Expert Systems With Applications 263 (2025) 125719

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

Human vs. Machines: Who wins in semiconductor market forecasting?

Louis Steinmeister [a,b,*], Markus Pauly [a,c]

[a] Department of Statistics, TU Dortmund University, Vogelpothsweg 87, 44227 Dortmund, Germany
[b] Graduate School of Logistics, Leonhard-Euler-Straße 5, 44227 Dortmund, Germany
[c] Research Center Trustworthy Data Science and Security, Joseph-von-Fraunhofer-Straße 25, 44227 Dortmund, Germany

A R T I C L E   I N F O

Dataset link: wsts.org

Keywords:
Prediction
Sales forecast
Semiconductor cycle
Demand forecast
Machine learning
Statistical learning

A B S T R A C T

"*If you ask ten experts, you will get ten different opinions.*" This common proverb illustrates the common association of expert forecasts with personal bias and lack of consistency. On the other hand, digitization promises consistency and explainability through data-driven forecasts employing machine learning (ML) and statistical models. Despite the importance of the semiconductor industry being widely recognized, little research has gone into forecasting the whole semiconductor market including all major product categories. Instead, analysts have generally relied on expert forecasts such as those provided by the World Semiconductor Trade Statistics (WSTS). In the following, we generate data-driven forecasts and evaluate whether existing industry expert forecasts can be further enhanced through statistical and ML models. This study contributes by systematically evaluating the accuracy of expert forecasts, examining comprehensive multi-granularity forecasts for the entire semiconductor market, and offering performance insights through out-of-sample error measures to guide future forecasting practitioners.

1. Executive summary

Objectives: The objective of this paper is to evaluate and contrast expert predictions of the World Semiconductor Trade Statistics (WSTS), a leading semiconductor market data provider, with data-driven forecasts. In this context, we compare expert forecasts with different data-driven forecast approaches with respect to three research hypotheses detailed in the introduction.

Motivation: WSTS plays a crucial role in the semiconductor industry. According to their website, WSTS is the "*most respected source of market data and forecasts for the semiconductor industry*" and their forecasts "*are the only ones that leverage the collective experience of the industry's major players with the market intelligence of a large portion of the semiconductor industry*" (WSTS.org, 2024). As one of the top providers of comprehensive semiconductor industry data and indicators, WSTS plays a pivotal role in business decision making and industry analyst research. Additionally, the well-being of the semiconductor industry, which lies upstream in the supply chain, has been identified to be a leading indicator for the broader economy (Chow & Choy, 2006). This highlights the importance of accurate and reliable semiconductor industry forecasts even outside of this specific industry.

Methods: Several popular statistical and ML methods for time series forecasting are evaluated against official forecasts provided by WSTS by means of a time series cross validation.

Results: This paper finds that the expert forecasts provided by WSTS compare favorably to ML forecasts on a quarterly horizon but can nevertheless be enhanced by data driven forecasts. However, the performance of WSTS forecasts is put in perspective when the WSTS algorithmic updates, which are published bi-quarterly, are included. Furthermore, it can be argued that additional information should be incorporated into the forecasts, which results in a clear outperformance of the data-driven methods in comparison to the official WSTS forecasts. This discovery remains consistent regardless of the length of available data points. In other words, the outcome remains unaffected whether we analyze product categories with short histories or those with long histories.

Contribution: This study contributes in the following ways: (1) It provides a novel evaluation of the accuracy of expert forecasts within the semiconductor market, (2) it shows that comprehensive forecasting across various levels of granularity for the entire semiconductor industry is feasible, even with simple models, and (3) it provides valuable guidance to forecasting practitioners by supplying out-of-sample error measures for all analyzed models and product categories.

Conclusion: While WSTS forecasts provide a strong starting point, it is possible to improve the forecast accuracy through data-driven approaches.

* Correspondence to: Graduate School of Logistics, TU Dortmund University, Leonhard-Euler-Straße 5, 44227 Dortmund, Germany.
E-mail addresses: louis.steinmeister@tu-dortmund.de (L. Steinmeister), pauly@statistik.tu-dortmund.de (M. Pauly).

https://doi.org/10.1016/j.eswa.2024.125719
Received 2 August 2024; Received in revised form 4 October 2024; Accepted 4 November 2024
Available online 13 November 2024

2. Introduction

2.1. Motivation

In the today's dynamic business landscape, accurate forecasts play an increasingly important role in shaping operative and strategic decisions. The article "Bringing a real-world edge to forecasting" released by McKinsey & Company in 2020 makes the case that "*[a 'good' forecasting process] should be accurate enough to inform a range of critical business decisions – capital reallocation, hiring, strategy, production, and more*" (Agrawal et al., 2020).

For example, Wang et al. (2024) propose a freight rate index forecasting model to help industry players mitigate risks in the shipping market and inform investment opportunities and Wu et al. (2024) developed a model to forecast container throughput to inform port logistics industry decision makers. Additionally, Yu et al. (2012) propose a forecasting model to predict the fashion color trend leading to a higher success rate of new fashion products and an example of how resource allocation can be improved by increased operational efficiency through more accurate short to mid-term demand forecast is provided by Pauly and Kuhlmann (2023). Furthermore, forecasts play an important role in anticipating technological change (Foster, 1986; Modis, 1999) and estimating technology and product life cycles, which inform important portfolio and product development decisions (Modis, 1994; Petropoulos et al., 2022; Steinmeister et al., 2023). This is particularly true for the semiconductor industry with long lead times, a dynamic technological environment, and shortening product life cycles (Lv et al., 2018; Macher, 2006; Wu & Chien, 2008).

Accurate forecasts of the semiconductor market are relevant to the broader economy. The global semiconductor market reached sales of \$618 billion in 2022 according to Alsop (2024b). This amounted to about 0.61% of global GDP in 2022 compared to 0.22% in 1990, highlighting the increasing importance of the industry to the global economy (Alsop, 2024a; World Bank, 2024). Semiconductors are ubiquitous in modern life. They enable AI applications, modern defense equipment, data centers, automobiles, wireless communications, all the way to home appliances like your washing machine, gaming console, and electronic toothbrush. As Chow and Choy (2006) observe, the semiconductor market is a leading indicator for the broader economy.

The strategic importance of the semiconductor industry has further been recognized by governments. The US authorized about \$280 billion for research and manufacturing of semiconductors in the US with the CHIPS and Science Act passed in August 2022 (Taylor, 2023). Likewise, the EU subsidizes the industry with roughly 43 billion Euros (roughly \$46.3 billion) (European Comission, 2023). Even more impressively, Sam Altman, the current CEO of OpenAI, is seeking \$5 trillion to \$7 trillion of investments to "*boost the world's chip-building capacity and expand its ability to power AI, among other things*" according to Reuters (Rajan, 2024). To add perspective: this amounts to the combined market capitalization of Microsoft and Apple, the two largest American companies by market capitalization, at the time of writing.[1]

This motivates a detailed study of semiconductor market forecasts, which often inform industry and financial analysts. Furthermore, semiconductor market forecasts often inform internal goal setting and benchmarking. Semiconductor market data also factors into the calculation of market share by product groups. Projections of these quantities can have strategic implications.

The World Semiconductor Trade Statistics (WSTS)[2] is a premier provider of such data and forecasts. Several academic studies and industry reports cite WSTS as an authority for semiconductor market forecasts, see Corder et al. (2024), Nagao (2019) and Simons (2024) to name a few. Furthermore, the standing of WSTS as a data provider is highlighted by the common use of their data in academic fields such as the semiconductor cycle prediction, as outlined in Table 2 which provides a selection of related works in the field.

Industry forecasts, such as those from WSTS, serve as the basis for important strategic decisions and are often based on expert judgement. A popular method for consolidating these forecasts is the Delphi method (Armstrong, 2008; Delphi, 1975; Hyndman & Athanasopoulos, 2018). However, several researchers have discovered that these collective predictions, also referred to as "wisdom of the crowds" or "collective intelligence", are susceptible to inaccuracies and low precision, particularly when the individuals surveyed are pundits or uninformed laypersons (Atanasov et al., 2015; Modis, 1999).

However, the industry members polled by WSTS are industry experts and have access to detailed insider information, such as customer orders and sentiment, customer-related project status, the status of customer contracts, and new product development. Expert forecasts are generally thought to perform well when systems are complex, dynamic, and available history is sparse (Hyndman & Athanasopoulos, 2018). The semiconductor industry is heavily intertwined in the global economy, with complex upstream, and downstream supply chains. This means that the semiconductor industry is exposed to the bullwhip effect (Lee et al., 1997). It is also increasingly subjected to geopolitical considerations. While these factors highlight the importance of accurate semiconductor industry forecasts, the dynamic technological environment, geopolitical factors, and the complex supply chains concurrently complicate the data-driven forecast process due to the amount of potentially relevant extrinsic information. Therefore, expert forecasts from industry insiders queried by WSTS, who possess extensive access to quantitative and qualitative insider information, are anticipated to be highly reliable. Moreover, the lack of research into the area of detailed semiconductor market forecasts, illustrated in the related works section (Section 2.2), raises the question whether industry experts may simply be more accurate than data-driven models in this field. This prompts inquiry into the potential competitiveness of data-driven forecasts, and if and how they might enhance existing expert forecasts. To our knowledge the accuracy of semiconductor industry forecasts has not been systematically evaluated despite their regular use. The present paper closes this gap. To this end, the following three research hypotheses are examined in the following:

(H1) Expert forecasts exhibit higher accuracy compared to autoregressive data-driven forecasts.

WSTS publishes quarterly forecasts for each product category in an alternating pattern: expert forecasts are issued in May and November, while algorithmically computed updates are provided in February and August. These algorithmic updates are derived from the preceding quarter's results. Industry experts also only have access to official WSTS figures dating back to the previous quarter. Nevertheless, the numbers for the initial month of the forecasted quarter are disclosed simultaneously with the forecasts. Furthermore, it can be anticipated that industry experts possess internal information regarding the first month's data (for a detailed account of the data, see Section 4.1). This gives rise to the second hypothesis:

(H2) The incorporation of additional autoregressive information in data-driven forecasts enhances their competitiveness against expert forecasts.

Fig. 1 within Section 4.1 illustrates that certain product categories exhibit significantly shorter historical data compared to others. This observation holds significance as it aligns with the understanding that experts possess the capability to generate accurate forecasts when historical data is limited (Hyndman & Athanasopoulos, 2018). Consequently, this observation motivates the formulation of the third hypothesis:

[1] Based on: https://www.tradingview.com/markets/world-stocks/worlds-largest-companies/. Accessed 03 April 2024.

[2] wsts.org.

Table 1
Selection of related works in semiconductor sales and demand forecasting including response variables and utilized forecasting methods.

Authors	Methods	Response variable
Wang and Chen (2019)	ARIMA, VAR[a]	Quarterly sales time series from Taiwanese semicondoctor companies from Q1 2009 to Q4 2018
Kapur et al. (2019)	Technology diffusion	Sales and price data for DRAM, LCD monitors, and room air-conditioners
Chen and Chien (2018)	Technology diffusion	27 quarters of shipments for two technology generations of non-volatile memory products from a semiconductor company
Xu and Sharma (2017)	XGBoost, Linear model, RF[b], ARIMA, Ensemble	Weekly Intel CPU sales in 2012
Chien and Lin (2012)	Rolling grey forecasting method	Annual sales of companies in Hsinchu Science Park from 1983 to 2010
Aytac and Wu (2013)	Extended, Bayesian logistic growth model	Monthly sales data of about 5300 short life-cyle products from three semiconductor companies
Chien et al. (2013)	Technology diffusion	36 quarters of demand data for four technologies of a leading foundry from Hsinchu Science Park

[a] Vector autoregression.
[b] Random Forest.

(H3) Expert forecasts outperform data-driven forecasts particularly in the context of short time series.

Structure: The following two subsections summarize the related work (Section 2.2) and highlight the research gap and our contributions (Section 2.3). To investigate the hypotheses, the ML and statistical methods used are summarized in Section 3. The results are discussed in Section 5, with Section 5.1 addressing (H1), Section 5.2 examining (H2), Section 5.3 (H3), and finally Section 5.4, which offers insights of the results at a product category level. Section 6 completes this work with a brief discussion of the findings.

2.2. Related works

Despite the recognition of the semiconductor industry's importance in politics and business, little research was dedicated to data-driven forecasting the semiconductor market. A Scopus search (“market forecast” AND “semiconductor” AND (statistic* OR “machine learning”)) yields only one search result which discusses front-end drivers of changes in the semiconductor market (Nagao, 2019). However, this paper does not apply statistical or machine learning methods to generate forecasts directly. The small number of academic studies in the sector of the semiconductor industry was also noted by Aubry and Renou-Maissant (2014).

Likewise, a Scopus search for the “world semiconductor trade statistics” yields 11 results which largely cite WSTS as an authority for market data or forecasts. However, none of these use WSTS data as a basis for industry wide forecasts, nor do they assess the accuracy of WSTS' forecasts.

More related results can be found in the field of company-specific sales and demand forecasting and operational planning. Several contributions in this domain are summarized in Table 1. However, it should be noted that, while these studies often analyze different product or technology groups, the analyses are specific to one or several companies and usually do not encompass the whole semiconductor market with its different levels of granularity.

Another related field is the prediction of the semiconductor cycle. For a short summary, see Table 2. The semiconductor cycle, similarly to the economic cycle, describes cyclical fluctuations in semiconductor industry. These cycles are characterized by growth and contraction phases. Contributions in this domain usually focus on the overall semiconductor market, particularly WSTS global semiconductor sales, as the target variable. However, total semiconductor sales are usually considered in this scenario while a break down into finer product categories is often of interest for analysts and internal benchmarking.

2.3. Contributions

(1) **Evaluation of the Accuracy of Expert Semiconductor Market Forecasts:** Despite the frequent reliance on expert forecasts for semiconductor market forecasts, the accuracy of these forecasts has not, to our knowledge, been openly or systematically evaluated. This study addresses this gap by providing an evaluation of short-term forecasts through a comparative analysis with forecasts derived from several statistical and machine learning models.

(2) Comprehensive, Multi-Granularity Forecasting of the Semiconductor Market: While recent studies have focused on granular demand and sales forecasts for specific products or companies, they do not comprehensively cover the semiconductor market across all segments and as a whole, see Table 1 in Section 2.2 for an overview. Additionally, research into the semiconductor cycle often concentrates on high-level market trends and the identification of leading indicators, see Table 2 in Section 2.2. This study, however, provides a comprehensive analysis of the semiconductor market across various levels of granularity, utilizing the WSTS product categorization hierarchy. Forecasts are systematically generated for 110 product categories, covering the entire semiconductor market. This approach captures higher-granularity product groups as well as broader market trends, offering a comprehensive perspective that takes a step towards unifying these two areas of research.

(3) Guidance for Forecasting Practitioners through Model Performance Insights: In addition to forecasting across various levels of the semiconductor market, this study offers valuable guidance for forecasting practitioners. Out-of-sample error measures for each statistical and machine learning model are presented for all 110 product categories. These performance insights allow practitioners to assess the accuracy and applicability of different models in forecasting specific segments of the semiconductor market. To the best of our knowledge, providing such detailed model performance data for a wide range of product categories is a novel contribution.

3. Used data-driven methods

This section gives a brief introduction to the data-driven models used in the subsequent analysis. It starts with the description of traditional models based on statistical time series analysis (Section 2.1), continues with ML methods (Section 2.2), and concludes with a brief note on ensemble methods (Section 2.3).

The selection of the models presented here is partially influenced by their performance in the Makridakis Competitions (M-Competitions), which are renowned forecasting competitions conducted on diverse and realistic datasets (Hyndman, 2020). The results based on the M3

Table 2
Selection of related works in semiconductor cycle prediction including goals, response variables as well as utilized forecasting methods and indicators.

Authors	Goal	Methods	Response variable	Indicators[g]
Aubry and Renou-Maissant (2014)	Identification of best model for the semiconductor cycle prediction	ARMA, VAR[a], BVAR[b], VECM[c], MRSM[d], SF[e], ES[f]	WSTS global semiconductor sales (Jan 1991–Jun 2010)	SOX, NI, TI, BOOK
Aubry and Renou-Maissant (2013)	Prediction and description of the global semiconductor industry cycle	VECM[c]	WSTS global semiconductor sales (Jan 1991–Jun 2010)	SOX, NO, TI, IP, BOOK
Chow and Choy (2006)	Identification of leading indicators of semiconductor sales to predict the global semiconductor cycle	VAR[a], BVAR[b], BECM[c]	World semiconductor sales (Feb 1992–Jan 2005)	NASDAQ, NO, SI, PPI
Liu and Chyi (2006)	Prediction of the semiconductor cycle turning points	MRSM[d]	WSTS global semiconductor sales growth (Jan 1990–Aug 2003)	–
Liu (2005)	Identification of explanatory factors for the semiconductor cycle	VAR[a]	WSTS global semiconductor sales growth (Jan 1990–Dec 2001)	IP, FF, CS, SOX, NO, TI, UTL, EQO, CAP, PPI, SIP, VS

[a] Vector autoregression.
[b] Bayesian vector autoregression.
[c] Vector error correction models.
[d] Markov regime switching model.
[e] Spectral forecasting.
[f] Exponential smoothing.
[g] SOX: Philadelphia Semiconductor Index, IP: U.S. Industrial Production, FF: Federal Funds Rate, CS: U.S. Consumer Sentiment, NO: New Orders, TI: Total Inventories, UTL: Capacity Utilization, EQO: North American Equipment Orders, CAP: Capacity, PPI: Producer Price Index, SIP: Industry Production Index, VS: Value of Shipments, BOOK: Global Bookings of N.A. Semiconductor Equipment Producers, NASDAQ: NASDAQ Stock Index, SI: U.S. Shipments to Inventories Ratio.

competition are of particular interest: the dataset comprised 3003 individual time series with 14 to 126 observations featuring various levels of seasonality (Hyndman, 2020; Makridakis & Hibon, 2000).

Furthermore, this dataset has served as a benchmark for evaluating popular data-driven forecasting methods. Both Ahmed et al. (2010) and Makridakis et al. (2018) utilized a subset of the M3 dataset (consisting of 1045 time series) with a minimum length of 81 observations for their analyses.

3.1. Statistical models

Comparisons based on the M3-Competition data have generally favored statistical models over ML approaches (Hyndman, 2020; Makridakis et al., 2018). One suggested reason for this trend is the relatively short length of the time series involved. Unfortunately, this limitation is common in forecasting applications and reflects a realistic constraint. The time series examined in this paper, ranging from 92 to 392 monthly observations (further details in Section 4.1), are longer compared to those studied in the papers based on the M3-Competition, where lengths typically spanned from 81 to 126 observations. Nevertheless, these lengths are still comparable, especially when contrasted with datasets such as the M5 competition, which feature significantly longer time series, reaching up to approximately 2000 observations (Makridakis et al., 2022b).

3.1.1. SARIMA

The **s**easonal **a**uto**r**egressive **i**ntegrated **m**oving **a**verage (ARIMA) model is a traditional statistical time series model. Its advantages are its interpretability, its wide spread use and that many of its mathematical properties are well known (Brockwell & Davis, 2002). According to Brockwell and Davis (2002), a time series $X = \{X_t\}$ is said to be a SARIMA$(p, d, q) \times (P, D, Q)$ process with period s if

$$Y_t := (1 - B)^d (1 - B^s)^D X_t$$

is an causal ARMA process defined as

$$\phi(B)\Phi(B^s) Y_t = \theta(B)\Theta(B^s) Z_t,$$

where B is the back-shift operator defined as $BY_t = Y_{-1}$, $\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p$, $\Phi(z) = 1 - \Phi_1 z - \cdots - \Phi_P z^P$, $\theta(z) = 1 + \theta_1 z + \cdots + \theta_q z^q$, $\Theta(z) = 1 + \Theta_1 z + \cdots + \Theta_Q z^Q$, and $Z = \{Z_t\}$ being a white noise process.

Generally, an ARMA(p, q) process $Y = \{Y_t\}$ is characterized as

$$Y_t - \sum_{i=1}^{p} \phi_i Y_{t-i} = Z_t + \sum_{i=1}^{q} \theta_i Z_{t-i},$$

with $Z = \{Z_t\}$ being a white noise process. The left-hand side of this equation is the autoregressive part, while the right-hand side is the moving average part (moving average of the error process Z). More details on the SARIMA model can be found in Brockwell and Davis (2002).

These models are often used to describe and to generate data of a wide range of processes. But they can also be used as a predictive model when the parameters are estimated. To this end, we use the `auto.arima` function of the `forecast` library in R (Hyndman & Khandakar, 2008). A similar implementation for Python is available through the `StatsForecast` library (Garza et al., 2022).

The inclusion of the SARIMA model in this work is motivated by its ubiquity in time series analysis and its strong performance on the M3-Competition dataset (Makridakis et al., 2018).

3.1.2. Simple exponential smoothing

Exponential smoothing models range back to the 1950's (Gardner, 1985). Despite their simplicity, they often achieve high predictive performance (Hyndman, 2001; Satchell & Timmermann, 1995). Simple exponential smoothing (SES) only requires two quantities: the initial forecast $\hat{X}_0$ and the smoothing constant α. Consecutive forecasts can then be calculated via

$$\hat{X}_t = (1 - \alpha)\hat{X}_{t-1} + \alpha X_{t-1},$$

where $\hat{X}_t$ denotes the one-step forecast for X_t based upon the history up to X_{t-1}. An R implementation is available with the SES function of the `forecast` library (Hyndman & Khandakar, 2008). A similar implementation for Python is available through the `StatsForecast` library (Garza et al., 2022).

SES' simplicity, ease of implementation, and computational efficiency make it a popular forecasting tool for practitioners. Additionally, the model performed well on the M3 and M5 Competition datasets (Makridakis et al., 2018, 2022a).

3.1.3. Error, trend, and seasonality

Error, Trend, and Seasonality (ETS) approaches are a flexible class of exponential smoothing models that go beyond SES (see above). As their name suggests, they are capable of modeling time series with trends and seasonality (Hyndman & Athanasopoulos, 2018). ETS was the best performing model in the (Makridakis et al., 2018) comparison based on the M3-Competition data. It is also implemented as part of the `forecast` library in R (Hyndman & Khandakar, 2008). As for the previous two models, a similar implementation for Python is available through the `StatsForecast` library (Garza et al., 2022).

3.2. ML models

This subsection introduces the used ML models. Makridakis et al. (2018) found that ML methods performed worse than classical statistical models for relatively short time series - a finding that was confirmed by Cerqueira et al. (2022). This is particularly the case for artificial neural networks and deep learning models, which are well known to require large sample sizes to produce the desired results (Goodfellow et al., 2016). This was also verified by the NN3-Competition, which extended the M3-Competition to include neural network approaches (Crone et al., 2011; Hyndman, 2020). Therefore, following (Cerqueira et al., 2022), this paper does not discuss neural network models despite their considerable popularity in recent years. Likewise, boosting models are not included.

