

Article

Automobile-Demand Forecasting Based on Trend Extrapolation and Causality Analysis

Zhengzhu Zhang ¹, Haining Chai ¹, Liyan Wu ², Ning Zhang ¹ and Fenghe Wu ^{1,3,*}

¹ College of Mechanical Engineering, Yanshan University, Qinhuangdao 066004, China; zhangzhengzhu@stumail.ysu.edu.cn (Z.Z.); chnwsj@stumail.ysu.edu.cn (H.C.); niza@stumail.ysu.edu.cn (N.Z.)

² Xinfu Group Co., Ltd., Liaocheng 252000, China; liyanwu2875@163.com

³ Hebei Heavy-Duty Intelligent Manufacturing Equipment Technology Innovation Center, Qinhuangdao 066004, China

* Correspondence: risingwu@ysu.edu.cn

Abstract: Accurate automobile-demand forecasting can provide effective guidance for automobile-manufacturing enterprises in terms of production planning and supply planning. However, automobile sales volume is affected by historical sales volume and other external factors, and it shows strong non-stationarity, nonlinearity, autocorrelation and other complex characteristics. It is difficult to accurately forecast sales volume using traditional models. To solve this problem, a forecasting model combining trend extrapolation and causality analysis is proposed and derived from the historical predictors of sales volume and the influence of external factors. In the trend-extrapolation model, the historical predictors of sales series was captured based on the Seasonal Autoregressive Integrated Moving Average (SARIMA) and Polynomial Regression (PR); then, Empirical Mode Decomposition (EMD), a stationarity-test algorithm, and an autocorrelation-test algorithm were introduced to reconstruct the sales sequence into stationary components with strong seasonality and trend components, which reduced the influences of non-stationarity and nonlinearity on the modeling. In the causality-analysis submodel, 31-dimensional feature data were extracted from influencing factors, such as date, macroeconomy, and promotion activities, and a Gradient-Boosting Decision Tree (GBDT) was used to establish the mapping between influencing factors and future sales because of its excellent ability to fit nonlinear relationships. Finally, the forecasting performance of three combination strategies, namely the boosting series, stacking parallel and weighted-average parallel strategies, were tested. Comparative experiments on three groups of sales data showed that the weighted-average parallel combination strategy had the best performance, with loss reductions of 16.81% and 4.68% for data from the number-one brand, 25.60% and 2.79% for data from the number-two brand, and 46.26% and 14.37% for data from the number-three brand compared with the other combination strategies. Other ablation studies and comparative experiments with six basic models proved the effectiveness and superiority of the proposed model.

Keywords: automobile demand; combination forecasting; trend extrapolation; causality analysis



Citation: Zhang, Z.; Chai, H.; Wu, L.; Zhang, N.; Wu, F. Automobile-Demand Forecasting Based on Trend Extrapolation and Causality Analysis. *Electronics* **2024**, *13*, 3294. <https://doi.org/10.3390/electronics13163294>

Academic Editors: Stanley C. Ahalt, Guanfeng Liu, Ping-Feng Pai, Igor Mekterović, Ljiljana Brkić and Marko Horvat

Received: 27 June 2024

Revised: 1 August 2024

Accepted: 17 August 2024

Published: 19 August 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In Sales and Operations Planning (S&OP), automobile-demand forecasting can provide important guidance for supply planning and production planning, which play a key role in achieving coordination among production, supply, and sales in the automotive-manufacturing industry, with significant implications for intelligent manufacturing in the automobile-manufacturing industry [1]. An inaccurate demand forecast may have negative impacts on automobile enterprises, such as oversupply, shortage, or the disruption of production rhythms [2].

With the growth of methods using data-driven analytics, forecasting technology is commonly used in supply-chain management (SCM). Researchers have applied prediction

technologies to SCM across various industries. Table 1 briefly lists the related research. Mitra et al. [3] propose a framework using Gaussian Mixture Model (GMM) clustering and Hierarchical Agglomerative Clustering (HAC) to analyze demand patterns and use a Random Forest algorithm to forecast sales of food and beverages. Fu et al. [4] use K-means to identify the demand status and LSTM to forecast the demand. Van Belle et al. [5] compare explanatory methods with extrapolative methods and argue that LASSO has the best performance in forecasting drug demand. Liu et al. [6] proposed a BO-CNN-LSTM-based method for sales forecasting for stores and supply chains. Moalem et al. [7] forecast electrical-power demand by ELATLBO-LSTM. The studies above prove that demand forecasting provides valuable information for optimal supply-chain management and that an understanding of the demand patterns and influencing factors is essential for forecasting.

Table 1. State-of-the-art forecasting research in the automobile industry.

