. *Management Science*, vol. 31, no. 12, pp. 1499–1508, 1985.

Balzer, W. K., Doherty, M. E., and O'Connor, R., Effects of cognitive feedback on performance. *Psychological Bulletin*, vol. 106, no. 3, pp. 410–433, 1989.

Bates, J. M., and Granger, C. W. J., The Combination of Forecasts. *OR*, vol. 20, no. 4, pp. 451, 1969.

Beiderbeck, D., Frevel, N., von der Gracht, H. A., Schmidt, S. L., and Schweitzer, V. M., Preparing, conducting, and analyzing Delphi surveys: Cross-disciplinary practices, new directions, and advancements. *MethodsX*, vol. 8, no. 101401, 2021.

Bendoly, E., Croson, R., Goncalves, P., and Schultz, K., Bodies of Knowledge for Research in Behavioral Operations. *Production and Operations Management*, vol. 19, no. 4, pp. 434–452, 2009.

Benson, P. G., and Önkal, D., The effects of feedback and training on the performance of probability forecasters. *International Journal of Forecasting*, vol. 8, no. 4, pp. 559–573, 1992.

Blattberg, R. C., and Hoch, S. J., Database Models and Managerial Intuition: 50% Model + 50% Manager. *Management Science*, vol. 36, no. 8, pp. 887–899, 1990.

Bolton, G. E., Ockenfels, A., and Thonemann, U. W., Managers and Students as Newsvendors. *Management Science*, vol. 58, no. 12, pp. 2225–2233, 2012.

Brau, R., Aloysius, J., and Siemsen, E., Demand planning for the digital supply chain: How to integrate human judgment and predictive analytics. *Journal of Operations Management*, 2023.

Bunn, D. W., and Salo, A. A., Forecasting with scenarios. *European Journal of Operational Research*, vol. 68, no. 3, pp. 291–303, 1993.

Carbone, R., Andersen, A., Corriveau, Y., and Corson, P. P., Comparing for Different Time Series Methods the Value of Technical Expertise Individualized Analysis, and Judgmental Adjustment. *Management Science*, vol. 29, no. 5, pp. 559–566, 1983.

Carbone, R., and Gorr, W. L., ACCURACY OF JUDGMENTAL FORECASTING OF TIME SERIES. *Decision Sciences*, vol. 16, no. 2, pp. 153–160, 1985.

Clemen, R. T., Combining forecasts: A review and annotated bibliography. *International Journal of Forecasting*, vol. 5, no. 4, pp. 559–583, 1989.

Collopy, F., and Armstrong, J. S., Expert opinions about extrapolation and the mystery of the overlooked discontinuities. *International Journal of Forecasting*, vol. 8, no. 4, pp. 575–582, 1992.

Davydenko, A., and Fildes, R., Measuring forecasting accuracy: The case of judgmental adjustments to SKU-level demand forecasts. *International Journal of Forecasting*, vol. 29, no. 3, pp. 510–522, 2013.

De Baets, S., and Harvey, N., Using judgment to select and adjust forecasts from statistical models. *European Journal of Operational Research*, vol. 284, no. 3, pp. 882–895, 2020.

De Baets, S., and Harvey, N., Forecasting from time series subject to sporadic perturbations: Effectiveness of different types of forecasting support. *International Journal of Forecasting*, vol. 34, no. 2, pp. 163–180, 2018.

Diamantopoulos, A., and Mathews, B., Factors affecting the nature and effectiveness of subjective revision in sales forecasting: An empirical study. *Managerial and Decision Economics*, vol. 10, no. 1, pp. 51–59, 1989.

Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., and Lim, W. M., How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research*, vol. 133, pp. 285–296, 2021.

Donthu, N., Kumar, S., Pandey, N., Pandey, N., and Mishra, A., Mapping the electronic word-of-mouth (eWOM) research: A systematic review and bibliometric analysis. *Journal of Business Research*, vol. 135, pp. 758–773, 2021.

Edmundson, B., Lawrence, M., and O'Connor, M., The use of non-time series information in sales forecasting: A case study. *Journal of Forecasting*, vol. 7, no. 3, pp. 201–211, 1988.

Fildes, R., and Goodwin, P., Against Your Better Judgment? How Organizations Can Improve Their Use of Management Judgment in Forecasting. *Interfaces*, vol. 37, no. 6, pp. 570–576, 2007.

Fildes, R., and Goodwin, P., Stability in the inefficient use of forecasting systems: A case study in a supply chain company. *International Journal of Forecasting*, vol. 37, no. 2, pp. 1031–1046, 2021.

Fildes, R., Goodwin, P., and De Baets, S., Forecast value added in demand planning. *International Journal of Forecasting*, vol. 41, no. 2, pp. 649–669, 2025.

Fildes, R., Goodwin, P., Lawrence, M., and Nikolopoulos, K., Effective forecasting and judgmental adjustments: an empirical evaluation and strategies for improvement in supply-chain planning. *International Journal of Forecasting*, vol. 25, no. 1, pp. 3–23, 2009.

Flores, B. E., Olson, D. L., and Wolfe, C., Judgmental adjustment of forecasts: A comparison of methods. *International Journal of Forecasting*, vol. 7, no. 4, pp. 421–433, 1992.

Goodwin, P., and Fildes, R., Judgmental forecasts of time series affected by special events: does providing a statistical forecast improve accuracy? *Journal of Behavioral Decision Making*, vol. 12, no. 1, pp. 37–53, 1999.