3.2.1. Random forest

Random forests (RF) are a bagging algorithm, a specific kind of ensemble learning, which combines the outputs of multiple decision trees (Breiman, 2001). Ahmed et al. and Makridakis et al. included Categorization and Regresssion Trees (CART), which generate single decision trees for regression or classification purposes (Breiman, 1984). However, since its introduction, RF has proven to be an incredibly versatile and successful model for both regression and classification (Biau & Scornet, 2016; Grinsztajn et al., 2022; Huang et al., 2020). We use the `ranger` implementation of this model as provided by Wright and Ziegler (2017). An alternative Python implementation is available through the `skranger` library. During each CV-step (see below), a grid search was conducted to tune the three hyper-parameters (Probst et al., 2019):

$$\begin{aligned} mtry &\in \{2,7,12,16,23\} \\ min.node.size &\in \{5,7,10\} \\ splitrule &\in \{variance, extratrees\}. \end{aligned}$$

3.2.2. Extremely randomized trees

Extremely Randomized Trees (ExtraTrees, also referred to as ET for brevity) is a model similar to Random Forest (see above). The difference lies in randomizing the splitting point and the feature to split on during training. It is computationally more efficient and promises greater accuracy on a range of problems (Geurts et al., 2006). The `ranger` library is also used for this model (Wright & Ziegler, 2017). However, in contrast to the RF model, the parameter `splitrule` remained fixed as `extratrees`. A Python implementation is available through `sklearn` (Pedregosa et al., 2011). During each CV-step, a grid search was conducted to tune the hyper-parameters (Probst et al., 2019):

$$\begin{aligned} mtry &\in \{2,7,12,16,23\} \\ min.node.size &\in \{5,7,10\}. \end{aligned}$$

3.2.3. Gaussian processes regression

Gaussian Processes Regression (GPR) is a probabilistic regression model incorporating Bayesian ideas: A prior distribution of possible regression functions is narrowed down as evidence (observed data points) are incorporated to yield a posterior distribution (Wang, 2023). It has been shown that GPR can be viewed as a limit of many artificial neural network designs and ARMA processes can be viewed as a Gaussian process under the right conditions (Williams & Rasmussen, 1995). Furthermore, due to their probabilistic nature, GPR easily provides uncertainty quantification for the forecasts in terms of prediction intervals. GPR showed a promising performance on the M3-Competition dataset (Ahmed et al., 2010; Makridakis et al., 2018). This analysis uses the GPR implementation of the `kernlab` R library (Karatzoglou et al., 2004). The model is also implemented for Python in the `sklearn` library (Pedregosa et al., 2011).

3.2.4. K-Nearest Neighbors

K-Nearest Neighbors (KNN) is a popular non-parametric model classification and regression model. It bases estimates on the K nearest neighbors in the covariate space (Cover & Hart, 1967; Fix & Hodges, 1989). K represents a hyper-parameter to be tuned. For regression, a mean of these K nearest neighbors is usually used as the predictor in regression. This model can be employed with a kernel. In the following, the kernel is automatically chosen. While KNN has not been among the best performing models on the M3-Competition data (Ahmed et al., 2010; Makridakis et al., 2018), it is nevertheless popular as simple a non-parametric model. Here, the implementation of the `kknn` library is used (Schliep & Hechenbichler, 2016). KNN is implemented for Python as the `NearestNeighbors` model in the `sklearn` library (Pedregosa et al., 2011). During each CV-step, a grid search was conducted to tune hyper-parameters among:

$$\begin{aligned} K &\in \{1,2,3,4,5,7,9\} \\ distance &\in \{L1, L2, L3\}, \end{aligned}$$

where the $L1, L2, L3$ distances are given by

$$d_{L^p}(x,y) := \|x-y\|_p$$

given that $\|x\|_p := \left(\sum_{i=1} x_i^p\right)^{1/p}$ is the $\mathcal{L}^p$ norm.

3.2.5. Support vector regression

Support Vector Regression (SVR) was proposed as an extension to the classical Support Vector Machine (SVM) for classification (Cortes & Vapnik, 1995) to tackle regression problems (Drucker et al., 1996). Instead minimizing all residuals, such as in ordinary least square regression, the distance of observations outside a margin of error to this margin of error (the ϵ-insensitive tube) are minimized. Analogous to SVMs, these observations are called support vectors, because the regressor only depends on these observations. To model non-linear dependencies, kernels can be used (Awad & Khanna, 2015). We use the radial kernel to add another non-linear method. The used model is implemented in the `kernlab` library for R or the `sklearn` library for Python (Karatzoglou et al., 2004; Pedregosa et al., 2011). A grid search was conducted to tune hyper-parameters among:

$$\begin{aligned} \sigma &\in \left\{\frac{1}{16}, \frac{1}{8}, \frac{1}{4}, \frac{1}{2}, 1\right\} \\ C &\in \left\{\frac{1}{4}, \frac{1}{2}, 1, 2, 4\right\}. \end{aligned}$$

3.3. Ensemble

Ensemble models consist of several individual models which are combined to produce a single output (Opitz & Maclin, 1999). In addition to the tree-based ensembles Random Forests and ExtraTrees, this paper also analyzes a simple ensemble of all the employed data-driven models by taking the median prediction.

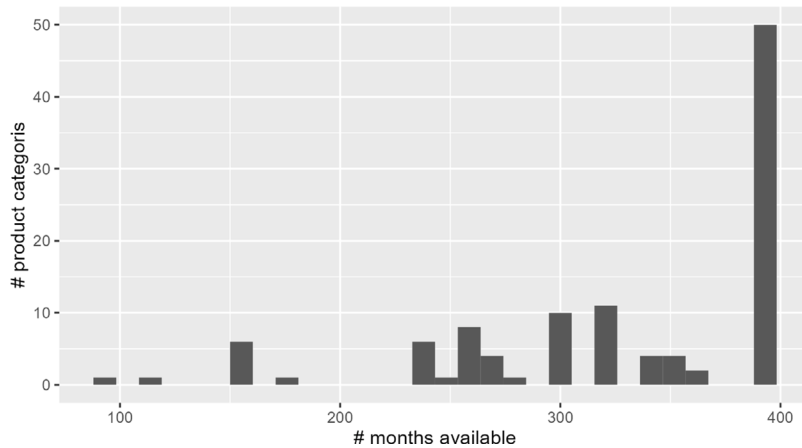

Fig. 1. Histogram of time series lengths.

4. Experimental setup

This section provides an overview over the data used in the analysis (Section 4.1) and the methodology behind the time series training and model evaluation (Section 4.2).

4.1. Data

This paper analyzes time series aggregated sales of 110 WSTS product categorizations, which were reported monthly and measured in USD. The data is accessible through a WSTS membership[3] or a subscription.[4] A major challenge was the consistency of the data:

Given the dynamic nature of the semiconductor market, product categorizations changed over time. The historical consistency of the current product categorizations was investigated and resolved dating back to Jan 2010 by merging the C7a and C7b product categorizations to C7 (Field Effect General Purpose Power Transistors), the P51 and P52 categorizations into P5 (Automotive and General Purpose MCU), and the L7a, L7b, L7c, L7d and L7f categorizations into L7a/b/c/d/f (Wireless Communication Total). The newest category (F10) dates back to Jan 2016. Fig. 1 provides an overview of the months of history of the categorizations as used in the analysis. These are the categories which are consistent until August 2023 and they provide a comprehensive overview of the semiconductor market. We refer to WSTS.org (2024) for an exact description of the market and each category.

Generally, the categories positioned higher in the hierarchy exhibit greater consistency. Fig. 2 illustrates the hierarchical structure of the product categorizations (starting with T99 as the highest aggregation). To effectively incorporate seasonal components, model fitting necessitated a minimum of 24 months (two seasonal cycles worth) of training data. The shortest time series comprised 92 monthly data points, thus leaving 68 months (or about 17.4% of the complete time series from 1991) as a test set for the first step of the rolling time series cross-validation (CV, see below). Hence, the first training set spanned all data from January 1991 (or whenever available) to December 2017. Consequently, the test set spanned the time frame from January 2018 to August 2023.

Official forecasts from WSTS were released quarterly from Q1 2016 to Q3 2023 (midway through the first forecasted quarter). Expert forecasts are consolidated during a global WSTS meeting twice a year – each May and November. WSTS additionally issues forecast updates semiannually, in February and August, derived from the preceding meeting's expert forecasts and updated algorithmically with new data reported for the prior quarter. For example, the forecast for Q2 2024 relies on the upcoming global WSTS meeting scheduled for May 20th–23rd, while the February 2024 forecast update drew upon forecasts from the November 2023 meeting and the reported data from Q4 2023. Thus, 11 expert forecasts and 12 updated forecasts were considered in the analyzed time span from January 2018 to August 2023.

4.2. Methodology

Training and evaluating the ML models: As described in the last subsection, each time series is split into first training and test sections. Time series with a longer available history consequently have more data points for training than shorter time series, i.e. newer product categories.

To obtain reliable forecast performance estimates for all of them, rolling time series cross-validation (CV) as in Hyndman and Athanasopoulos (2018) is performed on each time series and for each model. This is illustrated in Fig. 3, which also shows the most extreme training periods for the different time series (only 24 months for the first forecast of category F10 up to 390 months for the last forecast of T99).

Within each iteration, the training data is automatically transformed with the Box–Cox transformation (Box & COX, 1964) given by

$$X_t^{(\lambda)} = \begin{cases} (X_t^{\lambda} - 1)/\lambda & \lambda \neq 0, \\ \log(X_t), & \lambda = 0. \end{cases}$$

This transformation is incorporated and automatically estimated by the used libraries "forecast" and "caretForecast" (Akay, 2022; Hyndman & Khandakar, 2008). Applying the Box–Cox transformation is standard practise, especially when residual distributions are skewed, and when non-negative forecasts are desired (Hyndman & Athanasopoulos, 2018). The considered time series report aggregated sales in USD. Thus, there is unlimited upside potential whereas the lower bound is always zero since all time series are positive. Note that λ can always be chosen close to one if the transformation is not particularly helpful. Automatically applying it to all cases therefore does not hurt. This is also the default setting in the `forecast` library (Hyndman & Khandakar, 2008).

Additionally, hyper-parameters of all data-driven models introduced in Section 3 are optimized using the default grid search setting of the "caret" and "caretForecast" libraries (Akay, 2022; Kuhn, 2008) if no hyper-parameter optimization is conducted through the learning algorithm (one example where hyper-parameters are tuned automatically is the SARIMA model using the "auto.arima" function). After each training iteration, a forecast is generated up to three months in advance. These forecasts can be compared against the reported numbers, providing an estimate of the performance of the model, and the WSTS forecasts.

[3] https://www.wsts.org/61/membership.

[4] https://www.wsts.org/61/subscription.

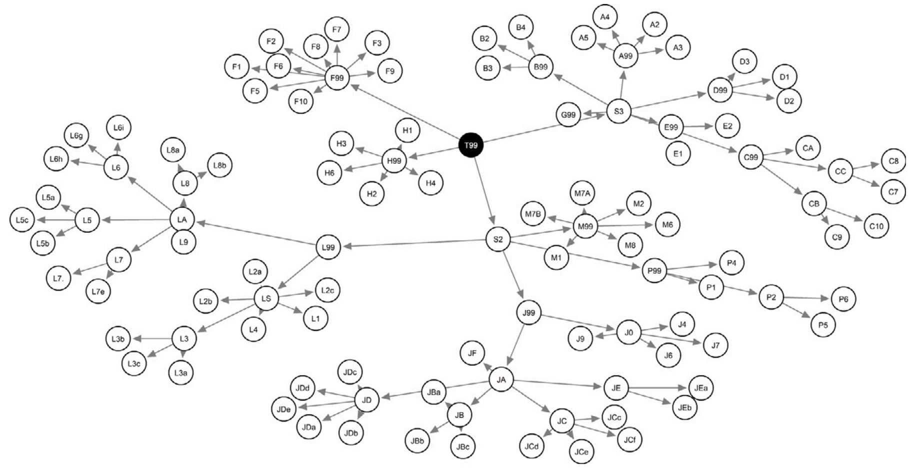

Fig. 2. WSTS product categorization hierarchy. The highest aggregation level is T99, the node colored in black with white print, slightly to the top of the center of the illustration. The arrows point to the subsumed product categories.

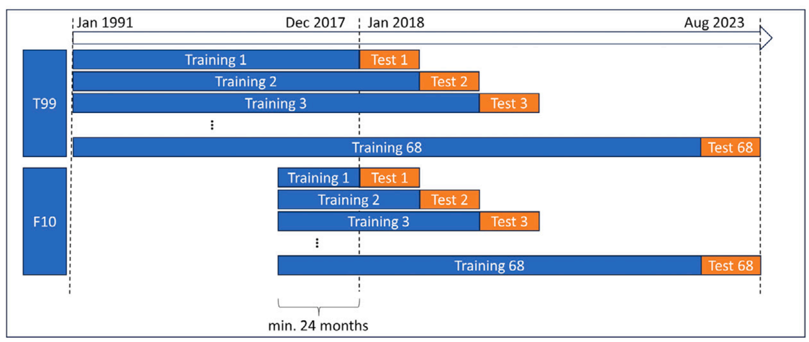

Fig. 3. Illustration of time series cross validation for two product categories with differing lengths: T99 has a much longer history of reported aggregate sales than F10.

Evaluation and comparison with WSTS' forecasts: The forecasts provided by WSTS are evaluated by type ("meeting" corresponding to the two WSTS expert forecast per year, "alg. update" corresponding to the two WSTS algorithmic forecasts per year and "overall" for all four WSTS forecasts per year) and compared to the corresponding data-driven forecasts. The first comparison evaluates all forecasts with a forecasting horizon of $h = 3$ months. This corresponds to the information timestamp available to WSTS' updating algorithm and the industry experts.

At the same time, it is safe to assume that the industry experts have access to the sales data (among many more) of the first month of each quarter when the meeting convenes (the meeting is held in about the middle of the first quarter to be forecasted). Additionally, the forecasts are disclosed at the same time as the first month's results are published. Hence, the information time stamp of the forecast and the first month's results is the same. Utilizing this information reduces the required forecast horizon to $h = 2$ months. This is analyzed in a second step to investigate whether using more of the available information might boost the forecasting accuracy.

Similar to Hyndman and Koehler (2006) and Pauly and Kuhlmann (2023), the forecasting accuracy is evaluated using the mean squared error (MSE), mean absolute percentage error (MAPE), and the mean absolute error (MAE). These are given by the following equations

$$MSE = \frac{1}{T}\sum_{t=1}^{T}(\hat{X}_t - X_t)^2$$

$$MAPE = \frac{1}{T}\sum_{t=1}^{T}\left|\frac{\hat{X}_t - X_t}{X_t}\right|$$

$$MAE = \frac{1}{T}\sum_{t=1}^{T}|\hat{X}_t - X_t|,$$

where again $\hat{X}_t$ is the one-step forecast for the tth observation X_t of the test set and $T = 68$ is the number of evaluated forecasts. We note that the MAPE is applicable as the values of all time series are positive.

Table 3 provides a brief overview over the key methods which are employed in this study and provides references to motivating studies and deeper methodological discussions.

5. Results

As discussed in Section 4.2, first, the quarterly forecasts ($h = 3$ months) are discussed in Section 5.1 to examine the first research hypothesis (H1) that expert forecasts exhibit higher accuracy compared to data-driven forecasts. This is followed by a comparison with the model performance when an additional month of available information ($h = 2$ months) is incorporated in Section 5.2, addressing the second research hypothesis (H2). Lastly, the results are contrasted by time series length

5.1. Quarterly forecast performance

WSTS' forecasts are provided on a quarterly basis. Each quarter, the previous quarter's numbers are known when the forecasts are compiled. Therefore, as a first step, the data-driven models are compared against WSTS' forecasts on a forecasting horizon of $h = 3$ months, i.e. one quarter.

Table 4 presents the average performance of data-driven forecasts across 110 product categories, relative to the average performance of forecasts provided by the World Semiconductor Trade Statistics (WSTS). Hence, the first data column (for WSTS) always reads 1. Each row represents a different error measure, organized according to

Table 3
Overview of used methods: Purposes and motivations for chosen techniques, with references.

Methods	Purpose and motivation	Authors
SARIMA, SES, ETS, RF, ET, GPR, KNN, SVR	Purpose: generating point forecasts. Motivation: the selection of these models was largely motivated by studies based on the M3-Competition dataset. These models achieved outstanding performance on time series similar to the ones in this study.	Ahmed et al. (2010) and Makridakis et al. (2018)
Cross-Validation	Purpose: estimating out-of-sample performance estimates. Motivation: The application of time series cross validation is a standard procedure.	Cerqueira et al. (2020)
MSE, MAPE, MAE	Purpose: measures for prediction accuracy. Motivation: these are standard measures commonly employed to assess the accuracy of (time series) models with continuous targets	Hyndman and Koehler (2006)

Table 4
Average performance of the data-driven forecasts across all 110 product categories and relative to the World Semiconductor Trade Statistics' (WSTS). Each row refers to a different error measure, sorted by WSTS' forecast type: algorithmic update, meeting (expert forecast), and overall. Lower values are preferable. The best value per row is bold and italic.

		WSTS	SARIMA	ETS	ET	GPR	KNN	RF	SES	SVM	Ensemble
Alg. Update	MSE	1.00	***0.34***	0.55	1.55	0.37	4.47	1.49	0.94	3.20	0.64
	MAPE	***1.00***	1.01	1.01	1.20	1.01	1.71	1.18	1.06	1.69	1.04
	MAE	1.00	***0.75***	0.81	1.34	0.77	2.14	1.29	1.08	1.94	0.95
Meeting	MSE	***1.00***	1.61	1.57	2.78	1.33	7.23	2.49	1.79	6.91	1.77
	MAPE	***1.00***	1.14	1.09	1.32	1.13	1.76	1.28	1.10	1.82	1.14
	MAE	***1.00***	1.22	1.22	1.53	1.17	2.32	1.47	1.24	2.32	1.28
Overall	MSE	1.00	0.73	0.86	1.93	***0.66***	5.31	1.79	1.20	4.33	0.98
	MAPE	***1.00***	1.08	1.05	1.26	1.07	1.73	1.23	1.08	1.75	1.09
	MAE	1.00	0.97	1.00	1.42	***0.95***	2.22	1.37	1.15	2.11	1.10

WSTS' forecast types: algorithmic update, meeting (expert forecast), and overall. Lower values are preferable in all cases.

For WSTS' "Algorithmic Update" forecasts, three error measures are reported: Mean Squared Error (MSE), Mean Absolute Percentage Error (MAPE), and Mean Absolute Error (MAE) – see Section 4.2. Among these measures, the data-driven forecasts outperform WSTS in terms of MSE and MAE, with the best-performing model indicated by bold and italic formatting. The SARIMA model exhibits the lowest MSE (0.34 relative to WSTS) and MAE (0.75), suggesting superior performance in this category. Almost as good was the GPR model (relative MSE of 0.37 and MAE of 0.77). In term of MAPE, WSTS performed slightly better than several data driven forecast methods: SARIMA, ETS, and GPR each scored 1% worse. Overall, these results indicate potential for improvement of WSTS' algorithmic update protocol.

Similarly, for the "Meeting" forecasts, the same three error measures are provided. Consistent with the first research hypothesis (H1), industry experts demonstrated superior performance across all three error measures. However, among the data-driven models, the best performers were GPR, exhibiting a 33% increase in MSE compared to expert forecasts, while ETS and SES displayed 9% and 10% higher MAPE values respectively. Additionally, GPR, again, was the best performing data-driven model in terms of MAE, showing a 17% higher MAE relative to the expert forecasts.

Table 4 concludes with the "Overall" forecasts, providing the average performance of all models relative to the average performance of the combined algorithmic and expert forecasts by WSTS (from 2 meeting and 2 algorithmic forecast per year). Once more, the GPR model emerges as the top-performing data-driven model, showing a 34% improvement in MSE and a 5% enhancement in MAE compared to WSTS. However, WSTS outperforms all models in terms of MAPE, with ETS, the top-performing data-driven model, recording a 5% higher error than WSTS.

Table 5 provides insights into the mean ranks of the various forecasting models across all 110 product categories, structured similarly to Table 4. Contrary to Table 4, the columns in Table 7 are not standardized by the WSTS column. A rank of 1 indicates the model performed the best for that product category based on the respective error measure, thus lower mean ranks are preferred.

In general, the observations from Table 7 are similar to those from Table 6. The forecasts provided by WSTS demonstrate excellence across most error metrics and scenarios. However, there is an exception with WSTS' Algorithmic Update forecasts, where SARIMA and GPR achieved slightly lower average ranks (3.8 vs. 3.9 for WSTS). Considering that the average MSE for the forecasts based on the SARIMA model was 66% lower than the MSE of WSTS' algorithmic updates, it suggests SARIMA's strong performance is primarily driven by specific product categories where it outperforms WSTS (Section 5.4 contains a detailed discussion on individual product categories). Although several data-driven models exhibited a superior average error in terms of MSE and MAE for the overall forecasts (for example, SARIMA and GPR) this superiority does not necessarily translate to lower average ranks, highlighting WSTS' robust performance across product categories.

Comparing only the data-driven forecasts, GPR emerges as the most accurate data-driven model in the Algorithmic Update and Overall scenarios, while Simple Exponential Smoothing (SES) demonstrates the best performance among data-driven models during quarters where WSTS provided expert forecasts (Meeting), particularly in terms of MSE and MAE. ETS achieved a slightly lower average rank in terms of Mean Absolute Percentage Error (MAPE).

One plausible explanation for the strong performance of the industry experts is their access insider information. In contrast, the data-driven models rely solely on aggregated historical sales data from the specific product category being forecasted. Another factor to consider is the timing of the WSTS meetings where expert forecasts are consolidated. These meetings typically occur in the middle of the forecasted quarter, implying that experts are likely aware of their first-month figures. Moreover, these figures are available simultaneously with the official forecast release. In contrast, the bi-quarterly algorithmic updates do not integrate this information, though they could potentially benefit from it. Consequently, data-driven models that incorporate this timely information, necessitating forecasts with a horizon of only $h = 2$ months, are evaluated in the subsequent subsection.

5.2. Forecast performance with additional information

Considering the timing of the WSTS meetings in the middle of the quarter, it is reasonable to assume that industry experts take into account the numbers and internal information pertaining to the first month when formulating their forecasts for the first quarter. Moreover, the consolidated results of the first month are released simultaneously with the forecasts by WSTS. Therefore, utilizing all available information for forecasts seems appropriate. This approach allows for an

Table 5
Mean ranks of the forecasts with horizon h = 3 months across all 110 product categories. Each row refers to a different error measure, sorted by WSTS' forecast type: algorithmic update, meeting (expert forecast), and overall. Lower values are preferable. The best value per row is bold and italic.

		WSTS	ARIMA	ETS	ET	GPR	KNN	RF	SES	SVM	Ensemble
Alg. Update	MSE	3.9	***3.8***	4.2	6.6	***3.8***	8.7	6.2	4.4	8.8	4.5
	MAPE	***3.6***	4.1	4.1	6.6	3.9	8.8	6.1	4.5	8.7	4.5
	MAE	***3.6***	4.1	4.2	6.6	3.9	8.8	6.1	4.4	8.9	4.5
Meeting	MSE	***2.9***	4.5	3.9	7.0	4.1	8.9	6.4	3.8	8.9	4.6
	MAPE	***2.8***	4.3	4.0	7.0	4.2	8.9	6.4	4.1	8.8	4.5
	MAE	***2.9***	4.3	4.0	7.2	4.2	8.9	6.4	3.8	8.8	4.5
Overall	MSE	***3.2***	4.4	4.2	6.9	3.7	9.0	6.3	4.1	8.9	4.4
	MAPE	***2.8***	4.1	3.9	7.0	3.9	9.1	6.5	4.2	8.9	4.5
	MAE	***2.9***	4.2	4.0	7.0	4.0	9.1	6.4	4.1	8.8	4.5

Table 6
Average performance of the data-driven forecasts with horizon $h = 2$ months across all 110 product categories and relative to WSTS. Each row refers to a different error measure, sorted by WSTS' forecast type: algorithmic update, meeting (expert forecast), and overall. Lower values are preferable. The best value per row is bold and italic.