Method	Case Study	Publishing Author
Gaussian mixture model, Hierarchical agglomerative clustering, Random Forest	Food and beverage	Mitra et al. [3]
K-means, LSTM	Semiconductors	Fu et al. [4]
LASSO	Medicine	Van Belle et al. [5]
Bayesian optimization-CNN-LSTM	Stores and supply chains	Liu et al. [6]
ELATLBO-LSTM	Electrical power	Moalem et al. [7]

Based on actual production needs, the focus of this study is on forecasting the monthly sales of automobiles for the next quarter. For a better understanding of how forecasting technology has been used in the automobile industry, Table 2 briefly lists representative forecasting studies from the automobile industry. Matsumoto et al. [8] examined the performance of the Holt–Winters model and the ARIMA model on the time-series data comprising the sales of 160 types of remanufactured alternators and starters. Gao et al. [9] proposed a novel hybrid method based on particle-swarm optimization and ant-colony optimization (HPA) to forecast automobile sales. Rozanec et al. [10] evaluated 21 baseline, statistical, and machine learning algorithms used to forecast automobile demand. Vijayakumar et al. [11] forecast automobile demand using a combination model composed of the Seagull Optimization-based Holt–Winters (SO-HW) method and a Quantile Regression Neural Network. Ou-Yang et al. [12] forecast car sales by using online sentiment data and the CNN-LSTM method. Sales forecasting is an important method for evaluating automobile demand, and most researchers have pointed out that sales series have the complex characteristics of seasonality, trends, and non-stationarity.

Table 2. State-of-the-art forecasting research in the automobile industry.

Method	Types of Models Used	Publishing Author
Holt–Winters and ARIMA	statistical model	Matsumoto et al. [8]
Polynomial regression with optimization method	statistical model	Gao et al. [9]
21 basic statistic and machine learning models	statistical model/machine learning model	Rozanec et al. [10]
Holt–Winters and QRNN	statistical model/deep learning model	Vijayakumar et al. [11]
CNN-LSTM	deep learning model	Ou-Yang et al. [12]

Common forecasting methods can be categorized into two main groups: trend-extrapolation methods [13,14] and causal-analysis methods [15,16]. Trend-extrapolation methods treat the predicted variable as a time series, extracting sequential patterns from historical data and then extending these patterns to future time points for accurate predictions. Causal-analysis methods assume that future changes in the predicted variable

are influenced by a set of external variables, thus basing predictions on the establishment of a mathematical model of the relationship between the predicted variable and exogenous factors. As an economic variable, sales are typically impacted by various internal and external factors with complex characteristics such as seasonality, autocorrelation, non-linearity, trending, and non-stationarity. These complexities pose significant challenges in constructing an automobile-sales-predictive model with good performance.

The trend-extrapolation method requires the selection of an appropriate time-series model that can summarize the historical predictors of the forecast quantity. Autoregressive integrated moving average (ARIMA) [17], seasonal autoregressive integrated moving average (SARIMA) [18,19], exponential smoothing (ES) [20,21], grey model (GM) [22] and others are forecasting models that are commonly used in trend-extrapolation methods. In particular, SARIMA effectively handles autocorrelation and seasonality in series, and the amount of data required for training is small; this approach is often used to deal with sales series that are greatly affected by seasonal factors [23,24]. However, SARIMA has high requirements with regard to the stationarity of series and cannot handle non-linear factors in the data, which affect the robustness and accuracy of prediction. Aiming to manage the problem that SARIMA model has high requirements with regard to sequence stationarity, Williams et al. [25] proposed a time-series-modeling method based on multi-SARIMA. In that approach, locally estimated scatterplot smoothing is used to perform multi-seasonal-trend decomposition of the series, then SARIMA is used to model the decomposed multi-seasonal components. Zhao et al. [26] proposed a wind-power-prediction method based on EEMD-SARIMA-LSTM; the aggregate-empirical-mode decomposition method is used to decompose the wind-power series into stationary series, then SARIMA and LSTM are used to predict various sub-sequences, and finally, the results are obtained by reconstruction of each predicted value. To address the problem that SARIMA cannot deal with nonlinear factors, Lippi et al. [27] predicted traffic flow based on SARIMA, then introduced the Kalman filter algorithm to modify the predictions. The study points out that Kalman filter could improve the predictions of the SARIMA model. Zhang et al. [28] proposed a short-term prediction method for offshore wind power based on DWT-SARIMA-LSTM. In that method, the power sequence is decomposed into multiple sub-sequences based on discrete wavelet transform and the linear and nonlinear sub-sequences are processed by SARIMA and LSTM, respectively. The above research indicates that the time-series characteristics of the forecasting quantity should be fully considered and that the appropriate model should be selected when the trend-extrapolation method is used.