Goodwin, P., and Wright, G., Improving judgmental time series forecasting: A review of the guidance provided by research. *International Journal of Forecasting*, vol. 9, no. 2, pp. 147–161, 1993.

Han, W., Wang, X., Petropoulos, F., and Wang, J., Brain imaging and forecasting: Insights from judgmental model selection. *Omega*, vol. 87, pp. 1–9, 2019.

Khosrowabadi, N., Hoberg, K., and Imdahl, C., Evaluating human behaviour in response to AI recommendations for judgemental forecasting. *European Journal of Operational Research*, vol. 303, no. 3, pp. 1151–1167, 2022.

Kourentzes, N., Saayman, A., Jean-Pierre, P., Provenzano, D., Sahli, M., Seetaram, N., and Volo, S., Visitor arrivals forecasts amid COVID-19: A perspective from the Africa team. *Annals of Tourism Research*, vol. 88, no. 103197, 2021.

Lawrence, M. J., Edmundson, R. H., and O'Connor, M. J., An examination of the accuracy of judgmental extrapolation of time series. *International Journal of Forecasting*, vol. 1, no. 1, pp. 25–35, 1985.

Lawrence, M., and O'connor, M., The anchor and adjustment heuristic in time-series forecasting. *Journal of Forecasting*, vol. 14, no. 5, pp. 443–451, 1995.

Legaki, N.-Z., Karpouzis, K., Assimakopoulos, V., and Hamari, J., Gamification to avoid cognitive biases: An experiment of gamifying a forecasting course. *Technological Forecasting and Social Change*, vol. 167, no. 120725, 2021.

Lim, J. S., and O'Connor, M., Judgemental adjustment of initial forecasts: Its effectiveness and biases. *Journal of Behavioral Decision Making*, vol. 8, no. 3, pp. 149–168, 1995.

Makridakis, S., Chatfield, C., Hibon, M., Lawrence, M., Mills, T., Ord, K., and Simmons, L. F., The M2-competition: A real-time judgmentally based forecasting study. *International Journal of Forecasting*, vol. 9, no. 1, pp. 5–22, 1993.

Makridakis, S., Petropoulos, F., and Kang, Y., Large Language Models: Their Success and Impact. *Forecasting*, vol. 5, no. 3, pp. 536–549, 2023.

Mitchell, V. W., The delphi technique: an exposition and application. *Technology Analysis & Strategic Management*, vol. 3, no. 4, pp. 333–358, 1991.

Niranjan, T. T., Ghosalya, N. K., Menon, R. R., Rotaru, K., and Gavirneni, S., Unpacking the cognitive processes of the boundedly rational newsvendor. *Production and Operations Management*, vol. 32, no. 10, pp. 3138–3157, 2023.

Perera, H. N., Hurley, J., Fahimnia, B., and Reisi, M., The human factor in supply chain forecasting: A systematic review. *European Journal of Operational Research*, vol. 274, no. 2, pp. 574–600, 2019.

Petropoulos, F., Fildes, R., and Goodwin, P., Do 'big losses' in judgmental adjustments to statistical forecasts affect experts' behaviour? *European Journal of Operational Research*, vol. 249, no. 3, pp. 842–852, 2016.

Petropoulos, F., Kourentzes, N., Nikolopoulos, K., and Siemsen, E., Judgmental selection of forecasting models. *Journal of Operations Management*, vol. 60, no. 1, no. 34–46, 2018.

Satopää, V. A., Pemantle, R., and Ungar, L. H., Modeling Probability Forecasts via Information Diversity. *Journal of the American Statistical Association*, vol. 111, no. 516, pp. 1623–1633, 2016.

Satopää, V. A., Salikhov, M., Tetlock, P. E., and Mellers, B., Bias, Information, Noise: The BIN Model of Forecasting. *Management Science*, vol. 67, no. 12, pp. 7599–7618, 2021.

Satopää, V. A., Salikhov, M., Tetlock, P. E., and Mellers, B., Decomposing the effects of crowd-wisdom aggregators: The bias–information–noise (BIN) model. *International Journal of Forecasting*, vol. 39, no. 1, pp. 470–485, 2023.

Sroginis, A., Fildes, R., and Kourentzes, N., Use of contextual and model-based information in adjusting promotional forecasts. *European Journal of Operational Research*, vol. 307, no. 3, pp. 1177–1191, 2023.

Syntetos, A. A., Kholidasari, I., and Naim, M. M., The effects of integrating management judgement into OUT levels: In or out of context? *European Journal of Operational Research*, vol. 249, no. 3, pp. 853–863, 2016.

Trapero, J. R., Kourentzes, N., and Fildes, R., On the identification of sales forecasting models in the presence of promotions. *Journal of the Operational Research Society*, vol. 66, no. 2, pp. 299–307, 2015.

Biographies

Shivani Jaiswal is a research scholar at the Department of Management Studies at the Indian Institute of Technology Roorkee, Uttarakhand, India. She holds a bachelor's degree in industrial and production engineering from the Institute of Engineering and Rural Technology, Prayagraj, Uttar Pradesh, India. Her research interest areas include practices in sales forecasting, judgmental forecasting, and multi-criteria decision-making techniques.

Sumit Kumar Yadav is an Assistant Professor at the Department of Management Studies, Indian Institute of Technology Roorkee, Uttarakhand, India. He holds a PhD in Production and Quantitative Methods from the Indian Institute of Management, Ahmedabad, Gujarat, and a bachelor's degree in mechanical engineering from the Indian Institute of Technology, Bombay, India. His research interest areas include machine learning, analytics and statistics, stochastic processes, optimization, and decision-making techniques.