		WSTS	SARIMA	ETS	ET	GPR	KNN	RF	SES	SVM	Ensemble
Alg. Update	MSE	1.00	***0.10***	0.18	0.52	***0.10***	1.31	0.49	0.46	1.29	0.32
	MAPE	1.00	0.62	0.62	0.76	***0.61***	1.09	0.75	0.68	1.08	0.65
	MAE	1.00	***0.42***	0.51	0.81	0.43	1.31	0.79	0.77	1.24	0.63
Meeting	MSE	1.00	0.68	0.66	1.41	***0.64***	3.65	1.25	0.96	3.18	0.91
	MAPE	1.00	***0.61***	0.62	0.82	0.65	1.15	0.79	0.65	1.15	0.68
	MAE	1.00	***0.72***	0.73	1.05	0.75	1.54	1.00	0.88	1.46	0.85
Overall	MSE	1.00	0.28	0.32	0.80	***0.27***	2.03	0.72	0.61	1.86	0.50
	MAPE	1.00	***0.61***	0.62	0.79	0.63	1.12	0.77	0.66	1.12	0.66
	MAE	1.00	***0.56***	0.61	0.92	0.58	1.41	0.89	0.82	1.34	0.73

additional month of data to forecast the quarterly result, rendering a forecasting horizon of $h = 2$ months sufficient. Considering that one out of three months' numbers do not require estimation, a plausible anticipation would be to observe approximately a 33% reduction in MSE (assuming an unbiased estimator and uncorrelated errors). Such a decrease could already position several data-driven methods as competitive alternatives to WSTS, which is examined in this section.

The average performance of these forecasts relative to WSTS' is presented in Table 6, structured equivalently to Table 4. Table 6 presents the average performance of data-driven forecasts with a horizon of $h = 2$ months across 110 product categories, relative to the forecasts provided by WSTS. Each row represents a different error measure, categorized by WSTS' forecast types: algorithmic update, meeting (expert forecast), and overall. As before, lower values indicate better performance, with the best value per row highlighted in bold and italic.

Contrasting these results with those presented in previous table (Table 4), several notable differences emerge. Firstly, upon a cursory glance of the results, it becomes evident that the data-driven approaches have exhibited markedly superior performance in this context. Whereas WSTS' expert forecasts (Meeting) previously outperformed the data-driven forecasts in terms of MSE, MAPE, and MAE, the tables have now turned, with the data-driven forecasts consistently showcasing superior forecast accuracy in the new scenario (with the exception of SVM and KNN for all as well as ET and RF for the Meeting MSE and MAE comparisons). Specifically, the GPR model achieved a 36% lower MSE than WSTS' expert forecasts, followed by ETS (34% lower) and SARIMA (32% lower). In terms of MAPE, SARIMA outperformed WSTS' experts by 39%, followed by ETS (38% lower), and GPR and SES (both 35% lower). Additionally, SARIMA demonstrated the best performance in terms of MAE (28% lower than WSTS), trailed by ETS (27% lower) and GPR (25% lower). Even the simple ensemble, which even incorporates the forecasts of the worse performing models, surpassed WSTS' experts by 9% in MSE, 32% in MAPE, and 15% in MAE. This suggests that data-driven models incorporating the latest available information are highly effective in forecasting outcomes within a shorter horizon. Furthermore, WSTS' expert forecasts attained superior average performance in terms of MAPE across all three forecast types with a forecasting horizon of $h = 3$ months. However, with a reduced horizon of $h = 2$ months, the top-performing data-driven forecasts now outperform WSTS by up to 39%. ARIMA and ETS emerged as the top performers, closely followed by GPR.

Secondly, concerning the algorithmic updates, according to Table 6, SARIMA and GPR once again emerge as one of the top-performing models. In MSE, both SARIMA and GPR achieved errors 90% lower than WSTS. Additionally, in terms of MAE, SARIMA attained a 58% lower error, closely followed by GPR with a 57% reduction. While in terms of MAPE, where WSTS previously outperformed data-driven models, GPR achieved a 39% lower error, with SARIMA and ETS achieving a 38% lower MAPE.

Table 7 offers insights into the mean ranks of the various forecasting models across all 110 product categories, organized similarly to Tables 4 and 6. It is important to note that, in contrast to Tables 4 and 6, the columns in Table 7 are not standardized by the WSTS column. A rank of 1 indicates the model performed the best for that product category based on the respective error measure, thus lower mean ranks are preferred. Overall, the observations from Table 7 parallel those from Table 6. Models such as SARIMA, ETS, and GPR consistently garnered high ranks. Notably, most data-driven methods outperformed WSTS' expert forecasts, except for KNN and SVM, which exhibited poorer performance.

Finally, Fig. C.4 in the Appendix illustrates the frequency of the best-performing forecasts across 110 WSTS product categories, with colors indicating performance metrics: red for Mean Squared Error (MSE), blue for Mean Absolute Percentage Error (MAPE), and green for Mean Absolute Error (MAE). The left panel reflect the $h = 3$ months forecast while $h = 2$ is presented on the right side. It is evident that WSTS forecasts rarely emerge as the top performers within any given product category for $h = 2$. While WSTS' expert forecasts generally outperform its algorithmic updates, data-driven models consistently outshine both. When comparing against WSTS' expert forecasts (second row), SARIMA emerges as the best performer between 20% (MSE) and 30% (MAE) of the time, followed by ETS between 19% (MSE) and 27% (MAPE). Additionally, GPR and SES each excel between 13% (MAE and MAPE) and 17% and 18% (MSE), respectively. In contrast, WSTS' expert forecasts demonstrate the best performance between 7% (MAE and MAPE) and 13% (MSE) of the time.

Table 7
Mean ranks of the forecasts with horizon h = 2 months across all 110 product categories. Each row refers to a different error measure, sorted by WSTS' forecast type: algorithmic update, meeting (expert forecast), and overall. Lower values are preferable. The best value per row is bold and italic.

		WSTS	SARIMA	ETS	ET	GPR	KNN	RF	SES	SVM	Ensemble
Alg. Update	MSE	8.0	3.6	3.7	5.8	***3.3***	8.4	5.7	3.9	8.7	3.7
	MAPE	8.2	***3.4***	3.7	5.8	3.6	8.6	5.6	4.1	8.4	3.5
	MAE	8.0	***3.4***	3.8	5.9	3.5	8.6	5.7	4.1	8.5	3.6
Meeting	MSE	6.9	3.5	***3.3***	6.5	3.7	8.9	5.9	3.5	8.5	4.2
	MAPE	7.2	***2.9***	3.4	6.7	3.4	8.8	6.2	3.7	8.6	4.2
	MAE	7.0	***2.9***	3.5	6.8	3.5	8.8	6.2	3.7	8.5	4.1
Overall	MSE	7.8	***3.3***	3.4	6.3	***3.3***	8.8	5.9	3.6	8.7	3.8
	MAPE	8.0	***2.9***	3.2	6.4	3.3	9.0	5.9	3.8	8.7	3.7
	MAE	7.8	***2.9***	3.3	6.4	3.4	9.0	6.0	3.8	8.6	3.8

Table 8
Overview of the time series lengths by category.

	Count	Av. Length	Min. Length	Max. Length
Long	50	392	392	392
Medium	31	324	296	359
Short	29	222	92	284
All	110	328	92	392

These findings corroborate the trends observed in the preceding analyses of this section, suggesting the reliability and effectiveness of certain data-driven models over expert forecasts in near-term forecasting scenarios when additional available information is incorporated, supporting the second research hypothesis (H2).

5.3. Impact of time series length on forecast performance

This section aims to investigate the possibility that expert forecasts perform better on shorter time series due to the limited data available for training data-driven models (Hypothesis (H3)). To explore this third hypothesis, the 110 product categories are divided into long, medium, and short categories based on the length of available observations. An overview of this categorization is presented in Table 8.

In total, the time series had an average length of 328 monthly observations. Among these, the 50 product categories with available monthly data points for the entire examined time period were classified as "long" time series. The "medium" category comprised 31 time series with observations ranging from 296 to 359, averaging 324 available data points. The remaining 29 product categories were designated as "short" time series, with observations ranging from 92 to 284, averaging 222 available data points.

Similar to Table 6, Table 9 provides an overview of the performance metrics Mean Squared Error (MSE), Mean Absolute Percentage Error (MAPE), and Mean Absolute Error (MAE) for the 2-month forecasts. These metrics are segmented based on time series lengths: Short, Medium, and Long, as delineated in Table 8, and further categorized by WSTS' forecast types: algorithmic update, meeting (expert forecast), and overall.

For short time series, GPR demonstrated superior performance in terms of both MSE and MAPE, showcasing reductions of 76% and 43% respectively compared to WSTS' combined forecasts. Regarding MAPE, ETS exhibited the lowest error rate (37% lower than WSTS), trailed by SES (36% lower), SARIMA (34% lower), and GPR (33% lower). When contrasting the data-driven forecasts with WSTS' expert forecasts, ETS emerged as the top-performing model, tying with GPR in terms of MSE (both 43% lower than expert forecasts) and outperforming WSTS' experts by 32% in terms of MAE. However, SES marginally outperformed ETS in terms of MAPE, showing reductions of 35% and 34% respectively. When comparing data-driven forecasts solely to algorithmic forecasts, GPR emerged as the most accurate model across all metrics, boasting reductions of 93% in MSE, 43% in MAPE, and 58% in MAE. This explains the robust performance of GPR relative to WSTS' combined forecasts.

It is also noteworthy to highlight the disparity in the performance of data-driven methods for short time series depending on whether quarters with expert forecasts or algorithmic updates were used for the benchmarks. Given that these are derived from the same time series, a consistent ranking of data-driven methods might have been expected. The best-performing models for medium and long time series remain largely consistent, regardless of whether they are evaluated using algorithmic forecasts or expert forecasts. Even in cases where there are differences, the margin is minimal.

In the Appendix, Table A.10 outlines the mean ranks of the analyzed models, once again categorized by Mean Squared Error (MSE), Mean Average Percentage Error (MAPE), and Mean Absolute Error (MAE), and segmented by time series lengths. These rankings are further delineated by WSTS forecast type, mirroring the structure of Table 9.

Consistent with the findings in Table 9, GPR demonstrated the most favorable performance in terms of mean ranks across MSE (2.8), MAPE (2.8), and MAE (2.9) for short time series when compared to WSTS' algorithmic forecasts. However, while ETS excelled in average MSE and MAE, and SES in MAPE, this outcome is reversed when considering mean ranks: SES exhibited the best performance in terms of mean ranks for MSE and MAE, while ETS performed best for MAPE. It is worth noting that the differences between them are minimal in each case. Furthermore, ETS attained the lowest mean rank when all quarters were taken into account (overall), suggesting that exponential smoothing models perform well when data is limited. Furthermore, the high accuracy of the data-driven forecasts compared to WSTS' expert forecasts is evidence contrary to the third research hypothesis (H3) that expert forecasts would outperform data-driven forecast in the context of short time series.

In fact, a slight increase in error relative to the expert forecasts is observed for the best forecasts when medium-length time series are considered in terms of MSE (0.58 vs. 0.57). Likewise, the data-driven forecasts fared slightly worse in terms of MSE (0.64 vs. 0.57) and MAE (0.72 vs. 0.68) for long time series. Given the additional information which was utilized in training the models, the opposite might have been expected. This was the case when the forecasts were evaluated by MAPE (0.57 medium and long time series vs. 0.65 for short ones).

For medium time series, SARIMA generated the most reliable forecasts in terms of relative average error measures – 42% more accurate than the experts polled by WSTS in terms of MSE, 43% more accurate in terms of MAPE, and 36% more accurate in terms of MAE, followed by ETS. This is also reflected in the mean ranks of the SARIMA forecasts (2.9, 2.2, and 2.8 vs. 7.1, 7.8, and 7.5 respectively). Similarly, the SARIMA forecasts demonstrated superior performance in comparison to the algorithmic updates and when evaluated overall.

In the case of the long time series, the most accurate forecasts resulted from the GPR and SARIMA models (Table 9). GPR resulted in 36% lower MSE and a 43% lower MAPE compared to the experts. SARIMA excelled in terms of MAE – 38% lower than WSTS. A similar

Table 9
Average performance of the data-driven forecasts with horizon $h = 2$ months across the different time series lengths and relative to WSTS. Each row refers to a different error measure, sorted by WSTS' forecast type: algorithmic update, meeting (expert forecast), and overall. Lower values are preferable. The best value per row is bold and italic.

			WSTS	SARIMA	ETS	ET	GPR	KNN	RF	SES	SVM	Ensemble
Short	Alg. Update	MSE	1.00	0.12	0.11	0.25	***0.07***	0.58	0.22	0.29	0.60	0.14
		MAPE	1.00	0.63	0.61	0.73	***0.57***	1.08	0.72	0.63	1.07	0.63
		MAE	1.00	0.49	0.49	0.70	***0.42***	1.09	0.68	0.67	1.11	0.56
	Meeting	MSE	1.00	0.74	***0.57***	0.86	***0.57***	1.35	0.77	0.65	1.26	0.66
		MAPE	1.00	0.69	0.66	0.88	0.77	1.23	0.85	***0.65***	1.24	0.74
		MAE	1.00	0.77	***0.68***	0.89	0.74	1.23	0.85	0.72	1.20	0.76
	Overall	MSE	1.00	0.33	0.27	0.46	***0.24***	0.85	0.41	0.42	0.83	0.32
		MAPE	1.00	0.66	***0.63***	0.80	0.67	1.15	0.78	0.64	1.16	0.68
		MAE	1.00	0.63	0.58	0.79	***0.57***	1.16	0.76	0.70	1.16	0.65
Medium	Alg. Update	MSE	1.00	***0.14***	0.17	0.42	***0.14***	1.33	0.39	0.27	1.23	0.27
		MAPE	1.00	***0.53***	0.57	0.76	0.54	1.08	0.75	0.61	1.06	0.63
		MAE	1.00	***0.45***	0.48	0.81	***0.45***	1.33	0.79	0.63	1.32	0.64
	Meeting	MSE	1.00	***0.58***	0.61	1.05	0.66	2.17	0.93	0.73	2.21	0.78
		MAPE	1.00	***0.57***	0.65	0.83	0.64	1.20	0.79	0.66	1.15	0.67
		MAE	1.00	***0.64***	0.67	0.97	0.72	1.45	0.92	0.76	1.48	0.81
	Overall	MSE	1.00	***0.31***	0.34	0.66	0.34	1.66	0.60	0.45	1.61	0.47
		MAPE	1.00	***0.55***	0.61	0.79	0.59	1.14	0.77	0.63	1.10	0.65
		MAE	1.00	***0.54***	0.57	0.89	0.58	1.39	0.85	0.69	1.39	0.72
Long	Alg. Update	MSE	1.00	***0.10***	0.18	0.54	***0.10***	1.34	0.50	0.47	1.31	0.32
		MAPE	1.00	***0.67***	0.68	0.78	0.68	1.09	0.77	0.76	1.11	0.68
		MAE	1.00	***0.41***	0.52	0.83	0.43	1.34	0.81	0.80	1.25	0.64
	Meeting	MSE	1.00	0.68	0.66	1.45	***0.64***	3.80	1.28	0.98	3.29	0.93
		MAPE	1.00	0.58	0.58	0.78	***0.57***	1.07	0.75	0.64	1.09	0.64
		MAE	1.00	***0.72***	0.75	1.08	0.76	1.60	1.03	0.93	1.50	0.88
	Overall	MSE	1.00	***0.27***	0.33	0.81	***0.27***	2.08	0.74	0.63	1.91	0.51
		MAPE	1.00	***0.62***	0.63	0.78	***0.62***	1.08	0.76	0.70	1.10	0.66
		MAE	1.00	***0.55***	0.63	0.94	0.58	1.46	0.91	0.86	1.36	0.75

picture arises when mean ranks (Table A.10) are considered. An exception is the algorithmic update category, where forecasts based on the ensemble method achieved the lowest mean ranks.

Fig. C.5 in the Appendix illustrates the frequency of the best-performing forecasts with a forecast horizon of $h = 2$ months. The chart is divided into 3 columns corresponding to the time series lengths: short, medium, and long. Rows are arranged by WSTS forecast type (algorithmic update, meeting (expert), and overall). The colors differentiate between Mean Squared Error (MSE) in green, Mean Absolute Percentage Error (MAPE) in blue, and Mean Absolute Error (MAE) in red. In concordance with the analysis of Tables 9 and A.10, GPR shows the highest frequency of top model performance for short sime series when compared to the algorithmic updates. Overall and for the expert forecasts ETS and SES had the highest frequencies of highest accuracy. For the medium and long time series, SARIMA, GPR, and the exponential smoothing models excelled most often.

Additional details pertaining to the overall performance of the various data-driven models for each product category categorized by time series lengths is available in Table B.12.

5.4. Additional results

In addition the aggregated results presented in Sections 5.1–5.3, this Section elaborates the results on a product category level.

Tables B.11, B.12, and B.13 in the Appendix, provide the Root Mean Squared Errors (RMSE), which is the square root of the Mean Squared Error (MSE) described in Section 4.2 and was chosen for readability here, mean absolute percentage errors (MAPE), and Mean Absolute Errors (MAE) respectively for various forecasting models across 110 different product categories. Each time, these are organized by time series lengths, consistent with the categorization in Section 5.3. The error measures for the forecasts with horizon $h = 3$ months are presented in the first half of each table and those of forecasts with horizon $h = 2$ months are presented in the second halves. Lower RMSE, MAPE, and MAE values indicate higher forecasting accuracy, with the best performing model per product category printed in bold and italic in each row of each table.

In the 3-month forecast category, the results are consistent with those discussed in Section 5.1. As can be seen in Tables B.11 and B.13, data-driven methods, particularly GPR and SARIMA, are overall able to outperform WSTS in terms of RMSE and MAE. Table B.11 reveals that this is in large part due to the strong performance of the GPR ans SARIMA forecasts in terms of RMSE for a few product categories such as M99 (37.72 for GPR vs 46.77 for WSTS), T99 (56.83 for GPR vs. 71.85 for WSTS), and S2 (49.17 for SARIMA vs 70.79 for WSTS). Similarly, scrutinizing Table B.13 indicates that among all forecasts, WSTS' was the most reliable for most product categories in terms of MAE. But GPR, SARIMA, and ETS produced forecasts which excelled for specific product categories, such as T99 (53.76 for WSTS vs. 46.24 for SARIMA, 47.35 for GPR, and 50.85 for ETS), resulting in a higher average performance for long time series. In terms of MAPE, Table B.12 illustrates WSTS' strong performance on the $h = 3$ month horizon across all time series lengths. Nevertheless, upon scrutinizing Table B.12, it becomes apparent that data-driven models outperform WSTS' combined expert and algorithmic update forecasts in terms of MAPE for specific product categories. For instance, in categories such as J99 and L8a, the SARIMA model produces the lowest MAPE, demonstrating its effectiveness in forecasting these particular products. In some categories like L1, where the time series might exhibit unique patterns or complexities, traditional statistical models such as SARIMA and ETS perform inadequately compared to specific machine learning models, in this case: RF.

Consistent with the observations in Section 5.2, these results are reversed when the forecasts with a horizon of $h = 2$ months are considered. In this scenario, data-driven forecasts excelled across the vast majority of product categories in terms of RMSE (Table B.11), MAPE (Table B.12), and MAE (Table B.13). Nevertheless, WSTS' forecasts were superior for some select product categories with long histories in terms of RMSE, such as C7 (1.18 vs. 1.32 for the best performing data-driven model: GPR), CC (1.29 vs. 1.47 for the best performing data-driven model: GPR), and S3 (2.77 vs. 2.94 for the best performing data-driven model: SES).

Moreover, despite the dominance of the SARIMA, GPR, and ETS among the data-driven models, it is noteworthy that other models with poorer average performance still yielded in strong forecasts for select product categories: ET was amongst the top performing models for the A5 product category in terms of RMSE and MAPE, and SVM performed best for the JCd product category in terms of RMSE and MAE. This finding is underlined by Fig. C.4 located in Appendix. The bar chart visualizes the frequency of the best-performing forecasts across 110 WSTS product categories. Differentiated by colors, green signifies the best performance based on Mean Squared Error (MSE), blue represents Mean Absolute Percentage Error (MAPE), and red indicates Mean Absolute Error (MAE). The chart is divided into two sections: the left side displays outcomes for forecasts with a horizon of $h = 3$ months, while the right side portrays forecasts with a horizon of $h = 2$ months (to be discussed in the subsequent section). Rows are arranged by algorithmic update, expert forecasts (meeting), and overall performance.

Finally, Fig. C.4 in the Appendix illustrates the frequency of the best-performing forecasts across 110 WSTS product categories, with colors indicating performance metrics: red for Mean Squared Error (MSE), blue for Mean Absolute Percentage Error (MAPE), and green for Mean Absolute Error (MAE). The left panel reflect the $h = 3$ months forecast while $h = 2$ is presented on the right side.

For $h = 3$ months the superior performance of the WSTS models is evident, particularly their Meeting and Overall forecasts. Depending on the error measure, WSTS' experts outperformed all data driven models between 46% and 48% of all product categories. The most frequent best performing data driven models were SARIMA, GPR and the exponential smoothing models, ETS and SES. In aggregate, the forecasts provided by WSTS (Overall row) achieved a top performance in 38% to 39% of the cases. In contrast, the playing field was more even when only the algorithmic updates are considered. WSTS achieved a top performance for 21% to 23% of all product categories, followed by SARIMA with 17% to 23% and GPR with 13% to 19% of all product categories.

It is evident that WSTS forecasts rarely emerge as the top performers within any given product category for $h = 2$ months. While WSTS' expert forecasts generally outperform its algorithmic updates, data-driven models consistently outshine both. When comparing against WSTS' expert forecasts (second row), SARIMA emerges as the best performer between 20% (MSE) and 30% (MAE) of the time, followed by ETS between 19% (MSE) and 27% (MAPE). Additionally, GPR and SES each excel between 13% (MAE and MAPE) and 17% and 18% (MSE), respectively. In contrast, WSTS' expert forecasts demonstrate the best performance between 7% (MAE and MAPE) and 13% (MSE) of the time.

This highlights that there is no one model for all situations but that the model choice should depend on the individual product category and the error measure with the greatest business relevance.

6. Discussion

6.1. Summary

Section 2 made the case that the semiconductor industry plays a crucial role in the broader economy and stressed the importance of reliable forecasts for operational and strategic decision making. Furthermore, the rapidly evolving technologies, complicated geopolitical considerations, and complex supply chains exposing the industry to the bullwhip effect make data-driven forecasting more challenging. Concurrently, industry insiders, such as those queried by the World Semiconductor Trade Statistics (WSTS) – a leading provider of semiconductor market data and forecasts, promise reliable forecasts based on a wealth of quantitative and qualitative insider information.