Machine learning or deep learning models are often used in causal-analysis methods to construct nonlinear regression relationships between dependent variables and predictor variables [29,30]. However, beyond the model's expressive capability, clarity about the factors influencing automobile sales and the construction of reasonable characteristics play a more important role in the predictive ability of the model. Zhang et al. [31] considered the influence of public opinion, as expressed online, and search index on automobile sales to establish a predictive model of automobile sales based on LSTM. Xia et al. [32] considered the influence of macroeconomics, date, price, and other factors on auto sales and constructed a predictive model of auto sales based on XGBoost. Dai et al. [33] extracted 17-dimensional characteristics of influencing factors from three categories (economy, society and policy) and constructed a multivariate predictive model for trends in the development of China's automobile market based on a BP neural network. The above studies indicate that the nonlinear influence of multi-dimensional factors on the forecast quantity should be considered when using causal-analysis methods.

By analyzing the timing features of monthly automobile sales and the influence of external factors, this paper constructs forecasting models from the perspective of trend extrapolation and causal analysis, respectively, then improves the prediction performance by combination forecasting [34,35]. Combination forecasting is a method that integrates the forecasting information of a single model, effectively avoiding the problems of poor robustness and high uncertainty associated with using one single forecasting model [36].

Weighted-average combination combines the predictions by assigning different weights to the predictions of multiple models [37]. Ensemble combination combines multiple models into a meta-learner through supervised learning; this meta-learner then outputs the combination forecasting results. Ensemble-combination strategies can be divided into three types: boosting, stacking, and bagging [38].

In summary, a combination model for automobile-demand forecasting based on trend extrapolation and causality analysis is proposed. The main contributions of this paper are as follows:

- (1) SARIMA and PR are used to mine the historical predictors of sales series, and several methods, such as empirical mode decomposition, a stationarity test, and an autocorrelation test, are introduced to separate the stationary components with strong seasonality from the original series. Then, a trend-extrapolation submodel with high robustness to nonlinearity and non-stationarity is constructed.
- (2) Based on GBDT, the mapping of three types of influencing factors to future sales is established, and causality-analysis submodel with the capacity for nonlinear prediction is constructed.
- (3) The forecasting effects of three kinds of combination strategies, boosting series, stacking parallel and weighted-average parallel, are tested, and the effectiveness of the proposed combination model is verified using actual sales data.

To support better understanding of the study, the flow chart of the research is shown in Figure 1. First, timing features in the historical sales data and correlations among influencing factors were analyzed and a case study was used to improve understanding of the characteristics of automobile-sales data. Then, according to the analysis results, two submodels, trend extrapolation and causality analysis, were developed. Finally, three combination strategies were developed and tested to determine whether they yielded more accurate forecasts.

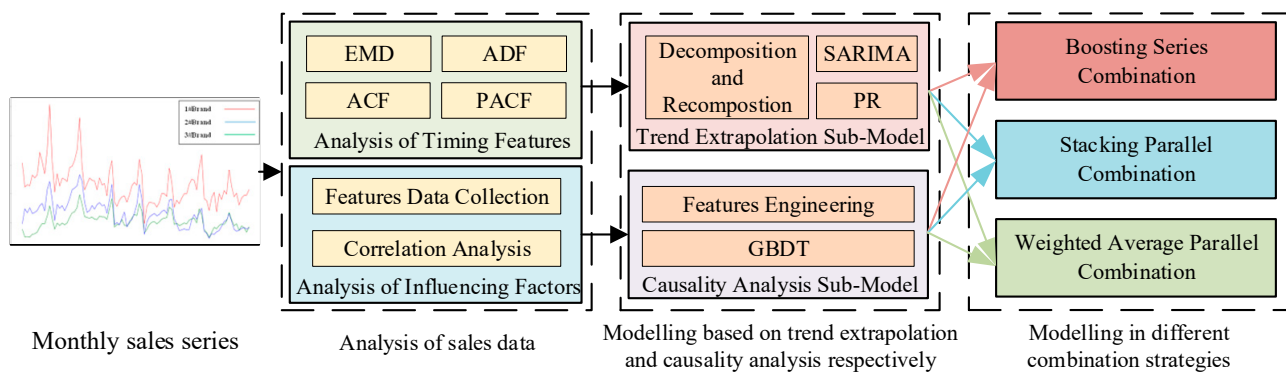


Figure 1. Flowchart of the research in the paper.

The rest of this paper is organized as follows. In Section 2, basic theories of the algorithm involved are introduced, and timing characteristics and factors influencing monthly automobile sales are also analyzed. In Section 3, the methods involving two submodels and three combination strategies are introduced in detail. In Section 4, two groups of ablation experiments and two groups of comparison experiments are conducted to prove the effectiveness and superiority of the proposed model. Finally, the conclusion is given in Section 5.