This motivated the first research hypothesis (H1) that expert forecasts exhibit higher accuracy compared to data-driven forecasts. This hypothesis was extensively examined in Section 5.1 for a forecast horizon of $h = 3$. The analysis of the benchmark concluded that the bi-quarterly expert forecasts indeed demonstrated superior accuracy on a quarterly horizon in terms of Mean Squared Error (MSE), Mean Absolute Percentage Error (MAPE), and Mean Absolute Error (MAE). In contrast to the superior performance of the expert forecasts, it was also found that the bi-quarterly algorithmic forecasts provided by WSTS showed potential for further improvement.

Furthermore, it was noted that industry insiders may have access to additional information owing to the timing of WSTS meetings, during which the forecasts are formulated. This observation prompted the formulation of the second research hypothesis (H2) that the additional information yields competitive data-driven forecasts, which was investigated in Section 5.2. It was found that the additional data and the shorter horizon of $h = 2$ months significantly improved forecasts based on quantitative models. Forecasts based on the SARIMA, GPR, ETS, and SES models consistently demonstrated superior accuracy. As a consequence, it is recommended that practitioners complement expert forecasts with data-driven methods to enhance the forecasts reliability.

General wisdom holds that experts excel in situations with limited historical data (Hyndman & Athanasopoulos, 2018). Consequently, the third research hypothesis (H3), which postulates that industry insiders outperform in short time series due to restricted quantitative data available for model training, was examined in Section 5.3. The analysis revealed that data-driven forecasts exhibited superiority across all examined time series lengths. Nonetheless, different models showcased higher accuracy under varying circumstances. Specifically, exponential smoothing models attained the highest accuracy for short time series, whereas SARIMA dominated in the medium-length scenario. Conversely, GPR outperformed for longer time series. This implies that several diverse models should be evaluated and the one that aligns most effectively with the given circumstances should be selected.

6.2. Implications, limitations, and future directions

This paper has presented the first comprehensive comparative analysis of expert semiconductor market forecasts. Long production lead times of semiconductors means that fulfilled orders have to be placed and the production has to be planned months in advance. As a result, industry experts must be very accurate when it comes to short-term sales forecasts. WSTS consolidates such industry forecasts to derive market forecast. Hence, we hope that the strong performance of the data driven methods motivates analysts and industry practitioners to employ data-driven methods to enhance existing forecasts – even when time series are short and the models are simple.

Additionally, this study shows that comprehensive multi-granularity modeling of the semiconductor market is feasible. Therefore, our hope is that this paper presents a first step in the direction of reconciling the fields of semiconductor cycle prediction, which assumes a higher-level view, and semiconductor sales and demand forecasting, which is much more granular: (1) The semiconductor market consists of diverse products from processors and memory chips to switches and sensors. Therefore, a multi-granularity study of the semiconductor market could yield insights into possible sub-cycles. (2) Conversely, demand and sales forecasts could benefit from granular and reliable market forecasts as an indicator for future sales. The same holds for semiconductor cycle indicators. Both approaches remain open for future study.

Nevertheless, this study has several limitations. The first relating to the **short time series lengths**, which had several consequences.

1. *Model simplicity*. We observed that simpler models such as SARIMA, ETS, and GPR performed stronger than more complex ML models such as RF. Furthermore, we abstained from the use of even more complex models such as Long Short Term Memory (LSTM) and boosting algorithms such as XGBoost.
2. *Purely autoregressive modeling*. We did not incorporate explanatory variables and relied solely on autoregressive modeling to keep the number of parameters at a minimum.

Table A.10
Mean ranks of the forecasts with horizon h = 2 months across the different time series lengths. Each row refers to a different error measure, sorted by WSTS' forecast type: algorithmic update, meeting (expert forecast), and overall. Lower values are preferable. The best value per row is bold and italic.

			WSTS	SARIMA	ETS	ET	GPR	KNN	RF	SES	SVM	Ensemble
Short	Alg. Update	MSE	8.1	3.4	3.3	5.9	***2.8***	8.9	5.9	4.0	8.9	3.9
		MAPE	8.2	3.6	3.5	6.1	***2.8***	8.8	5.8	4.0	8.5	3.7
		MAE	7.9	3.3	3.4	6.1	***2.9***	8.9	5.9	4.1	8.7	3.9
	Meeting	MSE	7.2	3.8	2.8	6.6	4.0	9.1	5.9	***2.7***	8.6	4.2
		MAPE	7.2	3.3	***2.8***	6.8	3.8	8.9	6.3	3.0	8.7	4.3
		MAE	7.0	3.3	3.0	6.7	4.1	8.9	6.1	***2.9***	8.8	4.2
	Overall	MSE	8.2	3.4	***3.0***	6.2	3.3	9.1	5.9	3.3	8.8	3.8
		MAPE	8.2	3.2	***2.8***	6.4	3.4	9.1	5.8	3.2	9.0	3.9
		MAE	7.9	3.2	***3.0***	6.4	3.4	9.2	5.8	3.2	8.9	4.0
Medium	Alg. Update	MSE	8.1	***2.8***	3.5	6.5	2.9	8.4	6.5	3.7	8.5	4.1
		MAPE	8.4	***2.8***	3.4	6.2	3.1	8.6	6.2	3.7	8.6	4.0
		MAE	8.1	***2.8***	3.5	6.4	3.0	8.6	6.3	3.9	8.5	3.9
	Meeting	MSE	7.1	***2.9***	3.6	6.6	3.8	8.7	6.0	4.1	8.2	4.0
		MAPE	7.8	***2.2***	3.7	6.6	3.7	8.7	5.9	4.0	8.2	4.1
		MAE	7.5	***2.4***	3.5	6.8	3.7	8.7	6.2	3.9	8.2	4.1
	Overall	MSE	8.1	***2.3***	3.4	6.7	3.1	8.8	6.4	3.7	8.5	4.1
		MAPE	8.4	***2.1***	3.5	6.4	3.1	9.1	6.2	3.9	8.4	3.9
		MAE	8.2	***2.0***	3.4	6.5	3.1	9.0	6.4	4.0	8.3	4.0
Long	Alg. Update	MSE	8.0	4.3	4.2	5.3	3.9	8.2	5.1	4.1	8.6	***3.3***
		MAPE	8.1	3.7	4.1	5.4	4.3	8.5	5.1	4.3	8.3	***3.1***
		MAE	7.9	3.8	4.2	5.5	4.2	8.4	5.1	4.3	8.4	***3.2***
	Meeting	MSE	6.7	3.7	3.5	6.4	***3.4***	8.9	5.9	3.7	8.6	4.2
		MAPE	6.8	***3.1***	3.5	6.7	***3.1***	8.8	6.2	3.9	8.7	4.1
		MAE	6.7	***3.1***	3.7	6.8	***3.1***	8.7	6.1	4.0	8.6	4.1
	Overall	MSE	7.4	3.9	3.7	6.2	***3.5***	8.7	5.6	3.7	8.7	3.6
		MAPE	7.7	***3.3***	***3.3***	6.4	***3.3***	8.9	5.8	4.0	8.8	3.5
		MAE	7.5	***3.3***	3.4	6.3	3.6	8.8	5.9	3.9	8.7	3.6

3. *Short forecast horizon.* The focus on quarterly forecasts comparisons limited the forecast horizon under study.

The use of simple models and purely autoregressive modeling could be viewed as strengthening our arguments in the context of the comparative analysis: they provide a natural benchmark, and they can be easily implemented by practitioners. The use of more complex models and the incorporation of indicators are possible extensions of this work. Future contributions could also emphasize longer forecast horizons.

The **hierarchical nature of product categorizations** was not considered in this study. This means that forecasts may not be coherent, i.e., forecasts of sub-groups may not sum up to forecasts of higher-level groups. Forecast coherence is often desired because it enables decisions which are based on these forecasts to be consistent across hierarchical levels. Furthermore, studies have shown that forecast reconciliation can further enhance forecast accuracy (Wickramasuriya et al., 2019).

Finally, our study focuses on **point forecasts** to evaluate forecast accuracy. Probabilistic forecasts (e.g., see Gneiting and Katzfuss (2014)) or uncertainty quantification in the form of prediction intervals may provide a better gauge of risks and opportunities to practitioners. This perspective was not the primary goal of this study, but prediction intervals can easily be obtained from many of the models that were employed here. A study of their accuracy may be of interest to analysts and industry practitioners.

CRediT authorship contribution statement

Louis Steinmeister: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing. **Markus Pauly:** Conceptualization, Funding acquisition, Methodology, Supervision, Writing – review & editing.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used the ChatGPT 3.5 model from OpenAI for minor language edits, aiming to enhance readability. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Louis Steinmeister reports financial support was provided by Graduate School of Logistics at TU Dortmund University. This paper was supported through a research collaboration between Infineon Technologies AG and TU Dortmund University through the Graduate School of Logistics. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This paper was supported through a research collaboration between Infineon Technologies AG and TU Dortmund University through the Graduate School of Logistics.

Appendix A. Mean ranks by time series lengths

See Table A.10.

Appendix B. Error measures by product category

Table B.11
Root Mean Squared error (RMSE), the square root of the Mean Squared Error (MSE), for the 110 examined WSTS product categories, which are organized by time series length and presented for the 3-month forecasts (first) and 2-month forecast (second). Averages are contained in the long, medium, short, 3-months, and 2-months rows respectively. The best result in each row, i.e. for each time series, is printed bold and italic. The RMSE was chosen instead of the MSE for readability purposes in this table.

ID	WSTS	SARIMA	ETS	ET	GPR	KNN	RF	SES	SVM	Ensemble
3 months	4.14	3.87	4.07	5.51	***3.78***	8.93	5.36	4.50	8.61	4.27
Long	6.87	6.24	6.71	9.39	***6.06***	15.55	9.11	7.59	14.76	7.08
A2	0.17	0.19	***0.15***	0.25	0.17	0.38	0.23	0.16	0.51	0.19
A5	***0.11***	0.18	0.18	0.13	0.15	0.22	0.13	0.15	0.20	0.15
A99	***0.48***	0.56	0.53	0.80	***0.48***	1.29	0.82	0.53	1.16	0.59
B2	***0.18***	0.20	0.21	0.28	0.24	0.45	0.27	0.20	0.53	0.19
B4	***0.16***	0.24	0.22	0.21	0.21	0.24	0.21	0.19	0.20	0.19
B99	***0.33***	0.48	0.45	0.37	0.42	0.43	0.38	0.37	0.40	0.38
C7	***1.18***	1.58	1.27	2.03	1.30	3.75	1.81	1.34	4.04	1.70
C99	2.11	2.24	2.04	3.57	***1.95***	6.62	3.15	2.22	7.30	2.78
CA	0.33	0.32	0.31	0.33	***0.30***	0.44	0.34	0.35	0.43	0.33
CC	***1.29***	1.69	1.37	2.14	1.41	4.16	1.86	1.46	4.35	1.79
D1	***0.22***	0.23	0.23	0.28	0.23	0.35	0.27	0.24	0.33	0.23
D2	***0.18***	0.23	0.21	0.29	0.23	0.41	0.30	0.19	0.48	0.24
D3	0.16	***0.15***	0.17	0.23	0.17	0.38	0.22	***0.15***	0.50	0.20
D99	***0.42***	0.59	0.48	0.70	0.58	1.08	0.65	0.46	1.29	0.55
E1	***0.10***	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.14	0.12
E2	***0.11***	***0.11***	0.12	0.13	0.13	0.17	0.13	0.12	0.17	0.12
E99	***0.15***	***0.15***	0.16	***0.15***	***0.15***	0.18	***0.15***	0.17	0.18	0.16
F1	0.11	0.10	0.09	0.11	***0.08***	0.13	0.11	0.09	0.13	0.09
F2	1.59	1.53	1.52	1.50	1.68	1.93	1.50	***1.36***	2.10	1.43
F3	0.32	0.39	***0.30***	0.44	0.35	0.68	0.42	0.31	0.79	0.33
F5	5.80	5.94	5.69	6.18	5.85	6.68	6.23	7.84	7.36	***5.37***
F6	0.99	0.86	***0.84***	0.87	0.90	0.99	0.87	0.88	0.86	***0.84***
F8	0.21	0.15	0.17	0.18	0.16	0.28	0.18	***0.14***	0.19	0.16
F99	6.65	6.97	8.32	8.14	7.94	9.08	8.17	9.36	11.53	***6.53***
G99	0.11	0.11	0.11	0.11	0.11	***0.10***	0.11	0.11	0.13	0.11
J0	***3.75***	3.82	4.74	6.72	4.08	10.73	6.79	4.26	10.54	5.13
J4	***0.48***	0.50	0.55	0.71	0.49	1.15	0.63	0.50	1.06	0.57
J6	***2.73***	3.22	3.52	4.85	2.84	7.93	4.74	3.21	7.38	3.69
J7	0.72	0.95	0.74	0.93	0.73	1.32	0.88	***0.68***	1.10	0.80
J9	***0.49***	0.60	0.62	0.84	0.61	1.39	0.78	0.53	1.27	0.64
J99	9.10	***8.75***	10.23	15.80	9.52	27.52	16.16	10.53	29.76	12.28
JA	5.98	***5.41***	6.37	9.89	6.37	17.01	9.46	7.04	19.10	7.71
L1	0.03	0.03	0.03	0.03	***0.02***	***0.02***	***0.02***	***0.02***	0.03	***0.02***
L2a	0.35	0.40	0.41	0.34	***0.32***	0.39	0.35	0.35	0.36	0.33
L2b	0.04	0.04	0.04	0.04	0.04	0.04	0.04	***0.03***	0.05	***0.03***
L2c	3.20	3.20	***3.05***	5.47	3.71	10.27	4.70	3.83	10.54	4.24
L99	21.83	***15.97***	20.19	27.02	18.43	54.18	25.59	22.41	53.28	21.82
LA	18.64	***13.17***	17.50	21.95	16.86	44.46	21.00	18.75	40.56	17.96
LS	***4.37***	5.93	5.51	7.91	5.90	16.21	7.18	4.96	17.44	6.40
M1	28.93	35.32	27.90	33.67	27.67	53.12	34.56	30.86	52.37	***26.74***
M2	0.13	0.13	0.13	0.15	0.13	0.16	0.14	***0.11***	0.15	***0.11***
M8	0.43	0.44	0.50	***0.40***	0.49	0.41	0.49	0.66	0.44	0.46
M99	46.77	49.26	40.48	49.73	***37.72***	75.78	53.49	45.91	75.55	39.99
P1	***10.39***	10.71	12.80	13.02	10.59	14.24	12.71	13.42	14.83	11.98
P2	***2.83***	3.55	3.63	4.78	3.54	8.08	4.52	3.03	9.06	3.97
P5	***2.55***	3.17	3.24	4.39	3.46	7.82	4.26	2.77	9.48	3.72
P99	10.93	11.10	12.96	13.92	***10.73***	19.96	12.74	13.19	16.38	12.28
S2	70.79	***49.17***	65.42	99.61	53.28	166.49	94.91	77.94	146.56	68.48
S3	***2.77***	3.83	3.49	4.86	3.12	9.10	3.92	2.97	10.56	3.65
T99	71.85	57.77	66.25	113.07	***56.83***	189.21	106.86	83.20	164.85	76.28
Medium	***1.77***	1.78	1.81	2.27	1.81	3.51	2.22	1.86	3.60	1.92
A3	0.26	0.23	0.29	0.41	***0.22***	0.60	0.36	0.29	0.51	0.27
A4	0.09	***0.08***	***0.08***	***0.08***	***0.08***	0.09	***0.08***	***0.08***	***0.08***	***0.08***
B3	***0.08***	0.10	***0.08***	0.10	0.09	0.15	0.09	0.09	0.17	0.09
C10	0.84	0.98	***0.72***	1.15	0.76	1.97	1.08	0.87	2.03	0.94
C8	0.22	***0.20***	0.21	0.31	***0.20***	0.50	0.28	0.22	0.52	0.25
C9	0.37	0.35	0.33	0.57	***0.31***	0.90	0.53	0.33	1.11	0.41
CB	1.11	1.22	***0.91***	1.61	***0.91***	2.90	1.42	1.04	3.05	1.24

(*continued on next page*)

Table B.11 (*continued*).

F7	0.23	0.21	***0.18***	0.27	0.19	0.32	0.29	0.22	0.32	0.21
H99	***2.11***	2.41	2.49	3.28	2.41	6.23	3.21	2.42	6.95	2.66
JB	0.82	0.94	0.86	0.83	0.82	***0.77***	0.83	0.97	0.89	0.79
JBa	0.56	0.56	***0.52***	0.57	0.55	0.59	0.58	0.59	0.60	0.55
JBb	***0.36***	0.50	0.49	0.60	0.46	0.83	0.65	0.50	0.74	0.46
JC	0.58	0.76	0.72	0.57	0.71	0.70	***0.54***	0.73	0.58	0.58
JCc	***0.15***	0.20	0.22	0.25	0.23	0.30	0.24	0.17	0.30	0.21
JCd	0.49	0.66	0.59	0.48	0.62	0.50	0.47	0.58	***0.45***	0.52
JCe	0.12	0.13	0.13	0.13	***0.11***	0.20	0.12	0.13	0.12	0.12
JD	3.90	4.08	3.95	5.57	***3.89***	9.52	5.55	4.36	10.23	4.49
JE	***2.19***	2.42	2.29	3.64	2.51	5.91	3.61	2.44	6.81	3.07
JF	***0.48***	0.66	0.60	0.81	0.58	1.60	0.73	0.59	1.83	0.69
L3	3.28	2.99	2.99	3.72	3.16	6.59	3.92	***2.96***	6.13	3.26
L5	3.13	2.88	3.17	4.03	3.29	7.02	4.45	***2.64***	6.00	3.38
L6	9.68	***8.82***	9.54	10.84	9.19	16.44	10.50	9.64	16.51	9.75
L6g	8.76	***8.17***	9.36	10.05	9.13	13.19	9.67	8.90	12.70	9.23
L6h	1.16	***1.13***	1.25	1.49	1.16	2.81	1.47	1.24	2.72	1.31
L6i	0.79	***0.69***	0.84	0.85	0.77	1.04	0.82	0.74	0.88	0.76
L7	***9.72***	10.91	10.09	13.75	10.12	20.27	13.06	11.63	22.29	10.43
L8	1.06	***1.05***	1.09	1.55	1.14	3.52	1.48	1.17	4.16	1.32
L9	0.58	***0.54***	0.63	1.13	1.07	1.43	1.06	0.57	1.50	0.90
M6	0.23	***0.21***	0.28	0.22	***0.21***	0.28	0.23	0.23	0.26	0.22
M7A	0.98	0.56	0.59	0.83	***0.50***	0.81	0.76	0.75	0.66	0.63
P4	***0.53***	0.66	0.69	0.65	0.66	0.80	0.63	0.61	0.64	0.61
Short	1.98	2.02	***1.92***	2.28	1.97	3.31	2.26	1.97	3.34	1.95
F10	***0.82***	1.27	0.94	1.38	1.86	1.57	1.40	1.15	1.44	1.16
F9	0.41	0.41	0.43	0.52	***0.37***	1.29	0.52	0.46	1.21	0.48
H1	0.85	0.85	***0.80***	0.93	0.82	1.30	0.89	0.85	1.34	0.85
H2	0.62	0.73	0.73	0.67	0.64	0.83	0.70	0.72	0.87	***0.60***
H3	***0.56***	0.63	0.70	0.81	0.62	0.89	0.81	0.65	0.98	0.67
H4	***0.46***	0.58	0.48	0.55	0.50	0.79	0.51	0.47	0.91	0.49
H6	***1.77***	1.84	2.13	2.37	2.04	3.06	2.37	1.99	5.25	2.10
JBc	***0.02***	0.03	0.03	***0.02***	0.03	0.03	***0.02***	0.03	0.03	***0.02***
JCf	***0.06***	0.07	0.07	0.08	0.07	0.11	0.08	0.07	0.09	0.07
JDa	***3.23***	3.27	3.34	4.40	3.44	6.83	4.33	3.58	6.96	3.69
JDb	***0.43***	0.47	0.50	0.60	0.46	1.46	0.57	0.44	1.84	0.51
JDc	***0.58***	0.67	0.76	0.83	0.78	1.17	0.80	0.70	1.09	0.68
JDd	0.63	0.88	***0.61***	0.77	0.77	0.93	0.76	0.68	1.06	0.72
JDe	0.40	0.37	0.38	0.37	0.36	0.70	0.35	0.35	***0.33***	0.34
JEa	***0.27***	0.34	0.41	0.47	0.40	0.72	0.42	0.40	0.94	0.42
JEb	***1.99***	2.13	2.58	3.57	3.48	5.08	3.52	2.49	5.39	3.14
L3a	1.53	1.45	1.42	1.28	1.30	1.61	1.36	1.42	***1.25***	1.31
L3b	***1.94***	2.13	2.09	2.37	2.11	3.07	2.68	2.08	2.86	2.23
L3c	0.55	1.23	0.47	0.62	0.55	1.41	0.62	***0.43***	1.50	0.54
L4	1.02	0.96	1.56	1.25	***0.81***	2.07	1.24	0.82	1.96	0.88
L5a	1.97	1.85	1.94	1.86	1.91	3.14	1.84	***1.79***	2.33	1.80
L5b	1.27	1.34	1.27	1.61	1.94	3.15	1.56	***0.94***	2.60	1.43
L5c	0.23	0.23	***0.20***	0.26	0.22	0.38	0.26	0.22	0.26	0.21
L7a/b/c/d/f	***9.28***	10.65	9.41	12.60	9.58	16.70	11.97	10.93	18.00	9.96
L7e	3.90	3.45	***3.10***	3.92	3.28	7.63	3.84	3.81	9.17	3.80
L8a	0.69	***0.59***	0.86	1.03	0.77	2.46	1.02	0.78	2.54	0.91
L8b	***0.52***	0.58	0.57	0.84	0.65	1.25	0.82	0.56	1.34	0.69
M7B	20.62	18.75	17.11	19.15	16.66	25.32	19.42	17.74	22.21	***16.17***
P6	0.73	0.73	0.68	0.98	0.81	1.04	0.93	***0.64***	1.21	0.81
2 months	4.14	2.33	2.44	3.51	***2.31***	5.70	3.39	3.06	5.60	2.87
Long	6.87	3.87	4.12	6.01	***3.83***	9.87	5.80	5.33	9.66	4.89
A2	0.17	0.13	0.12	0.14	0.12	0.20	0.13	***0.10***	0.31	***0.10***
A5	0.11	0.14	0.13	***0.10***	0.13	0.16	***0.10***	***0.10***	0.12	0.12
A99	0.48	0.47	0.42	0.51	0.44	0.80	0.51	***0.37***	0.67	0.40
B2	***0.18***	0.24	0.21	0.21	0.24	0.29	0.20	0.19	0.36	0.19
B4	0.16	0.13	0.16	0.13	0.16	0.16	0.14	0.13	0.13	***0.12***
B99	0.33	0.42	0.37	0.29	0.39	0.32	0.30	0.29	***0.28***	0.29
C7	***1.18***	1.68	1.39	1.57	1.32	2.54	1.45	1.48	2.71	1.51
C99	2.11	2.28	2.06	2.64	***1.86***	4.50	2.42	2.47	4.77	2.48
CA	0.33	0.25	0.23	0.23	***0.21***	0.28	0.22	0.24	0.28	0.22
CC	***1.29***	1.85	1.57	1.72	1.42	2.82	1.58	1.66	2.94	1.67
D1	0.22	0.17	0.16	0.18	0.17	0.20	0.18	***0.15***	0.24	0.16
D2	0.18	0.21	0.18	0.20	0.19	0.30	0.19	***0.17***	0.32	0.18
D3	0.16	***0.12***	***0.12***	0.17	***0.12***	0.26	0.15	0.14	0.32	0.14
D99	0.42	0.44	0.40	0.47	0.43	0.75	0.46	***0.39***	0.87	0.42

(*continued on next page*)

Table B.11 (*continued*).

E1	0.10	0.07	0.08	0.08	0.07	0.09	0.07	***0.06***	0.09	0.07
E2	0.11	***0.07***	0.08	0.08	0.08	0.09	0.09	***0.07***	0.12	0.08
E99	0.15	***0.09***	0.10	0.10	***0.09***	***0.09***	***0.09***	0.10	0.12	***0.09***
F1	0.11	0.06	***0.05***	0.06	***0.05***	0.08	0.06	***0.05***	0.06	***0.05***
F2	1.59	0.68	0.74	0.85	0.85	1.12	0.85	***0.65***	1.35	0.74
F3	0.32	0.26	0.27	0.31	0.26	0.50	0.29	***0.23***	0.50	0.26
F5	5.80	***3.52***	4.01	4.63	4.04	4.76	4.64	4.25	5.33	3.77
F6	0.99	0.55	0.64	0.51	0.64	0.57	0.55	0.52	***0.49***	0.50
F8	0.21	***0.06***	0.08	0.10	0.08	0.18	0.11	0.08	0.10	0.08
F99	6.65	4.47	4.88	5.16	4.68	6.25	5.09	5.24	7.26	***4.43***
G99	0.11	***0.06***	***0.06***	0.07	***0.06***	0.09	0.07	***0.06***	0.08	0.07
J0	3.75	3.02	2.42	3.56	2.46	6.68	3.55	***2.32***	6.51	2.46
J4	0.48	0.24	0.25	0.36	0.26	0.59	0.35	***0.23***	0.54	0.29
J6	2.73	1.93	1.94	2.72	1.80	4.60	2.49	***1.71***	4.76	1.90
J7	0.72	0.47	0.41	0.48	0.43	0.82	0.45	0.47	0.63	***0.37***
J9	0.49	0.45	0.38	0.47	0.41	0.86	0.48	0.34	0.76	***0.32***
J99	9.10	5.07	5.08	9.07	***4.79***	18.55	9.15	5.31	18.52	6.11
JA	5.98	***2.69***	2.99	5.80	2.97	11.38	5.64	3.89	12.17	4.22
L1	0.03	***0.02***	0.03	***0.02***	***0.02***	***0.02***	***0.02***	***0.02***	0.03	***0.02***
L2a	0.35	0.25	0.24	0.22	0.26	0.24	0.23	***0.18***	0.23	0.20
L2b	0.04	***0.02***	***0.02***	0.03	***0.02***	0.03	***0.02***	***0.02***	0.04	***0.02***
L2c	3.20	1.45	***1.41***	3.19	1.61	6.72	2.99	2.23	6.41	2.46
L99	21.83	10.68	12.23	17.65	***10.63***	37.43	16.77	14.80	33.61	14.83
LA	18.64	***8.92***	11.23	14.76	10.11	30.80	14.38	12.43	26.03	12.73
LS	4.37	2.90	***2.76***	4.68	2.80	10.37	4.71	3.07	10.53	3.57
M1	28.93	19.44	***18.65***	20.98	19.17	32.64	22.87	25.83	35.41	19.31
M2	0.13	0.10	0.10	0.11	0.10	0.14	0.11	***0.09***	0.13	***0.09***
M8	0.43	***0.29***	0.30	0.30	0.34	0.32	0.32	0.50	0.31	0.30
M99	46.77	28.79	28.31	33.60	***26.37***	48.21	32.72	36.18	51.81	28.73
P1	10.39	***6.56***	7.12	8.19	6.64	10.02	7.64	8.46	10.63	7.66
P2	2.83	***1.88***	1.97	3.14	1.99	5.39	3.02	2.13	5.65	2.45
P5	2.55	1.70	***1.66***	2.85	1.87	5.30	2.77	2.04	6.14	2.25
P99	10.93	***5.97***	6.47	8.48	6.42	13.13	8.14	8.51	12.01	8.01
S2	70.79	***32.89***	37.13	64.12	33.11	101.47	59.85	54.16	94.54	49.90
S3	***2.77***	3.53	3.15	3.36	3.12	5.99	3.03	2.94	6.80	3.15
T99	71.85	***35.79***	41.38	72.01	35.88	114.29	68.42	59.43	108.95	55.01
Medium	1.77	***0.96***	1.01	1.43	1.01	2.32	1.37	1.14	2.28	1.19
A3	0.26	0.22	0.24	0.27	***0.21***	0.44	0.27	0.22	0.33	0.23
A4	0.09	***0.02***	0.03	0.04	0.03	0.05	0.05	***0.02***	0.04	0.03
B3	0.08	***0.05***	***0.05***	0.07	***0.05***	0.11	0.06	***0.05***	0.11	0.06
C10	0.84	***0.46***	0.50	0.84	0.47	1.26	0.80	0.69	1.31	0.74
C8	0.22	0.19	0.19	0.26	***0.18***	0.37	0.25	0.23	0.38	0.23
C9	0.37	0.32	0.29	0.41	***0.26***	0.62	0.40	0.30	0.75	0.35
CB	1.11	0.65	0.65	1.13	***0.60***	1.79	1.01	0.91	2.01	0.98
F7	0.23	0.13	***0.11***	0.16	***0.11***	0.22	0.16	***0.11***	0.19	0.12
H99	2.11	***1.01***	1.18	1.92	1.17	4.28	1.75	1.49	4.44	1.48
JB	0.82	0.28	***0.26***	0.44	0.30	0.40	0.44	0.37	0.43	0.31
JBa	0.56	***0.17***	***0.17***	0.29	0.18	0.28	0.29	0.20	0.30	0.18
JBb	0.36	0.25	***0.24***	0.34	***0.24***	0.49	0.34	0.26	0.48	0.28
JC	0.58	0.37	0.36	0.37	0.37	0.50	0.38	0.40	0.39	***0.35***
JCc	0.15	0.14	0.21	0.16	0.16	0.16	0.17	***0.13***	0.19	0.15
JCd	0.49	0.32	0.29	0.30	0.29	0.31	0.32	0.36	***0.28***	0.31
JCe	0.12	***0.08***	0.10	0.09	0.10	0.09	0.09	***0.08***	0.10	***0.08***
JD	3.90	2.37	***2.13***	3.46	2.20	6.78	3.48	2.51	6.28	2.72
JE	2.19	***0.78***	0.98	2.27	1.05	3.96	2.21	1.48	4.03	1.68
JF	0.48	0.40	***0.37***	0.48	0.41	0.90	0.45	0.41	1.10	0.40
L3	3.28	1.60	1.71	2.20	1.99	3.74	2.10	***1.59***	3.94	1.83
L5	3.13	1.22	1.49	1.77	1.49	3.49	1.89	***0.97***	3.00	1.36
L6	9.68	***5.61***	6.09	7.13	5.75	11.39	6.85	6.68	10.71	6.51
L6g	8.76	***5.29***	5.78	6.61	5.68	9.12	6.04	6.13	8.36	6.10
L6h	1.16	0.66	0.69	0.92	***0.65***	1.88	0.90	0.72	1.84	0.77
L6i	0.79	***0.40***	0.45	0.56	0.43	0.81	0.56	0.41	0.51	0.43
L7	9.72	5.00	***4.82***	9.04	5.23	13.81	8.62	6.53	14.40	6.92
L8	1.06	0.57	***0.46***	0.94	0.50	2.40	0.86	0.71	2.75	0.77
L9	0.58	0.26	0.45	0.61	0.31	0.80	0.62	***0.25***	0.77	0.42
M6	0.23	0.15	0.19	0.15	0.19	0.19	0.15	0.16	0.16	***0.14***
M7A	0.98	***0.34***	0.37	0.52	0.36	0.62	0.50	0.46	0.43	0.37
P4	0.53	0.45	0.48	0.48	0.47	0.64	0.48	***0.43***	0.58	0.45

(*continued on next page*)

Table B.11 (*continued*).

Short	1.98	1.14	***1.06***	1.43	1.07	2.12	1.39	1.22	2.17	1.20
F10	0.82	1.09	0.84	0.82	1.31	0.95	0.82	0.80	0.97	***0.73***
F9	0.41	0.30	0.33	0.41	***0.27***	0.88	0.41	0.36	0.82	0.38
H1	0.85	0.31	***0.28***	0.47	0.35	0.82	0.46	0.29	0.83	0.35
H2	0.62	0.31	***0.27***	0.43	***0.27***	0.48	0.43	0.33	0.55	0.33
H3	0.56	***0.28***	0.31	0.50	***0.28***	0.52	0.49	0.40	0.54	0.39
H4	0.46	0.33	0.31	0.37	0.33	0.56	0.36	***0.30***	0.60	0.33
H6	1.77	***0.81***	0.92	1.31	0.86	2.08	1.29	1.02	3.45	1.13
JBc	***0.02***	***0.02***	***0.02***	***0.02***	***0.02***	***0.02***	***0.02***	***0.02***	***0.02***	***0.02***
JCf	0.06	***0.04***	0.05	0.05	0.06	0.07	0.05	***0.04***	0.06	0.05
JDa	3.23	1.73	***1.48***	2.41	1.70	3.95	2.43	1.77	4.57	2.01
JDb	0.43	0.25	0.29	0.37	***0.22***	0.99	0.36	0.26	1.21	0.31
JDc	0.58	0.59	0.59	0.55	0.54	0.82	0.55	0.70	0.74	***0.53***
JDd	0.63	0.42	***0.39***	0.43	***0.39***	0.59	0.47	0.45	0.55	0.41
JDe	0.40	0.32	0.35	***0.28***	0.30	0.45	***0.28***	0.33	***0.28***	***0.28***
JEa	0.27	***0.21***	0.22	0.29	0.22	0.42	0.28	0.26	0.60	0.27
JEb	1.99	1.47	1.48	2.14	***1.14***	3.38	2.25	1.56	3.41	1.79
L3a	1.53	0.84	0.74	0.83	0.84	1.03	0.78	***0.73***	1.01	0.79
L3b	1.94	1.35	1.34	1.55	1.52	2.10	1.58	***1.29***	2.00	1.44
L3c	0.55	0.66	***0.20***	0.33	0.24	0.64	0.34	0.22	0.92	0.25
L4	1.02	0.55	0.55	0.65	0.47	1.41	0.66	***0.45***	1.27	0.49
L5a	1.97	0.72	0.82	0.86	0.78	1.65	0.80	***0.65***	1.05	0.70
L5b	1.27	0.46	0.65	0.86	0.98	1.72	0.87	***0.38***	1.37	0.64
L5c	0.23	0.13	0.14	0.16	0.14	0.22	0.16	***0.11***	0.18	0.13
L7a/b/c/d/f	9.28	4.29	***3.98***	7.51	4.31	10.11	7.23	4.84	11.17	5.43
L7e	3.90	2.34	***2.09***	2.89	2.23	5.30	2.96	2.71	6.12	2.77
L8a	0.69	***0.35***	0.37	0.53	0.37	1.67	0.56	0.39	1.61	0.46
L8b	0.52	0.42	***0.36***	0.59	0.37	0.89	0.57	0.44	0.94	0.51
M7B	20.62	12.21	10.98	13.22	***10.02***	17.07	12.23	13.84	15.39	11.33
P6	0.73	0.38	***0.35***	0.58	0.39	0.61	0.55	0.40	0.68	0.46

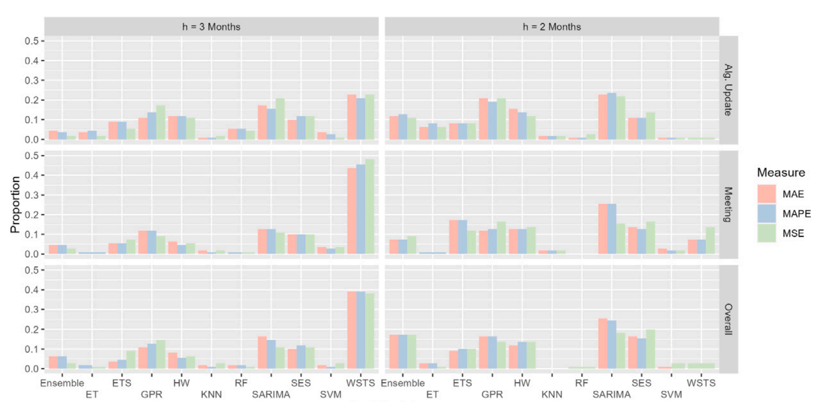

Fig. C.4. Bar chart illustrating the frequency of the best-performing forecasts based on Mean Squared Error (MSE) in red, Mean Absolute Percentage Error (MAPE) in blue, and Mean Absolute Error (MAE) in green, for the 110 WSTS product categories. The left side of the chart represents the outcomes of forecasts with a horizon of $h = 3$ months, while the right side represents those with a forecast horizon of $h = 2$ months. The rows are organized by algorithmic update, expert forecasts ("Meeting"), and overall performance, mirroring the structure of Tables 4 and 6.

Table B.12
Mean absolute percentage errors (MAPE) for the 110 examined WSTS product categories, which are organized by time series length and presented for the 3-month forecasts (first) and 2-month forecast (second). Averages are contained in the long, medium, short, 3-months, and 2-months rows respectively. The best result in each row, i.e. for each time series, is printed bold and italic.

ID	WSTS	SARIMA	ETS	ET	GPR	KNN	RF	SES	SVM	Ensemble
3 months	***8.4%***	9.0%	8.8%	10.6%	9.0%	14.5%	10.3%	9.0%	14.7%	9.1%
Long	***7.7%***	8.1%	8.1%	9.6%	8.0%	12.7%	9.5%	8.5%	13.3%	8.3%
A2	***5.5%***	6.6%	5.9%	8.3%	7.3%	12.7%	7.7%	6.1%	15.5%	6.8%
A5	***4.6%***	7.1%	6.5%	5.4%	6.3%	9.1%	5.7%	6.5%	8.3%	5.9%
A99	***4.2%***	5.2%	5.5%	7.6%	5.0%	12.3%	7.5%	5.4%	10.3%	5.8%
B2	5.2%	***4.9%***	5.2%	9.1%	6.1%	13.4%	8.8%	5.7%	16.3%	6.4%
B4	***6.0%***	8.5%	7.4%	7.6%	7.5%	9.7%	7.8%	6.8%	7.8%	7.3%
B99	***4.6%***	6.5%	5.8%	5.2%	5.6%	5.7%	5.5%	5.8%	5.3%	5.3%
C7	***4.0%***	4.8%	4.3%	7.1%	4.5%	13.0%	6.1%	4.8%	14.1%	6.0%
C99	***3.4%***	4.0%	3.6%	6.9%	3.7%	12.0%	6.0%	4.0%	13.4%	5.4%
CA	5.3%	5.2%	***4.6%***	5.3%	5.1%	7.5%	5.4%	5.7%	6.9%	5.4%
CC	***4.1%***	4.6%	4.2%	7.0%	4.6%	13.1%	6.2%	5.1%	13.5%	5.9%
D1	***4.1%***	5.1%	4.8%	5.2%	5.3%	6.8%	5.0%	4.9%	6.2%	5.0%
D2	***3.7%***	4.8%	4.8%	6.3%	5.0%	8.6%	6.5%	4.1%	9.5%	5.1%
D3	5.8%	5.5%	6.4%	8.8%	6.5%	14.6%	8.2%	***5.4%***	20.2%	7.5%
D99	***3.1%***	4.8%	4.0%	5.8%	4.7%	8.1%	5.3%	3.8%	10.3%	4.5%
E1	***8.2%***	9.2%	10.1%	9.8%	8.4%	9.4%	9.9%	9.3%	11.5%	8.6%
E2	***9.9%***	10.1%	10.3%	12.5%	11.4%	14.3%	12.7%	10.7%	14.8%	11.2%
E99	***6.6%***	6.9%	7.0%	6.9%	6.7%	8.2%	6.7%	7.9%	8.0%	7.2%
F1	8.3%	6.9%	6.2%	8.3%	***6.1%***	9.9%	8.5%	6.3%	11.0%	6.4%
F2	4.5%	4.0%	3.9%	4.0%	4.6%	5.2%	4.1%	***3.6%***	5.6%	3.9%
F3	5.5%	5.9%	5.5%	7.7%	5.5%	10.4%	7.3%	***5.3%***	12.3%	5.8%
F5	11.1%	11.7%	11.6%	10.1%	11.6%	11.3%	10.6%	15.7%	12.7%	***9.7%***
F6	11.8%	10.0%	10.7%	9.9%	11.2%	11.8%	10.4%	10.2%	9.9%	***9.7%***
F8	9.3%	7.3%	7.8%	9.4%	7.9%	14.5%	8.7%	***7.1%***	8.7%	7.8%
F99	***5.2%***	5.6%	7.2%	6.5%	6.8%	7.6%	6.6%	8.3%	9.8%	5.6%
G99	54.8%	53.4%	55.2%	56.6%	54.4%	***52.4%***	58.4%	55.9%	66.9%	55.3%
J0	4.8%	***4.7%***	5.7%	7.9%	5.0%	11.5%	7.7%	5.5%	11.1%	6.0%
J4	6.3%	6.3%	6.7%	7.8%	***6.1%***	11.5%	7.1%	6.9%	10.0%	6.8%
J6	***5.5%***	6.1%	6.9%	9.3%	5.7%	13.8%	9.1%	6.6%	13.4%	7.2%
J7	5.7%	7.5%	***5.4%***	7.6%	6.0%	9.5%	7.4%	6.0%	8.3%	6.6%
J9	***4.2%***	5.2%	5.0%	7.0%	4.9%	10.1%	6.4%	4.3%	10.3%	4.9%
J99	4.5%	***4.1%***	4.9%	7.8%	4.6%	12.5%	8.0%	5.5%	13.0%	6.1%
JA	4.9%	***4.4%***	4.9%	8.0%	4.9%	13.5%	7.8%	6.0%	14.0%	6.2%
L1	11.7%	13.5%	12.8%	11.6%	11.6%	11.7%	***10.5%***	12.6%	14.3%	11.0%
L2a	8.2%	8.4%	9.2%	8.0%	***6.9%***	9.7%	7.9%	7.8%	8.3%	7.5%
L2b	16.2%	18.5%	18.5%	18.7%	17.6%	23.5%	18.7%	***15.7%***	27.4%	16.8%
L2c	5.8%	***5.5%***	5.6%	9.5%	6.5%	14.1%	8.4%	6.4%	15.4%	7.3%
L99	4.8%	***3.5%***	4.2%	6.4%	4.2%	11.3%	6.0%	5.4%	11.4%	5.1%
LA	5.3%	***3.7%***	4.6%	6.5%	4.7%	12.3%	5.8%	5.6%	11.3%	5.0%
LS	***4.7%***	6.7%	6.0%	8.1%	6.5%	13.5%	7.5%	5.8%	15.8%	6.8%
M1	12.0%	14.1%	12.1%	16.1%	***11.0%***	22.9%	16.4%	14.0%	17.6%	12.6%
M2	11.1%	11.5%	10.1%	13.7%	10.5%	14.8%	13.1%	***9.2%***	15.0%	10.6%
M8	26.1%	24.3%	26.3%	***21.6%***	23.1%	24.1%	25.1%	32.8%	24.2%	23.1%
M99	11.4%	11.9%	9.6%	14.0%	***9.2%***	20.8%	14.4%	12.1%	17.3%	11.1%
P1	***6.7%***	7.5%	8.8%	8.4%	7.8%	9.2%	8.1%	9.1%	10.4%	7.9%
P2	***4.9%***	6.0%	6.0%	8.2%	5.9%	11.3%	7.8%	5.7%	13.6%	6.8%
P5	***4.8%***	6.0%	6.1%	8.8%	6.8%	12.3%	8.5%	5.5%	16.3%	7.2%
P99	***4.8%***	5.1%	6.1%	6.4%	5.0%	8.4%	5.9%	6.3%	6.9%	5.8%
S2	5.2%	***4.1%***	4.9%	7.9%	4.2%	11.4%	7.6%	6.4%	9.9%	5.5%
S3	***3.0%***	4.3%	4.1%	5.4%	3.7%	10.4%	4.5%	3.7%	11.9%	4.4%
T99	4.4%	***3.7%***	4.1%	7.5%	3.8%	11.0%	7.2%	5.8%	9.1%	5.1%
Medium	***8.0%***	8.2%	8.2%	10.1%	8.3%	14.1%	9.7%	8.3%	14.0%	8.6%
A3	5.4%	4.9%	6.5%	8.9%	***4.5%***	12.5%	7.8%	6.3%	9.9%	6.1%
A4	6.4%	6.7%	6.1%	7.1%	6.8%	9.4%	6.8%	***5.7%***	7.4%	6.2%
B3	10.0%	12.6%	***9.7%***	13.9%	11.3%	22.9%	13.1%	10.5%	26.1%	12.0%
C10	***5.1%***	6.8%	5.4%	7.8%	5.5%	13.5%	7.5%	5.8%	15.1%	5.8%
C8	8.8%	9.0%	***8.5%***	11.8%	9.0%	15.3%	10.7%	8.7%	16.1%	9.5%
C9	6.4%	6.0%	5.6%	9.4%	***5.5%***	16.8%	8.6%	5.8%	20.3%	6.8%
CB	4.6%	5.6%	***4.5%***	7.6%	***4.5%***	14.5%	6.6%	4.7%	16.9%	5.5%

(*continued on next page*)

Table B.12 (*continued*).

F7	6.6%	6.3%	6.0%	8.2%	***5.8%***	10.2%	8.4%	6.9%	9.9%	6.4%
H99	***3.9%***	4.1%	4.6%	5.5%	4.2%	10.1%	5.3%	4.3%	12.4%	4.6%
JB	10.2%	12.1%	10.9%	11.4%	***9.9%***	10.4%	11.0%	12.3%	11.4%	10.2%
JBa	12.5%	12.9%	***11.5%***	13.3%	11.8%	13.5%	12.9%	13.4%	13.5%	12.2%
JBb	***11.5%***	13.6%	13.0%	15.8%	13.0%	20.1%	15.9%	14.4%	17.2%	13.3%
JC	6.8%	7.2%	8.2%	7.1%	7.8%	8.9%	***6.5%***	8.0%	7.2%	6.9%
JCc	10.5%	11.2%	13.4%	14.8%	14.6%	14.5%	13.7%	***10.4%***	17.5%	12.5%
JCd	***8.5%***	10.8%	9.3%	10.1%	10.5%	11.9%	9.6%	9.6%	9.5%	9.5%
JCe	***7.1%***	8.3%	8.2%	8.7%	7.7%	13.9%	8.7%	8.7%	7.9%	8.1%
JD	6.3%	6.2%	6.1%	8.7%	***5.6%***	14.0%	9.0%	6.8%	13.7%	6.9%
JE	***6.0%***	6.4%	6.4%	10.5%	6.9%	16.1%	10.3%	6.7%	19.5%	8.9%
JF	***5.0%***	5.8%	5.7%	8.0%	5.1%	14.0%	7.4%	6.3%	14.7%	6.7%
L3	***13.2%***	14.1%	14.3%	15.7%	14.9%	21.8%	16.4%	14.0%	20.1%	14.9%
L5	7.4%	6.9%	8.0%	9.2%	8.3%	13.3%	10.5%	***6.2%***	13.2%	8.2%
L6	***7.5%***	***7.5%***	7.7%	10.8%	8.0%	16.5%	9.9%	8.6%	15.4%	9.0%
L6g	8.6%	***7.8%***	8.8%	12.6%	9.6%	16.4%	11.5%	9.4%	15.6%	10.4%
L6h	6.2%	***5.6%***	6.1%	8.2%	6.0%	13.2%	8.0%	5.9%	12.5%	6.7%
L6i	11.1%	***10.5%***	11.9%	13.6%	11.5%	17.8%	13.3%	***10.5%***	13.0%	11.6%
L7	5.8%	6.7%	5.7%	7.3%	5.7%	9.7%	6.6%	6.8%	10.7%	***5.3%***
L8	***5.6%***	***5.6%***	6.4%	9.5%	6.3%	20.2%	8.7%	6.8%	24.9%	7.8%
L9	13.4%	12.7%	13.3%	15.5%	17.7%	20.1%	14.4%	***12.1%***	17.9%	14.8%
M6	5.7%	5.8%	7.6%	5.9%	5.6%	7.7%	5.9%	***5.4%***	7.2%	5.5%
M7A	14.0%	6.4%	8.1%	9.2%	***6.2%***	9.7%	8.7%	8.6%	8.8%	7.4%
P4	***6.4%***	8.0%	8.2%	7.8%	8.4%	9.4%	7.6%	7.2%	7.2%	7.4%
Short	***9.9%***	11.4%	10.4%	12.7%	11.4%	18.0%	12.5%	10.6%	17.8%	11.0%
F10	***15.8%***	24.8%	19.3%	26.0%	31.8%	26.8%	26.9%	23.6%	26.6%	22.8%
F9	5.6%	5.3%	5.7%	6.6%	***4.7%***	13.6%	6.5%	6.2%	16.0%	5.9%
H1	9.3%	***8.4%***	8.9%	11.5%	9.6%	18.6%	11.3%	8.9%	20.9%	9.7%
H2	7.5%	9.5%	9.0%	7.9%	8.1%	10.1%	7.8%	8.6%	10.8%	***7.1%***
H3	***5.2%***	6.3%	6.6%	8.3%	6.0%	9.3%	8.1%	6.3%	10.6%	6.4%
H4	***5.8%***	8.3%	6.4%	7.1%	6.8%	9.4%	6.6%	6.6%	11.5%	6.3%
H6	***7.1%***	***7.1%***	7.5%	9.1%	7.7%	10.9%	9.4%	8.2%	18.5%	7.8%
JBc	20.6%	***20.4%***	20.6%	25.5%	21.5%	31.1%	24.7%	23.2%	27.8%	21.7%
JCf	20.7%	23.4%	***19.5%***	28.8%	21.2%	39.9%	28.5%	23.4%	34.8%	25.3%
JDa	***8.1%***	8.2%	8.5%	10.8%	8.4%	15.4%	10.5%	9.1%	15.6%	9.3%
JDb	6.7%	6.9%	7.5%	10.0%	***6.2%***	22.6%	10.0%	7.1%	27.6%	8.0%
JDc	***7.1%***	8.4%	8.9%	9.2%	9.5%	13.3%	8.8%	8.4%	12.5%	7.5%
JDd	21.9%	25.2%	***18.2%***	21.6%	22.6%	24.8%	20.8%	20.4%	27.5%	19.8%
JDe	6.7%	5.9%	6.2%	6.7%	***5.4%***	10.9%	6.4%	5.6%	5.7%	5.6%
JEa	***6.7%***	8.3%	10.1%	11.6%	10.7%	15.5%	10.4%	9.2%	18.0%	10.2%
JEb	***6.0%***	6.3%	8.0%	11.5%	10.1%	16.0%	11.3%	7.9%	17.4%	10.3%
L3a	17.5%	14.8%	14.7%	15.7%	14.7%	17.8%	16.3%	14.7%	***14.1%***	15.3%
L3b	***22.1%***	23.6%	24.7%	23.5%	23.6%	33.4%	25.4%	24.3%	28.0%	24.3%
L3c	7.8%	22.2%	7.7%	11.6%	11.0%	19.3%	10.8%	***6.9%***	20.9%	9.5%
L4	7.4%	7.2%	9.6%	9.0%	***5.6%***	13.8%	8.8%	6.3%	12.4%	6.5%
L5a	7.1%	7.7%	7.6%	7.1%	7.6%	11.9%	7.2%	6.9%	9.4%	***6.4%***
L5b	9.7%	10.8%	10.8%	13.6%	16.7%	22.7%	12.1%	***7.9%***	21.6%	11.4%
L5c	8.9%	9.0%	***8.2%***	11.2%	9.1%	16.9%	10.7%	***8.2%***	11.2%	8.6%
L7a/b/c/d/f	6.8%	8.6%	6.8%	7.9%	6.8%	11.1%	7.5%	8.2%	11.0%	***6.5%***
L7e	6.6%	6.7%	***6.2%***	8.2%	6.5%	15.7%	7.7%	7.4%	16.1%	7.3%
L8a	5.6%	***5.0%***	7.1%	8.2%	6.5%	16.9%	8.2%	6.3%	20.2%	7.3%
L8b	***8.1%***	11.4%	9.8%	16.5%	13.0%	22.8%	15.9%	9.4%	24.8%	13.0%
M7B	13.1%	13.1%	11.7%	14.3%	11.8%	20.1%	14.0%	12.5%	14.6%	***11.2%***
P6	6.5%	7.3%	***6.3%***	9.1%	7.0%	10.6%	8.7%	7.0%	10.3%	8.0%
2 months	8.4%	***5.1%***	5.2%	6.6%	5.3%	9.4%	6.4%	5.5%	9.3%	5.5%
Long	7.7%	***4.8%***	***4.8%***	6.0%	***4.8%***	8.3%	5.8%	5.4%	8.5%	5.1%
A2	5.5%	4.6%	4.3%	5.0%	3.9%	7.0%	4.7%	***3.8%***	8.9%	4.0%
A5	4.6%	5.2%	5.2%	***4.3%***	5.4%	6.3%	4.4%	4.4%	4.8%	4.7%
A99	4.2%	3.4%	3.9%	4.9%	3.7%	7.7%	4.9%	***3.3%***	5.9%	3.9%
B2	5.2%	5.4%	5.4%	5.8%	5.7%	8.4%	5.7%	***4.9%***	10.7%	***4.9%***
B4	6.0%	***4.2%***	5.2%	4.5%	5.7%	6.1%	4.8%	4.7%	4.8%	4.4%
B99	4.6%	4.8%	4.6%	***3.5%***	4.9%	4.1%	4.0%	3.6%	3.7%	***3.5%***
C7	4.0%	4.0%	3.3%	5.1%	***3.1%***	8.4%	4.5%	4.8%	9.0%	4.7%
C99	3.4%	3.2%	3.2%	4.8%	***2.5%***	8.2%	4.4%	4.4%	8.7%	4.4%
CA	5.3%	4.0%	3.5%	3.5%	***3.2%***	4.4%	3.5%	3.7%	4.3%	3.3%
CC	4.1%	4.1%	3.4%	5.0%	***3.1%***	8.6%	4.5%	5.0%	9.0%	4.7%
D1	4.1%	3.7%	3.6%	3.9%	3.5%	4.2%	3.9%	***3.2%***	4.6%	3.4%
D2	3.7%	3.6%	3.4%	4.0%	3.3%	6.1%	3.9%	***3.2%***	6.2%	3.4%

(*continued on next page*)

Table B.12 (*continued*).

D3	5.8%	4.6%	***4.4%***	6.7%	4.5%	10.2%	5.8%	5.3%	12.9%	5.5%
D99	3.1%	3.1%	3.3%	3.7%	3.1%	5.5%	3.6%	***2.9%***	6.6%	3.0%
E1	8.2%	5.7%	6.1%	6.6%	6.4%	6.4%	6.1%	***5.3%***	8.2%	5.8%
E2	9.9%	6.5%	***6.3%***	7.3%	7.3%	8.7%	7.5%	6.6%	9.9%	6.4%
E99	6.6%	***3.6%***	***3.6%***	4.2%	3.7%	4.1%	4.1%	4.1%	5.0%	***3.6%***
F1	8.3%	4.5%	4.4%	4.8%	4.2%	5.4%	4.4%	4.4%	4.7%	***4.0%***
F2	4.5%	1.8%	2.0%	2.3%	2.4%	3.1%	2.4%	***1.7%***	3.3%	2.1%
F3	5.5%	***3.3%***	4.2%	5.4%	3.7%	7.7%	5.2%	3.6%	8.1%	4.3%
F5	11.1%	5.8%	5.8%	7.0%	6.1%	7.9%	7.0%	6.6%	9.6%	***5.5%***
F6	11.8%	***4.1%***	5.7%	6.0%	5.6%	6.9%	6.2%	4.7%	6.1%	5.1%
F8	9.3%	***3.1%***	3.5%	4.9%	3.7%	8.2%	5.3%	3.7%	4.6%	4.0%
F99	5.2%	3.7%	3.5%	3.5%	3.5%	5.2%	3.7%	3.6%	6.1%	***3.0%***
G99	54.8%	18.4%	19.0%	30.2%	***18.1%***	38.7%	27.5%	20.4%	43.1%	23.8%
J0	4.8%	3.6%	2.8%	4.4%	***2.7%***	6.7%	4.2%	2.9%	6.7%	2.9%
J4	6.3%	3.2%	***3.0%***	4.3%	3.5%	6.4%	4.1%	3.2%	5.8%	3.5%
J6	5.5%	3.1%	3.5%	5.3%	***2.9%***	8.0%	5.0%	3.2%	8.7%	3.6%
J7	5.7%	4.4%	3.5%	3.9%	3.8%	5.6%	3.7%	3.9%	4.8%	***2.8%***
J9	4.2%	3.6%	3.1%	4.0%	3.0%	6.5%	4.0%	3.0%	5.8%	***2.6%***
J99	4.5%	2.3%	2.3%	4.6%	***2.1%***	7.9%	4.6%	2.6%	7.5%	3.0%
JA	4.9%	***2.0%***	2.3%	4.9%	2.2%	8.3%	4.8%	3.3%	8.8%	3.5%
L1	11.7%	10.8%	10.3%	10.0%	***8.7%***	10.5%	9.1%	9.7%	11.5%	9.0%
L2a	8.2%	5.3%	5.5%	5.3%	5.6%	5.7%	5.6%	***4.1%***	5.3%	4.7%
L2b	16.2%	9.5%	9.4%	12.7%	8.5%	16.7%	12.6%	***8.1%***	18.8%	10.6%
L2c	5.8%	***2.3%***	***2.3%***	5.0%	2.7%	8.4%	4.8%	3.6%	8.4%	3.8%
L99	4.8%	***2.2%***	2.4%	4.0%	2.4%	7.8%	3.8%	3.4%	7.1%	3.2%
LA	5.3%	***2.3%***	2.9%	4.3%	2.7%	8.5%	4.0%	3.6%	7.3%	3.5%
LS	4.7%	***3.2%***	***3.2%***	4.7%	3.4%	8.7%	4.6%	3.6%	9.1%	4.0%
M1	12.0%	9.1%	***8.7%***	9.2%	9.1%	13.4%	9.3%	12.6%	12.2%	9.1%
M2	11.1%	9.6%	9.9%	10.9%	9.0%	12.6%	9.5%	9.0%	11.2%	***8.3%***
M8	26.1%	15.7%	15.2%	***14.7%***	18.3%	19.1%	16.1%	27.1%	16.4%	15.5%
M99	11.4%	7.8%	8.0%	9.0%	***7.4%***	11.8%	8.5%	10.4%	10.6%	7.6%
P1	6.7%	***4.3%***	4.7%	5.1%	4.5%	6.4%	4.9%	5.5%	7.3%	4.9%
P2	4.9%	***3.4%***	***3.4%***	5.0%	3.5%	6.8%	4.8%	3.7%	7.6%	3.7%
P5	4.8%	3.5%	***3.1%***	5.4%	3.7%	7.7%	5.2%	4.2%	9.9%	4.0%
P99	4.8%	***2.7%***	2.8%	3.9%	3.0%	5.6%	3.8%	3.9%	5.0%	3.7%
S2	5.2%	***2.0%***	2.8%	5.1%	2.4%	7.4%	4.9%	4.5%	5.9%	3.7%
S3	3.0%	2.7%	3.0%	3.7%	***2.6%***	6.4%	3.3%	3.2%	7.4%	3.1%
T99	4.4%	***1.9%***	2.6%	4.7%	2.1%	6.6%	4.6%	4.0%	5.9%	3.3%
Medium	8.0%	***4.4%***	4.8%	6.3%	4.7%	9.1%	6.1%	5.0%	8.7%	5.2%
A3	5.4%	***4.2%***	4.7%	6.2%	***4.2%***	9.4%	5.9%	4.6%	6.9%	4.9%
A4	6.4%	3.0%	4.3%	4.5%	4.6%	6.5%	4.4%	***2.6%***	4.6%	3.6%
B3	10.0%	6.2%	6.2%	9.4%	6.6%	15.9%	8.7%	***5.7%***	17.4%	7.4%
C10	5.1%	***3.3%***	3.7%	5.9%	***3.3%***	8.5%	5.6%	4.7%	9.7%	5.3%
C8	8.8%	***6.4%***	6.6%	8.8%	6.5%	10.6%	8.2%	8.1%	10.0%	7.2%
C9	6.4%	5.0%	4.8%	7.0%	***4.4%***	11.2%	7.1%	5.4%	13.4%	6.1%
CB	4.6%	***3.1%***	***3.1%***	5.8%	***3.1%***	8.7%	5.1%	4.5%	11.0%	5.0%
F7	6.6%	4.2%	3.3%	4.8%	***3.1%***	6.5%	4.8%	3.8%	6.1%	3.7%
H99	3.9%	***1.9%***	2.0%	3.4%	***1.9%***	7.2%	3.1%	2.8%	8.0%	2.6%
JB	10.2%	3.4%	***3.0%***	5.9%	3.4%	5.3%	6.0%	4.8%	5.8%	4.1%
JBa	12.5%	***3.5%***	***3.5%***	7.0%	4.0%	7.1%	6.7%	4.9%	7.0%	4.0%
JBb	11.5%	***6.3%***	7.1%	8.8%	6.7%	12.9%	9.1%	7.1%	11.8%	7.5%
JC	6.8%	4.1%	4.6%	4.1%	4.5%	5.7%	4.4%	5.1%	4.4%	***4.0%***
JCc	10.5%	8.5%	12.0%	8.2%	8.9%	9.3%	8.7%	8.0%	9.3%	***7.7%***
JCd	8.5%	5.6%	***5.5%***	6.0%	***5.5%***	7.1%	6.1%	7.0%	6.0%	5.9%
JCe	7.1%	***4.7%***	6.2%	5.2%	6.2%	6.6%	5.2%	5.2%	5.3%	4.9%
JD	6.3%	3.5%	***3.0%***	5.6%	***3.0%***	9.5%	5.6%	4.0%	8.0%	4.3%
JE	6.0%	***2.2%***	2.7%	6.8%	2.6%	10.4%	6.5%	4.2%	11.2%	4.9%
JF	5.0%	3.0%	***2.7%***	4.5%	3.0%	7.8%	4.2%	3.9%	8.8%	3.6%
L3	13.2%	***7.0%***	7.8%	9.9%	8.6%	12.4%	9.3%	***7.0%***	13.1%	8.0%
L5	7.4%	3.2%	4.5%	4.2%	4.5%	6.3%	4.3%	***2.7%***	7.3%	3.6%
L6	7.5%	***4.3%***	4.4%	7.2%	4.5%	11.1%	6.6%	5.8%	10.1%	5.9%
L6g	8.6%	***4.3%***	5.0%	8.0%	5.2%	11.2%	7.4%	6.2%	10.0%	6.7%
L6h	6.2%	3.6%	3.6%	5.2%	***3.5%***	8.7%	5.0%	3.8%	8.1%	4.2%
L6i	11.1%	***5.7%***	6.9%	8.9%	6.5%	12.8%	8.9%	5.8%	7.5%	6.7%
L7	5.8%	3.0%	***2.9%***	4.9%	***2.9%***	6.6%	4.7%	4.0%	7.4%	3.8%
L8	5.6%	2.9%	2.7%	5.3%	***2.5%***	13.3%	4.7%	3.6%	16.2%	4.0%
L9	13.4%	5.6%	6.9%	7.7%	6.1%	11.1%	7.4%	***5.3%***	9.7%	6.5%
M6	5.7%	4.1%	4.5%	3.7%	4.8%	5.2%	3.9%	4.7%	4.1%	***3.4%***
M7A	14.0%	***4.5%***	5.3%	6.3%	5.0%	8.2%	6.1%	5.7%	5.9%	5.0%
P4	6.4%	***5.1%***	5.5%	5.5%	5.5%	7.8%	5.7%	***5.1%***	7.0%	5.2%

(*continued on next page*)

Table B.12 (*continued*).

Short	9.9%	6.5%	***6.3%***	8.0%	6.6%	11.5%	7.8%	***6.3%***	11.5%	6.8%
F10	15.8%	20.7%	15.0%	15.3%	18.9%	15.6%	15.1%	14.0%	17.5%	***13.7%***
F9	5.6%	3.4%	4.0%	5.2%	***3.0%***	9.3%	5.1%	4.2%	11.2%	4.5%
H1	9.3%	4.1%	3.9%	6.1%	5.2%	11.5%	6.3%	***3.8%***	12.5%	4.5%
H2	7.5%	3.7%	***3.3%***	5.0%	3.4%	6.2%	5.2%	3.9%	6.9%	4.1%
H3	5.2%	***2.5%***	2.9%	5.2%	3.0%	5.3%	5.2%	3.9%	5.6%	4.0%
H4	5.8%	4.4%	***4.1%***	5.0%	4.3%	6.7%	4.9%	***4.1%***	7.7%	4.2%
H6	7.1%	***3.4%***	3.9%	5.3%	3.5%	7.5%	5.3%	4.2%	12.4%	4.4%
JBc	20.6%	14.0%	14.8%	17.5%	13.8%	24.0%	16.3%	***12.6%***	18.5%	13.7%
JCf	20.7%	***14.1%***	17.1%	21.2%	19.5%	26.3%	19.0%	14.2%	23.6%	17.8%
JDa	8.1%	4.2%	***3.6%***	6.1%	4.3%	9.3%	6.4%	4.6%	10.2%	5.2%
JDb	6.7%	***3.9%***	5.3%	6.6%	4.0%	15.7%	6.3%	4.4%	18.0%	5.6%
JDc	7.1%	6.6%	7.2%	***5.5%***	6.3%	9.0%	5.6%	8.2%	8.5%	5.8%
JDd	21.9%	***11.7%***	12.2%	13.5%	12.2%	16.6%	14.2%	13.1%	14.8%	12.3%
JDe	6.7%	***4.6%***	5.8%	4.9%	4.8%	8.4%	4.8%	5.3%	4.9%	4.7%
JEa	6.7%	***4.5%***	5.4%	6.7%	5.0%	9.5%	6.4%	6.0%	11.3%	6.2%
JEb	6.0%	4.2%	3.8%	7.1%	***3.0%***	10.1%	7.4%	5.0%	10.7%	5.9%
L3a	17.5%	8.5%	7.8%	9.2%	9.1%	10.9%	8.6%	***7.7%***	10.0%	8.5%
L3b	22.1%	15.1%	14.7%	16.7%	16.4%	21.6%	17.4%	***14.5%***	21.2%	16.3%
L3c	7.8%	9.8%	***2.8%***	5.7%	4.3%	9.2%	5.5%	3.4%	13.0%	4.2%
L4	7.4%	4.0%	4.2%	4.8%	3.7%	9.1%	4.7%	***3.5%***	7.7%	3.7%
L5a	7.1%	***2.9%***	4.0%	3.8%	3.5%	6.0%	3.7%	***2.9%***	5.2%	3.3%
L5b	9.7%	3.9%	5.2%	7.5%	9.0%	13.4%	6.9%	***3.1%***	11.3%	5.5%
L5c	8.9%	4.8%	4.4%	6.4%	4.9%	9.5%	6.4%	***4.3%***	7.5%	4.8%
L7a/b/c/d/f	6.8%	3.2%	3.0%	5.1%	***2.7%***	6.7%	4.9%	3.9%	6.9%	3.7%
L7e	6.6%	4.5%	***3.9%***	5.8%	4.4%	11.4%	5.6%	5.0%	11.2%	5.2%
L8a	5.6%	3.4%	***3.1%***	4.5%	3.2%	11.3%	4.6%	3.3%	12.9%	3.7%
L8b	8.1%	6.2%	***5.7%***	10.2%	6.3%	14.7%	10.0%	6.1%	16.4%	8.3%
M7B	13.1%	9.2%	7.7%	9.2%	***6.7%***	11.9%	8.4%	10.1%	9.3%	7.5%
P6	6.5%	4.3%	***4.1%***	5.8%	4.3%	6.4%	5.7%	4.7%	6.0%	4.9%

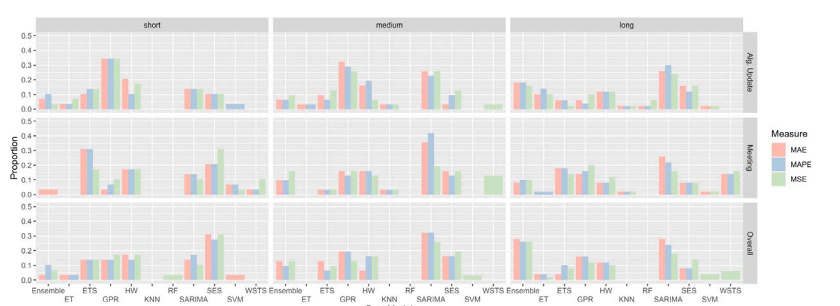

Fig. C.5. The bar chart provides insight into the comparative performance of forecasting methods across different time series lengths and evaluation metrics. It depicts the distribution of the best-performing forecasts across 110 WSTS product categories, categorized by Mean Squared Error (MSE) shown in red, Mean Absolute Percentage Error (MAPE) in blue, and Mean Absolute Error (MAE) in green. Each column represents a different time series length category: the left column represents short time series, the middle column represents medium-length series, and the rightmost column represents long series (those with full available history). The rows of the chart are arranged by algorithmic update, expert forecasts ("Meeting"), and overall performance, following the structure of Tables 4 and 6.

Table B.13
Mean absolute errors (MAE) for the 110 examined WSTS product categories, which are organized by time series length and presented for the 3-month forecasts (first) and 2-month forecast (second). Averages are contained in the long, medium, short, 3-months, and 2-months rows respectively. The best result in each row, i.e. for each time series, is printed bold and italic.

ID	WSTS	SARIMA	ETS	ET	GPR	KNN	RF	SES	SVM	Ensemble
3 months	3.12	3.01	3.10	4.43	***2.97***	6.93	4.27	3.59	6.59	3.44
Long	5.24	4.92	5.18	7.63	***4.83***	12.03	7.35	6.18	11.27	5.82
A2	***0.11***	0.13	***0.11***	0.16	0.14	0.26	0.15	0.12	0.34	0.13
A5	***0.09***	0.14	0.13	0.10	0.12	0.18	0.11	0.12	0.16	0.11
A99	***0.34***	0.43	0.44	0.62	0.40	1.03	0.62	0.43	0.90	0.48
B2	0.14	***0.13***	0.14	0.24	0.17	0.36	0.23	0.15	0.44	0.17
B4	***0.13***	0.20	0.17	0.16	0.16	0.21	0.17	0.15	0.17	0.16
B99	***0.25***	0.36	0.33	0.28	0.32	0.32	0.30	0.31	0.28	0.29
C7	***0.94***	1.15	0.98	1.65	1.04	3.16	1.41	1.10	3.45	1.40
C99	***1.58***	1.79	1.59	3.08	1.62	5.64	2.66	1.79	6.34	2.41
CA	0.26	0.26	***0.23***	0.26	0.25	0.37	0.26	0.28	0.35	0.27
CC	***1.03***	1.20	***1.03***	1.76	1.12	3.45	1.55	1.25	3.60	1.50
D1	***0.15***	0.19	0.18	0.20	0.20	0.26	0.19	0.18	0.24	0.19
D2	***0.14***	0.18	0.18	0.24	0.19	0.33	0.25	0.15	0.38	0.19
D3	0.12	***0.11***	0.13	0.19	0.13	0.33	0.17	***0.11***	0.45	0.16
D99	***0.30***	0.47	0.38	0.56	0.46	0.81	0.51	0.36	1.04	0.43
E1	***0.07***	0.08	0.09	0.09	***0.07***	0.08	0.09	0.08	0.10	0.08
E2	***0.09***	***0.09***	***0.09***	0.11	0.10	0.13	0.11	***0.09***	0.14	0.10
E99	***0.12***	0.13	0.13	0.13	***0.12***	0.15	***0.12***	0.14	0.15	0.13
F1	0.08	0.06	0.06	0.07	***0.05***	0.09	0.07	0.06	0.10	0.06
F2	1.34	1.18	1.17	1.20	1.36	1.56	1.22	***1.08***	1.71	1.17
F3	0.27	0.29	0.26	0.36	0.27	0.52	0.34	***0.25***	0.62	0.28
F5	5.02	5.27	5.23	4.78	5.20	5.32	4.91	7.17	5.98	***4.40***
F6	0.69	0.57	0.62	0.55	0.65	0.68	0.59	0.59	0.55	***0.54***
F8	0.15	***0.11***	0.13	0.15	0.13	0.23	0.14	***0.11***	0.14	0.12
F99	***5.27***	5.67	7.28	6.71	6.83	7.80	6.68	8.42	10.16	5.71
G99	0.07	0.07	0.08	0.07	0.08	***0.06***	0.07	0.07	0.08	0.07
J0	3.24	***3.17***	3.86	5.55	3.41	8.51	5.41	3.70	8.21	4.17
J4	***0.41***	0.42	0.44	0.54	***0.41***	0.82	0.48	0.44	0.74	0.47
J6	***2.34***	2.58	2.92	4.02	2.37	6.35	3.91	2.73	6.02	3.07
J7	0.54	0.71	***0.52***	0.77	0.58	0.98	0.74	0.60	0.87	0.66
J9	***0.39***	0.51	0.49	0.68	0.48	1.04	0.62	0.42	1.05	0.48
J99	7.54	***6.82***	8.06	13.25	7.55	22.57	13.55	9.18	23.72	10.31
JA	4.82	***4.34***	4.90	8.05	4.83	14.19	7.79	5.86	15.31	6.22
L1	***0.02***	***0.02***	***0.02***	***0.02***	***0.02***	***0.02***	***0.02***	***0.02***	***0.02***	***0.02***
L2a	0.29	0.30	0.34	0.28	***0.25***	0.33	0.28	0.28	0.29	0.26
L2b	***0.03***	***0.03***	***0.03***	***0.03***	***0.03***	0.04	***0.03***	***0.03***	0.04	***0.03***
L2c	2.54	***2.37***	2.41	4.33	2.86	7.14	3.73	2.89	7.68	3.31
L99	16.95	***12.38***	15.58	22.30	14.32	41.44	20.87	18.34	41.94	17.83
LA	14.34	***9.99***	13.15	17.50	12.99	35.16	15.81	14.65	31.89	13.93
LS	***3.63***	5.08	4.45	6.10	4.92	11.26	5.63	4.20	13.37	5.24
M1	22.04	24.78	20.41	27.80	***19.65***	42.46	28.39	23.18	37.20	21.69
M2	0.10	0.10	0.09	0.12	0.09	0.13	0.12	***0.08***	0.13	0.09
M8	0.36	0.34	0.37	0.31	0.33	***0.30***	0.36	0.49	0.32	0.32
M99	34.16	36.70	28.63	40.73	***27.69***	63.70	41.84	34.76	58.82	32.47
P1	***8.12***	9.11	10.55	10.41	9.40	11.84	9.95	10.80	13.19	9.88
P2	***2.38***	2.98	2.98	3.95	2.99	6.11	3.70	2.64	7.30	3.35
P5	***2.05***	2.69	2.69	3.73	3.03	6.01	3.55	2.25	7.88	3.14
P99	***8.74***	9.41	11.14	11.83	9.16	15.93	10.90	11.38	13.06	10.75
S2	52.17	***41.34***	49.89	80.01	43.02	123.26	76.37	63.37	107.91	56.24
S3	***2.16***	3.13	2.86	3.78	2.57	7.51	3.11	2.52	8.91	3.11
T99	53.76	***46.24***	50.85	91.83	47.35	141.27	87.40	69.58	119.65	63.31
Medium	***1.29***	1.33	1.33	1.79	1.36	2.74	1.71	1.40	2.82	1.47
A3	0.21	0.19	0.25	0.34	***0.17***	0.49	0.30	0.24	0.40	0.23
A4	***0.04***	***0.04***	***0.04***	***0.04***	***0.04***	0.05	***0.04***	***0.04***	***0.04***	***0.04***
B3	***0.06***	0.08	***0.06***	0.08	0.07	0.13	0.08	0.07	0.14	0.07
C10	***0.60***	0.76	***0.60***	0.90	0.63	1.56	0.85	0.64	1.74	0.67
C8	0.15	0.15	***0.14***	0.22	***0.14***	0.30	0.20	0.15	0.32	0.17
C9	0.28	0.27	0.25	0.42	***0.24***	0.75	0.38	0.25	0.93	0.30

(*continued on next page*)

Table B.13 (*continued*).

CB	0.76	0.92	***0.72***	1.21	***0.72***	2.32	1.05	0.75	2.67	0.88
F7	0.17	***0.16***	***0.16***	0.22	***0.16***	0.27	0.22	0.18	0.27	0.17
H99	***1.65***	1.72	1.98	2.45	1.84	4.70	2.36	1.87	5.53	2.05
JB	0.58	0.71	0.64	0.64	***0.56***	0.61	0.61	0.67	0.63	0.58
JBa	0.40	0.42	***0.37***	0.41	0.38	0.42	0.40	0.41	0.41	0.38
JBb	***0.28***	0.34	0.33	0.43	0.34	0.58	0.43	0.36	0.51	0.35
JC	0.45	0.47	0.56	0.48	0.54	0.59	***0.44***	0.54	0.46	0.47
JCc	***0.12***	0.14	0.17	0.19	0.19	0.20	0.18	0.13	0.23	0.16
JCd	***0.33***	0.43	0.37	0.39	0.42	0.45	0.37	0.38	0.36	0.37
JCe	***0.08***	0.10	0.10	0.10	0.09	0.15	0.10	0.10	0.09	0.10
JD	3.16	3.17	3.14	4.53	***2.92***	7.57	4.59	3.45	7.69	3.57
JE	1.57	1.56	***1.54***	2.87	1.71	4.85	2.81	1.65	5.79	2.36
JF	***0.43***	0.53	0.51	0.69	0.46	1.31	0.64	0.53	1.45	0.60
L3	***2.54***	2.65	2.65	3.22	2.83	4.97	3.38	2.57	4.63	2.92
L5	2.20	2.00	2.40	2.95	2.50	4.45	3.35	***1.77***	4.31	2.55
L6	***6.43***	6.57	6.72	9.13	6.93	14.40	8.43	7.44	13.71	7.78
L6g	5.61	***5.38***	6.04	8.25	6.58	11.14	7.61	6.31	10.80	7.05
L6h	0.90	***0.83***	0.90	1.20	0.88	2.05	1.18	0.87	1.93	0.99
L6i	0.54	***0.49***	0.60	0.66	0.57	0.86	0.63	0.51	0.62	0.55
L7	7.77	8.77	7.49	10.36	7.60	14.48	9.36	8.99	15.96	***7.48***
L8	0.78	***0.76***	0.86	1.26	0.86	2.95	1.16	0.91	3.61	1.04
L9	0.48	***0.42***	0.48	0.72	0.80	0.96	0.65	0.43	0.96	0.64
M6	0.16	0.17	0.22	0.16	0.16	0.22	0.16	***0.15***	0.20	***0.15***
M7A	0.76	0.38	0.45	0.56	***0.37***	0.58	0.52	0.51	0.52	0.44
P4	***0.44***	0.55	0.57	0.54	0.57	0.62	0.52	0.49	0.48	0.50
Short	***1.42***	1.52	***1.42***	1.73	1.47	2.62	1.68	1.48	2.56	1.44
F10	***0.64***	1.04	0.77	1.16	1.33	1.28	1.18	0.99	1.25	1.00
F9	0.34	0.32	0.34	0.42	***0.30***	0.94	0.41	0.38	1.01	0.38
H1	0.49	***0.44***	0.48	0.65	0.50	1.06	0.64	0.48	1.16	0.54
H2	0.50	0.62	0.59	0.53	0.53	0.68	0.53	0.57	0.72	***0.47***
H3	***0.39***	0.46	0.50	0.63	0.45	0.72	0.61	0.46	0.83	0.48
H4	***0.31***	0.44	0.34	0.39	0.36	0.53	0.36	0.34	0.65	0.34
H6	***1.32***	1.36	1.46	1.69	1.51	2.05	1.73	1.50	3.81	1.49
JBc	***0.02***	***0.02***	***0.02***	***0.02***	***0.02***	***0.02***	***0.02***	***0.02***	***0.02***	***0.02***
JCf	***0.05***	***0.05***	***0.05***	***0.05***	***0.05***	0.09	0.06	***0.05***	0.06	***0.05***
JDa	***2.53***	2.57	2.67	3.49	2.68	5.30	3.43	2.84	5.45	2.99
JDb	0.33	0.34	0.38	0.48	***0.30***	1.15	0.47	0.34	1.46	0.40
JDc	***0.44***	0.53	0.57	0.63	0.61	0.91	0.59	0.53	0.86	0.50
JDd	0.52	0.64	***0.45***	0.59	0.57	0.70	0.57	0.52	0.81	0.53
JDe	0.30	0.27	0.28	0.30	0.25	0.52	0.28	0.26	***0.24***	0.26
JEa	***0.21***	0.25	0.32	0.37	0.34	0.54	0.33	0.30	0.68	0.33
JEb	1.39	***1.33***	1.84	2.81	2.41	4.25	2.76	1.78	4.57	2.48
L3a	1.19	***1.02***	***1.02***	1.11	1.03	1.28	1.11	***1.02***	1.04	1.04
L3b	***1.63***	1.75	1.71	1.88	1.80	2.73	2.11	1.70	2.38	1.89
L3c	0.36	0.79	0.34	0.52	0.46	0.88	0.49	***0.32***	1.07	0.42
L4	0.72	0.69	0.99	0.87	***0.53***	1.45	0.84	0.60	1.36	0.62
L5a	1.26	1.32	1.33	1.26	1.35	2.23	1.27	1.18	1.71	***1.12***
L5b	0.91	0.98	1.01	1.32	1.64	2.35	1.13	***0.71***	2.12	1.10
L5c	0.18	0.18	***0.16***	0.21	0.18	0.31	0.20	***0.16***	0.22	0.17
L7a/b/c/d/f	7.20	8.80	7.26	8.78	7.14	12.73	8.21	8.61	12.72	***7.06***
L7e	2.35	2.29	***2.16***	2.72	2.24	5.55	2.62	2.53	6.08	2.54
L8a	0.53	***0.45***	0.66	0.78	0.58	1.80	0.77	0.59	2.07	0.69
L8b	***0.36***	0.44	0.42	0.71	0.55	1.07	0.68	0.39	1.17	0.55
M7B	14.08	14.25	12.70	15.10	12.47	21.94	14.56	13.18	17.81	***11.64***
P6	0.50	0.53	***0.47***	0.73	0.54	0.82	0.71	0.52	0.84	0.63
2 months	3.12	***1.74***	1.92	2.87	1.80	4.41	2.76	2.56	4.18	2.28
Long	5.24	***2.87***	3.28	4.95	3.02	7.64	4.77	4.51	7.13	3.90
A2	0.11	0.10	0.09	0.10	0.08	0.14	0.09	***0.07***	0.20	0.08
A5	0.09	0.10	0.10	***0.08***	0.11	0.12	0.09	***0.08***	0.09	0.09
A99	0.34	0.29	0.33	0.40	0.32	0.66	0.41	***0.27***	0.51	0.32
B2	0.14	0.15	0.15	0.16	0.16	0.23	0.15	***0.13***	0.29	***0.13***

(*continued on next page*)

Table B.13 (*continued*).

B4	0.13	***0.10***	0.12	***0.10***	0.13	0.13	***0.10***	0.11	***0.10***	***0.10***
B99	0.25	0.28	0.27	***0.20***	0.29	0.24	0.22	0.21	0.21	***0.20***
C7	0.94	1.04	0.84	1.21	***0.79***	2.05	1.06	1.16	2.21	1.12
C99	1.58	1.53	1.48	2.23	***1.18***	3.85	2.02	2.04	4.09	2.04
CA	0.26	0.20	0.17	0.18	***0.16***	0.22	0.17	0.18	0.22	***0.16***
CC	1.03	1.16	0.94	1.32	***0.84***	2.27	1.16	1.30	2.37	1.23
D1	0.15	0.14	0.13	0.15	0.13	0.16	0.15	***0.12***	0.18	0.13
D2	0.14	0.14	0.13	0.15	0.13	0.24	0.15	***0.12***	0.25	0.13
D3	0.12	***0.10***	***0.10***	0.14	***0.10***	0.23	0.13	0.12	0.29	0.12
D99	0.30	0.31	0.33	0.36	0.31	0.55	0.35	***0.29***	0.67	0.30
E1	0.07	***0.05***	0.06	0.06	0.06	0.06	***0.05***	***0.05***	0.07	***0.05***
E2	0.09	***0.06***	***0.06***	0.07	0.07	0.08	0.07	***0.06***	0.10	***0.06***
E99	0.12	***0.07***	***0.07***	0.08	***0.07***	0.08	0.08	0.08	0.10	***0.07***
F1	0.08	0.04	0.04	0.04	0.04	0.05	0.04	0.04	0.04	***0.03***
F2	1.34	***0.54***	0.61	0.71	0.72	0.94	0.74	***0.54***	1.03	0.63
F3	0.27	***0.17***	0.20	0.26	0.18	0.38	0.25	***0.17***	0.40	0.21
F5	5.02	2.83	2.86	3.44	2.94	3.75	3.47	3.26	4.49	***2.74***
F6	0.69	***0.28***	0.38	0.36	0.37	0.41	0.38	0.31	0.34	0.32
F8	0.15	***0.05***	0.06	0.08	0.06	0.14	0.09	0.06	0.08	0.07
F99	5.27	3.87	3.74	3.75	3.68	5.45	3.86	3.89	6.32	***3.21***
G99	0.07	***0.04***	***0.04***	0.05	0.05	0.06	0.05	***0.04***	0.06	0.05
J0	3.24	2.53	1.93	3.08	***1.84***	5.02	2.97	1.95	4.96	2.08
J4	0.41	0.20	***0.19***	0.29	0.22	0.45	0.27	0.20	0.41	0.23
J6	2.34	1.41	1.52	2.35	***1.30***	3.68	2.17	1.40	3.88	1.59
J7	0.54	0.40	0.31	0.38	0.35	0.58	0.36	0.37	0.49	***0.27***
J9	0.39	0.35	0.30	0.38	0.31	0.66	0.38	0.29	0.59	***0.25***
J99	7.54	4.04	3.93	7.90	***3.64***	14.52	7.88	4.48	13.99	5.18
JA	4.82	***2.01***	2.36	4.98	2.27	8.97	4.85	3.32	9.58	3.61
L1	0.02	0.02	0.02	0.02	***0.01***	0.02	***0.01***	0.02	0.02	***0.01***
L2a	0.29	0.19	0.20	0.18	0.20	0.19	0.20	***0.15***	0.18	0.16
L2b	0.03	0.02	0.02	0.02	***0.01***	0.03	0.02	***0.01***	0.03	0.02
L2c	2.54	***1.04***	1.05	2.38	1.21	4.38	2.25	1.67	4.37	1.81
L99	16.95	***7.76***	8.77	14.06	8.26	29.01	13.48	11.80	26.07	11.43
LA	14.34	***6.40***	8.29	11.67	7.47	24.51	11.02	9.63	20.64	9.71
LS	3.63	2.41	***2.32***	3.58	2.43	7.35	3.51	2.64	7.82	3.02
M1	22.04	16.54	***15.55***	17.62	16.82	25.79	18.30	23.26	24.94	16.91
M2	0.10	0.08	0.08	0.09	0.08	0.11	0.08	0.08	0.10	***0.07***
M8	0.36	0.22	0.22	***0.21***	0.26	0.24	0.23	0.41	0.22	0.22
M99	34.16	24.34	24.28	28.90	***23.20***	38.08	27.67	32.93	36.52	24.03
P1	8.12	***5.24***	5.66	6.43	5.42	8.13	6.08	6.82	9.24	6.19
P2	2.38	1.59	***1.58***	2.46	1.62	3.77	2.36	1.82	4.20	1.86
P5	2.05	1.47	***1.31***	2.34	1.53	3.80	2.26	1.83	4.86	1.82
P99	8.74	***4.96***	5.21	7.19	5.41	10.51	7.00	7.23	9.49	6.98
S2	52.17	***21.15***	29.51	52.70	25.16	79.43	50.38	46.24	65.43	39.02
S3	2.16	2.10	2.24	2.72	***2.01***	4.73	2.44	2.32	5.53	2.31
T99	53.76	***23.59***	33.75	59.66	27.24	85.43	56.83	49.92	78.19	42.75
Medium	1.29	***0.70***	0.74	1.15	0.74	1.79	1.09	0.89	1.80	0.93
A3	0.21	***0.16***	0.18	0.24	***0.16***	0.37	0.22	0.17	0.28	0.19
A4	0.04	***0.02***	***0.02***	0.03	0.03	0.04	0.03	***0.02***	0.03	***0.02***
B3	0.06	***0.04***	***0.04***	0.06	***0.04***	0.09	0.05	***0.04***	0.09	0.05
C10	0.60	0.38	0.40	0.68	***0.37***	0.99	0.64	0.54	1.12	0.61
C8	0.15	***0.12***	***0.12***	0.17	***0.12***	0.22	0.16	0.15	0.21	0.14
C9	0.28	0.23	0.21	0.32	***0.19***	0.51	0.32	0.24	0.61	0.27
CB	0.76	0.50	0.49	0.92	***0.47***	1.43	0.80	0.72	1.74	0.79
F7	0.17	0.11	***0.08***	0.13	***0.08***	0.18	0.13	0.10	0.16	0.10
H99	1.65	***0.81***	0.89	1.55	0.85	3.32	1.41	1.21	3.56	1.14
JB	0.58	0.21	***0.18***	0.34	0.21	0.30	0.34	0.29	0.32	0.24
JBa	0.40	***0.12***	***0.12***	0.22	0.14	0.22	0.21	0.17	0.21	0.13
JBb	0.28	***0.16***	0.18	0.25	0.17	0.37	0.25	0.18	0.34	0.20
JC	0.45	***0.27***	0.30	***0.27***	0.30	0.38	0.30	0.34	0.29	***0.27***
JCc	0.12	***0.10***	0.16	***0.10***	0.12	0.12	0.11	***0.10***	0.12	***0.10***
JCd	0.33	0.23	***0.21***	0.23	***0.21***	0.27	0.24	0.28	0.23	0.23
JCe	0.08	***0.06***	0.07	***0.06***	0.07	0.07	***0.06***	***0.06***	0.07	***0.06***
JD	3.16	1.75	***1.55***	2.88	***1.55***	5.25	2.88	2.00	4.52	2.23
JE	1.57	***0.62***	0.77	1.90	0.76	3.20	1.82	1.22	3.41	1.43
JF	0.43	0.28	***0.25***	0.39	0.30	0.72	0.36	0.34	0.86	0.32
L3	2.54	***1.28***	1.43	1.98	1.66	2.76	1.87	***1.28***	3.01	1.57
L5	2.20	0.92	1.30	1.27	1.27	2.09	1.32	***0.81***	2.30	1.09
L6	6.43	***3.92***	4.01	6.11	4.06	9.71	5.72	5.14	9.01	5.20
L6g	5.61	***3.08***	3.59	5.35	3.66	7.60	4.97	4.31	6.91	4.60
L6h	0.90	0.52	0.53	0.75	***0.50***	1.35	0.73	0.55	1.26	0.60
L6i	0.54	***0.30***	0.36	0.44	0.35	0.62	0.44	0.31	0.38	0.34
L7	7.77	4.04	***3.87***	6.92	3.98	9.69	6.68	5.48	10.84	5.33
L8	0.78	0.41	0.36	0.74	***0.34***	1.97	0.65	0.52	2.36	0.57
L9	0.48	0.21	0.31	0.35	0.23	0.55	0.35	***0.20***	0.49	0.27
M6	0.16	0.11	0.13	0.10	0.13	0.14	0.11	0.13	0.12	***0.09***

(*continued on next page*)

Table B.13 (*continued*).

M7A	0.76	***0.26***	0.29	0.38	0.28	0.48	0.37	0.33	0.34	0.29
P4	0.44	***0.36***	0.39	0.39	0.39	0.52	0.40	***0.36***	0.48	0.37
Short	1.42	0.89	0.83	1.12	***0.81***	1.64	1.08	0.99	1.64	0.92
F10	0.64	0.90	0.66	0.67	0.82	0.72	0.66	0.64	0.81	***0.60***
F9	0.34	0.22	0.25	0.32	***0.20***	0.64	0.31	0.26	0.70	0.28
H1	0.49	0.22	***0.21***	0.35	0.28	0.66	0.36	***0.21***	0.70	0.25
H2	0.50	0.25	***0.22***	0.34	0.23	0.41	0.35	0.26	0.47	0.27
H3	0.39	***0.20***	0.23	0.41	0.23	0.41	0.40	0.32	0.44	0.31
H4	0.31	0.24	***0.22***	0.28	0.23	0.38	0.27	***0.22***	0.43	0.23
H6	1.32	***0.61***	0.70	0.96	0.64	1.41	0.94	0.75	2.52	0.79
JBc	0.02	***0.01***	***0.01***	***0.01***	***0.01***	0.02	***0.01***	***0.01***	***0.01***	***0.01***
JCf	0.05	***0.03***	0.04	0.04	0.05	0.06	0.04	***0.03***	0.05	0.04
JDa	2.53	1.34	***1.18***	1.99	1.37	3.13	2.07	1.46	3.56	1.69
JDb	0.33	0.18	0.24	0.30	***0.17***	0.78	0.29	0.19	0.95	0.25
JDc	0.44	0.43	0.46	***0.37***	0.42	0.62	***0.37***	0.53	0.58	0.38
JDd	0.52	***0.29***	0.30	0.35	0.30	0.47	0.37	0.33	0.42	0.32
JDe	0.30	0.21	0.26	0.21	0.21	0.37	0.21	0.24	***0.20***	***0.20***
JEa	0.21	***0.15***	0.17	0.22	0.17	0.33	0.21	0.20	0.43	0.20
JEb	1.39	0.92	0.98	1.79	***0.79***	2.78	1.86	1.28	2.88	1.51
L3a	1.19	0.66	0.59	0.67	0.68	0.79	0.61	***0.58***	0.75	0.63
L3b	1.63	1.12	1.07	1.34	1.24	1.77	1.38	***1.06***	1.74	1.25
L3c	0.36	0.35	***0.13***	0.26	0.19	0.44	0.25	0.16	0.66	0.19
L4	0.72	0.38	0.42	0.48	0.36	0.98	0.48	***0.35***	0.86	0.37
L5a	1.26	0.51	0.69	0.65	0.62	1.10	0.63	***0.50***	0.89	0.56
L5b	0.91	0.36	0.49	0.70	0.80	1.33	0.64	***0.29***	1.09	0.50
L5c	0.18	0.10	***0.09***	0.13	0.10	0.19	0.13	***0.09***	0.15	0.10
L7a/b/c/d/f	7.20	3.29	3.15	5.63	***2.98***	7.67	5.34	4.06	8.02	4.09
L7e	2.35	1.52	***1.37***	1.94	1.51	3.96	1.91	1.72	4.14	1.78
L8a	0.53	***0.29***	***0.29***	0.42	0.30	1.21	0.42	0.30	1.32	0.35
L8b	0.36	0.29	***0.26***	0.47	0.29	0.71	0.46	0.29	0.78	0.38
M7B	14.08	10.31	9.02	10.82	***7.99***	13.65	9.91	12.01	11.40	8.90
P6	0.50	0.31	***0.29***	0.46	0.31	0.49	0.45	0.34	0.48	0.37

B.1. RMSE

See Table B.11.

B.2. MAPE

See Table B.12.

B.3. MAE

See Table B.13.

Appendix C. Best model frequencies

See Figs. C.4 and C.5.

Data availability

World Semiconductor Trade Statistics (WSTS) data and forecasts are available from wsts.org.

References

Agrawal, A., Khavkin, M., & Slonim, J. (2020). Bringing a real-world edge to forecasting: To improve the accuracy of corporate forecasts, build in the physical parameters from company operations. URL https://www.mckinsey.com/capabilities/strategy-and-corporate-finance/our-insights/bringing-a-real-world-edge-to-forecasting (Accessed 24 January 2024).

Ahmed, N. K., Atiya, A. F., Gayar, N. E., & El-Shishiny, H. (2010). An empirical comparison of machine learning models for time series forecasting. *Econometric Reviews*, *29*(5–6), 594–621. http://dx.doi.org/10.1080/07474938.2010.481556, URL http://www.tandfonline.com/doi/abs/10.1080/07474938.2010.481556.

Akay, R. (2022). Caretforecast: Conformal time series forecasting using state of art machine. URL https://CRAN.R-project.org/package=caretForecast.

Alsop, T. (2024a). Semiconductor market revenue worldwide from 1987 to 2024. URL https://www.statista.com/statistics/266973/global-semiconductor-sales-since-1988 (Accessed 19 February 2024).

Alsop, T. (2024b). Semiconductors - statistics & facts. URL https://www.statista.com/topics/1182/semiconductors (Accessed 19 February 2024).

Armstrong, J. S. (2008). Methods to elicit forecasts from groups: Delphi and prediction markets compared. *SSRN Electronic Journal*, http://dx.doi.org/10.2139/ssrn.1153124, URL http://www.ssrn.com/abstract=1153124.

Atanasov, P., Rescober, P., Stone, E., Swift, S. A., Servan-Schreiber, E., Tetlock, P. E., Ungar, L., & Mellers, B. (2015). Distilling the wisdom of crowds: Prediction markets versus prediction polls. *Academy of Management Proceedings*, *2015*(1), 15192. http://dx.doi.org/10.5465/ambpp.2015.15192abstract.

Aubry, M., & Renou-Maissant, P. (2013). Investigating the semiconductor industry cycles. *Applied Economics*, *45*(21), 3058–3067. http://dx.doi.org/10.1080/00036846.2012.697123, URL http://www.tandfonline.com/doi/abs/10.1080/00036846.2012.697123.

Aubry, M., & Renou-Maissant, P. (2014). Semiconductor industry cycles: Explanatory factors and forecasting. *Economic Modelling*, *39*, 221–231. http://dx.doi.org/10.1016/j.econmod.2014.02.039, https://linkinghub.elsevier.com/retrieve/pii/S0264999314000832.

Awad, M., & Khanna, R. (2015). *Efficient learning machines*. Berkeley, CA: A Press, http://dx.doi.org/10.1007/978-1-4302-5990-9, URL http://link.springer.com/10.1007/978-1-4302-5990-9.

Aytac, B., & Wu, S. D. (2013). Characterization of demand for short life-cycle technology products. *Annals of Operations Research*, *203*(1), 255–277. http://dx.doi.org/10.1007/s10479-010-0771-5, URL http://link.springer.com/10.1007/s10479-010-0771-5.

Biau, G., & Scornet, E. (2016). A random forest guided tour. *TEST*, *25*(2), 197–227. http://dx.doi.org/10.1007/s11749-016-0481-7, URL https://link.springer.com/article/10.1007/s11749-016-0481-7.

Box, G. E. P., & COX, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society. Series B. Statistical Methodology*, *26*(2), 211–243. http://dx.doi.org/10.1111/j.2517-6161.1964.tb00553.x.

Breiman, L. (1984). *Classification and regression trees*. The Wadsworth & Brooks/Cole.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*, 5–32.

Brockwell, P. J., & Davis, R. A. (2002). *Introduction to time series and forecasting*. Springer.

Cerqueira, V., Torgo, L., & Mozetič, I. (2020). Evaluating time series forecasting models: an empirical study on performance estimation methods. *Machine Learning*, *109*(11), 1997–2028. http://dx.doi.org/10.1007/s10994-020-05910-7, https://link.springer.com/10.1007/s10994-020-05910-7. arXiv:1905.11744.

Cerqueira, V., Torgo, L., & Soares, C. (2022). A case study comparing machine learning with statistical methods for time series forecasting: size matters. *Journal of Intelligent Information Systems*, *59*(2), 415–433. http://dx.doi.org/10.1007/s10844-022-00713-9, https://link.springer.com/10.1007/s10844-022-00713-9.

Chen, Y.-J., & Chien, C.-F. (2018). An empirical study of demand forecasting of non-volatile memory for smart production of semiconductor manufacturing. *International Journal of Production Research*, *56*(13), 4629–4643. http://dx.doi.org/10.1080/00207543.2017.1421783, https://www.tandfonline.com/doi/full/10.1080/00207543.2017.1421783.

Chien, C.-F., Chen, Y.-J., & Peng, J.-T. (2013). Manufacturing intelligence for semiconductor demand forecast based on technology diffusion and product life cycle. *International Journal of Production Economics*, *128*(2), 496–509. http://dx.doi.org/10.1016/j.ijpe.2010.07.022, URL https://linkinghub.elsevier.com/retrieve/pii/S092552731000263X.

Chien, C.-F., & Lin, K.-Y. (2012). Manufacturing intelligence for Hsinchu Science Park semiconductor sales prediction. *Journal of the Chinese Institute of Industrial Engineers*, *29*(2), 98–110. http://dx.doi.org/10.1080/10170669.2012.660200, http://www.tandfonline.com/doi/abs/10.1080/10170669.2012.660200.

Chow, H. K., & Choy, K. M. (2006). Forecasting the global electronics cycle with leading indicators: A Bayesian VAR approach. *International Journal of Forecasting*, *22*(2), 301–315. http://dx.doi.org/10.1016/j.ijforecast.2005.07.002.

Corder, G., Lafond, E., & Lorenzen, C. (2024). *Worldwide semiconductor market expected to recover strongly and hit $611 billion in 2024*: *Technical Report*, European Semiconductor Industry Association, https://www.eusemiconductors.eu/sites/default/files/ESIA_WSTS_Spring_Forecast2024.pdf (Accessed 01 October 2024).

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, *20*(3), 273–297. http://dx.doi.org/10.1007/BF00994018, URL http://link.springer.com/10.1007/BF00994018.

Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. *Institute of Electrical and Electronics Engineers. Transactions on Information Theory*, *13*(1), 21–27. http://dx.doi.org/10.1109/TIT.1967.1053964.

Crone, S. F., Hibon, M., & Nikolopoulos, K. (2011). Advances in forecasting with neural networks? Empirical evidence from the NN3 competition on time series prediction. *International Journal of Forecasting*, *27*(3), 635–660.

Linstone, H. A. (Ed.), (1975). *The Delphi method: Techniques and applications*. Reading, Mass.: Addison-Wesley.

Drucker, H., Burges, C. J. C., Kaufman, L., Smola, A., & Vapnik, V. (1996). Support vector regression machines. In M. Mozer, M. Jordan, & T. Petsche (Eds.), *Vol. 9*, *Advances in neural information processing systems* (pp. 155–161). MIT Press, URL https://proceedings.neurips.cc/paper_files/paper/1996/file/d38901788c533e8286cb6400b40b386d-Paper.pdf.

European Comission (2023). European chips act: Online factpage. URL https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/european-chips-act_en (Accessed 19 February 2024).

Fix, E., & Hodges, J. L. (1989). Discriminatory analysis. Nonparametric discrimination: Consistency properties. *International Statistical Review / Revue Internationale de Statistique*, *57*(3), 238. http://dx.doi.org/10.2307/1403797.

Foster, R. N. (1986). Working the S-curve: Assessing technological threats. *Research Management*, *29*(4), 17–20. http://dx.doi.org/10.1080/00345334.1986.11756976.

Gardner, E. S. (1985). Exponential smoothing: The state of the art. *Journal of Forecasting*, *4*(1), 1–28. http://dx.doi.org/10.1002/for.3980040103.

Garza, F., Canseco, M. M., Challú, C., & Olivares, K. G. (2022). StatsForecast: Lightning fast forecasting with statistical and econometric models. URL https://github.com/Nixtla/statsforecast. PyCon Salt Lake City, Utah, US 2022.

Geurts, P., Ernst, D., & Wehenkel, L. (2006). Extremely randomized trees. *Machine Learning*, *63*(1), 3–42. http://dx.doi.org/10.1007/s10994-006-6226-1.

Gneiting, T., & Katzfuss, M. (2014). Probabilistic forecasting. *Annual Review of Statistics and Its Application*, *1*(1), 125–151. http://dx.doi.org/10.1146/annurev-statistics-062713-085831, URL https://www.annualreviews.org/doi/10.1146/annurev-statistics-062713-085831.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.

Grinsztajn, L., Oyallon, E., & Varoquaux, G. (2022). Why do tree-based models still outperform deep learning on typical tabular data? In *Thirty-sixth conference on neural information processing systems datasets and benchmarks track* (pp. 507–520).

Huang, H., Pouls, M., Meyer, A., & Pauly, M. (2020). Travel time prediction using tree-based ensembles. In *Computational logistics: 11th international conference, ICCL 2020, enschede, the netherlands, September 28–30, 2020, proceedings 11* (pp. 412–427). Springer.

Hyndman, R. J. (2001). It's time to move from "what" to "why". *International Journal of Forecasting*, *17*(4), 567–570.

Hyndman, R. J. (2020). A brief history of forecasting competitions. *International Journal of Forecasting*, *36*(1), 7–14. http://dx.doi.org/10.1016/j.ijforecast.2019.03.015, https://linkinghub.elsevier.com/retrieve/pii/S016920701930086X.

Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and practice* (2nd ed.). Melbourne: OTexts, URL https://otexts.com/fpp2/.

Hyndman, R. J., & Khandakar, Y. (2008). Automatic time series forecasting: The forecast package for r. *Journal of Statistical Software*, *27*(3), 1–22. http://dx.doi.org/10.18637/jss.v027.i03, URL https://www.jstatsoft.org/article/view/v027i03.

Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, *22*(4), 679–688. http://dx.doi.org/10.1016/j.ijforecast.2006.03.001, URL https://linkinghub.elsevier.com/retrieve/pii/S0169207006000239.

Kapur, P., Panwar, S., & Singh, O. (2019). Modeling two-dimensional technology diffusion process under dynamic adoption rate. *J. Model. Manage.*, *14*(3), 717–737. http://dx.doi.org/10.1108/JM2-06-2018-0088, URL https://www.emerald.com/insight/content/doi/10.1108/JM2-06-2018-0088/full/html.

Karatzoglou, A., Smola, A., Hornik, K., & Zeileis, A. (2004). Kernlab - an S4 package for kernel methods in r. *Journal of Statistical Software*, *11*(9), http://dx.doi.org/10.18637/jss.v011.i09.

Kuhn, M. (2008). Building predictive models in r using the caret package. *Journal of Statistical Software*, *28*(5), http://dx.doi.org/10.18637/jss.v028.i05.

Lee, H. L., Padmanabhan, V., & Whang, S. (1997). The bullwhip effect in supply chains. *MIT Sloan Management Review*, *38*(3), 93–102.

Liu, W.-H. (2005). Determinants of the semiconductor industry cycles. *Journal of Policy Modeling*, *27*(7), 853–866. http://dx.doi.org/10.1016/j.jpolmod.2005.05.009, URL https://linkinghub.elsevier.com/retrieve/pii/S0161893805000852.

Liu, W.-H., & Chyi, Y.-L. (2006). A Markov regime-switching model for the semiconductor industry cycles. *Economic Modelling*, *23*(4), 569–578. http://dx.doi.org/10.1016/j.econmod.2006.02.007, URL https://linkinghub.elsevier.com/retrieve/pii/S0264999306000204.

Lv, S., Kim, H., Zheng, B., & Jin, H. (2018). A review of data mining with big data towards its applications in the electronics industry. *Applied Sciences*, *8*(4), 582. http://dx.doi.org/10.3390/app8040582.

Macher, J. T. (2006). Technological development and the boundaries of the firm: A knowledge-based examination in semiconductor manufacturing. *Management Science*, *52*(6), 826–843. http://dx.doi.org/10.1287/mnsc.1060.0511.

Makridakis, S., & Hibon, M. (2000). The M3-competition: results, conclusions and implications. *International Journal of Forecasting*, *16*(4), 451–476. http://dx.doi.org/10.1016/S0169-2070(00)00057-1, URL https://linkinghub.elsevier.com/retrieve/pii/S0169207000000571.

Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2018). Statistical and machine learning forecasting methods: Concerns and ways forward. In A. R. Hernandez Montoya (Ed.), *PLoS One*, *13*(3), Article e0194889. http://dx.doi.org/10.1371/journal.pone.0194889, URL https://dx.plos.org/10.1371/journal.pone.0194889{%}0A https://dspace.lboro.ac.uk/dspace-jspui/handle/2134/25091 https://dx.plos.org/10.1371/journal.pone.0194889.

Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2022a). M5 accuracy competition: Results, findings, and conclusions. *International Journal of Forecasting*, *38*(4), 1346–1364. http://dx.doi.org/10.1016/j.ijforecast.2021.11.013, https://linkinghub.elsevier.com/retrieve/pii/S0169207021001874.

Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2022b). The M5 competition: Background, organization, and implementation. *International Journal of Forecasting*, *38*(4), 1325–1336. http://dx.doi.org/10.1016/j.ijforecast.2021.07.007, https://linkinghub.elsevier.com/retrieve/pii/S0169207021001187.

Modis, T. (1994). Life cycles: forecasting the rise and fall of almost anything. *The Futurist*, *28*(5), 20–25.

Modis, T. (1999). A second lease on life for technological forecasting. *Technological Forecasting and Social Change*, *62*(1–2), 29–32. http://dx.doi.org/10.1016/S0040-1625(99)00036-0.

Nagao, M. (2019). The semiconductor trend from the front-end view. In *2019 pan Pacific microelectronics symposium (pan Pacific)* (pp. 1–6). IEEE, http://dx.doi.org/10.23919/PanPacific.2019.8696919, URL https://ieeexplore.ieee.org/document/8696919/.

Opitz, D., & Maclin, R. (1999). Popular ensemble methods: An empirical study. *Journal of Artificial Intelligence Research*, *11*, 169–198. http://dx.doi.org/10.1613/jair.614.

Pauly, M., & Kuhlmann, L. (2023). A dynamic systems model for an economic evaluation of sales forecasting methods. *Tehnički Glasnik*, *17*(3), 397–404. http://dx.doi.org/10.31803/tg-20230511175500.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine learning in python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Petropoulos, F., Apiletti, D., Assimakopoulos, V., Babai, M. Z., Barrow, D. K., Ben Taieb, S., Bergmeir, C., Bessa, R. J., Bijak, J., Boylan, J. E., Browell, J., Carnevale, C., Castle, J. L., Cirillo, P., Clements, M. P., Cordeiro, C., Cyrino Oliveira, F. L., De Baets, S., Dokumentov, A., ... Ziel, F. (2022). Forecasting: theory and practice. *International Journal of Forecasting*, *38*(3), 705–871. http://dx.doi.org/10.1016/j.ijforecast.2021.11.001, arXiv:2012.03854.

Probst, P., Wright, M. N., & Boulesteix, A. L. (2019). Hyperparameters and tuning strategies for random forest. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *9*(3), Article e1301. http://dx.doi.org/10.1002/widm.1301, URL https://onlinelibrary.wiley.com/doi/full/10.1002/widm.1301?casa_token=hyjhezUnLksAAAAA%3AvXjQPCMqR3ltQIm6WsK5Pk8y_xXwIosdDRnv9aOTF1KDIiFNUU4SH7UR-4hesN3wdjmy90AYMrvU9eQb. arXiv:1804.03515.

Rajan, G. (2024). OpenAI's Altman in talks to raise funds for chips, AI initiative - WSJ. URL https://www.reuters.com/technology/openais-altman-talks-raise-funds-chips-ai-initiative-wsj-2024-02-09/ (Accessed 19 February 2024).

Satchell, S., & Timmermann, A. (1995). On the optimality of adaptive expectations: Muth revisited. *International Journal of Forecasting*, *11*(3), 407–416. http://dx.doi.org/10.1016/0169-2070(95)00588-7.

Schliep, K., & Hechenbichler, K. (2016). Kknn: Weighted k-nearest neighbors. URL https://CRAN.R-project.org/package=kknn R package version 1.3.1.

Simons, C. (2024). *2024 semiconductor industry outlook: Trends and predictions for a cyclical industry*: *Technical Report*, Deloitte, URL https://www2.deloitte.com/us/en/pages/technology-media-and-telecommunications/articles/semiconductor-industry-outlook.html (Accessed 01 October 2024).

Steinmeister, L., Ramosaj, B., Schröter, L., & Pauly, M. (2023). Testing the limits : A robustness analysis of logistic growth models for life cycle estimation during the COVID-19 pandemic. *Vol. 2*, In *Proceedings of the conference on production systems and logistics: CPSL* (pp. 33–44). http://dx.doi.org/10.15488/15265.

Taylor, M. (2023). The US CHIPS and science act of 2022. *MRS Bulletin*, *48*(9), 874–879. http://dx.doi.org/10.1557/s43577-023-00581-w.

Wang, J. (2023). An intuitive tutorial to Gaussian processes regression. *Computing in Science & Engineering*, 1–8. http://dx.doi.org/10.1109/MCSE.2023.3342149.

Wang, C.-H., & Chen, J.-Y. (2019). Demand forecasting and financial estimation considering the interactive dynamics of semiconductor supply-chain companies. *Computers & Industrial Engineering*, *138*(June), Article 106104. http://dx.doi.org/10.1016/j.cie.2019.106104, https://linkinghub.elsevier.com/retrieve/pii/S036083521930573X.

Wang, W., He, N., Chen, M., & Jia, P. (2024). Freight rate index forecasting with prophet model based on multi-dimensional significant events. *Expert Systems with Applications*, *249*, Article 123451. http://dx.doi.org/10.1016/j.eswa.2024.123451, URL https://www.sciencedirect.com/science/article/pii/S0957417424003166.

Wickramasuriya, S. L., Athanasopoulos, G., & Hyndman, R. J. (2019). Optimal forecast reconciliation for hierarchical and grouped time series through trace minimization. *Journal of the American Statistical Association*, *114*(526), 804–819. http://dx.doi.org/10.1080/01621459.2018.1448825, https://www.tandfonline.com/doi/full/10.1080/01621459.2018.1448825.

Williams, C., & Rasmussen, C. (1995). Gaussian processes for regression. In D. Touretzky, M. Mozer, & M. Hasselmo (Eds.), *Vol. 8*, *Advances in neural information processing systems* (pp. 514–520). MIT Press, URL https://proceedings.neurips.cc/paper_files/paper/1995/file/7cce53cf90577442771720a370c3c723-Paper.pdf.

World Bank (2024). GDP (current US$) - United States, European union, world. URL https://data.worldbank.org/indicator/NY.GDP.MKTP.CD?locations=US-EU-1W (Accessed 19 February 2024).

Wright, M. N., & Ziegler, A. (2017). Ranger : A fast implementation of random forests for high dimensional data in C++ and R. *Journal of Statistical Software*, *77*(1), http://dx.doi.org/10.18637/jss.v077.i01.

WSTS. org (2024). More than 35 years authentic market monitoring by WSTS. URL https://www.wsts.org/ (Accessed 24 January 2024).

Wu, J. Z., & Chien, C. F. (2008). Modeling strategic semiconductor assembly outsourcing decisions based on empirical settings. *OR Spectrum*, *30*(3), 401–430. http://dx.doi.org/10.1007/s00291-007-0120-5.

Wu, G., Hu, Y.-C., Chiu, Y.-J., Jiang, P., & Chi, R. (2024). Forecast combination using grey relational analysis and choquet fuzzy integral for container throughput forecasting. *Expert Systems with Applications*, *252*, Article 124170. http://dx.doi.org/10.1016/j.eswa.2024.124170, URL https://www.sciencedirect.com/science/article/pii/S0957417424010364.

Xu, Q., & Sharma, V. (2017). Ensemble sales forecasting study in semiconductor industry. In P. Perner (Ed.), *Advances in data mining. applications and theoretical aspects* (pp. 31–44). Springer, http://dx.doi.org/10.1007/978-3-319-62701-4_3, URL http://link.springer.com/10.1007/978-3-319-62701-4{_}3.

Yu, Y., Hui, C.-L., & Choi, T.-M. (2012). An empirical study of intelligent expert systems on forecasting of fashion color trend. *Expert Systems with Applications*, *39*(4), 4383–4389. http://dx.doi.org/10.1016/j.eswa.2011.09.153, URL https://www.sciencedirect.com/science/article/pii/S095741741101476X.

Louis Steinmeister was awarded his MS degree in applied mathematics from Missouri S&T, where he was a member of the Applied Computational Intelligence Lab. His research interests include machine learning, statistics, and their application to business and financial problems. Currently, he is pursuing a Ph.D. in Statistics through the Graduate School of Logistics at TU Dortmund University, Germany.

Markus Pauly received his diploma degree in mathematics from the Heinrich-Heine University Düsseldorf (2005), where he also completed his Ph.D. (2008) and his habilitation (2013). From 2014 to 2019 he was full professor for Statistics at Ulm University. Since 2019 he is full professor at the Statistics Department at TU Dortmund University, where he investigates a wide range of statistical and machine learning topics.