2. Related Theories

2.1. Seasonal Autoregressive Integrated Moving Average

SARIMA, a method commonly used in seasonal time-series modeling, is a seasonal extension of ARIMA and has the ability to deal with seasonal components in time series.

The SARIMA model adds seasonal AR, MA and difference terms, so a SARIMA $(p, d, q)(P, D, Q)m$ model contains seven parameters; the meaning of each parameter is shown

in Table 3. It is necessary to determine the values of each parameter when using SARIMA to model the sequence.

Table 3. Description table of SARIMA parameters.

Parameters	Meaning of Parameters
p	The non-seasonal part of AR order
d	The non-seasonal part of difference degree
q	The non-seasonal part of MA order
P	The seasonal part of AR order
D	The seasonal part of difference degree
Q	The seasonal part of MA order
m	Length of a season

2.2. Gradient-Boost Decision Tree

GBDT is an iterative decision-tree algorithm that can deal with high-dimensional and nonlinear data in regression tasks and that is often used in prediction problems [39]. In each iteration, the main goal is to reduce the loss function. When the prediction result of the model is inconsistent with the actual observed value, GBDT will generate a decision tree in the direction of the gradient descent of the loss function to reduce the residual of the previous one until the output is basically consistent with the actual observed value. The principle is shown in Figure 2.

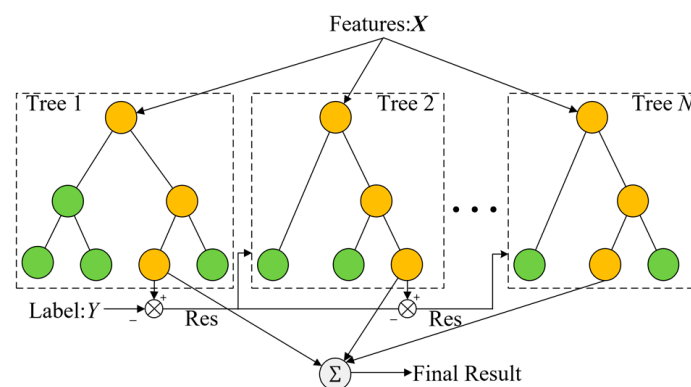


Figure 2. Principle of GBDT.

2.3. Related Time-Series-Analysis Methods

The following two methods for time-series analysis are used in the proposed model.

(1) Empirical mode decomposition

The EMD algorithm can adaptively decompose series with complex fluctuations into multiple Intrinsic Mode Function (IMF) components and one residual component, and each IMF component has a simple oscillation mode [40].

First, EMD constructs the upper and lower envelope of the time series by spline interpolation and calculates the average series. Then, the average sequence is subtracted from the original sequence to obtain the IMF. The above steps are repeated until the remaining sequence becomes monotonous or falls below the set threshold. Finally, all IMF and residual terms are obtained.

(2) Stationarity test

Stationarity is the basic assumption of many time-series models. If a time-series model were applied to non-stationary series without transformation, it would lead to inaccurate predictions. The augmented Dickey–Fuller (ADF) test is a strict statistical test method that can be used to judge the stationarity of the series. The Statsmodels module in Python

provides an ADF test function, which returns five sets of values. Table 4 shows the meaning of the returned values of this function. When there is 95% confidence that a sequence is stationary, the ADF test results need to meet two conditions: $T < T_5$ and $p \rightarrow 0$.

Table 4. Return-value table of ADF in the Statsmodels module.

Parameters	Meaning of Parameters
T	T statistics
P	Probability value of the T statistic
Lags Used	Lags
Number of Observations Used	Numbers of inputs to the test
[1%: T_1 , 5%: T_5 , 10%: T_{10}]	The significance level of T 's value

(3) Autocorrelation test

Autocorrelation is another important feature that is often of concern in time-series modeling. It represents the degree of correlation between variables at two different time points in the same time series. The Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) are often used to analyze the autocorrelation of time series.

3. Proposed Methodology

3.1. Data Description

Three time series of monthly automobiles sales were collected from the sales department of Geely Holding Group, which is one of the largest private automobile manufacturers in China. Due to the confidentiality requirements of the company, the data underwent desensitization processing via range scaling and numerical offsets that did not change the distribution and sequence characteristics of the data. The brands involved are all sub-brands under Geely, and the sales periods are all from January 2016 to August 2023. Detailed information on the sales period, power types, car types, price range, and architectures of the brands for which data on automobile sales were collected is shown in Table 5. The term “architecture” refers to the basic manufacturing platforms of Geely, among which are compact modular architecture (CMA), B-segment modular architecture (BMA), and sustainable experience architecture (SEA). Sales data for automobiles with SEA are scarce because this model was developed only in September 2020, so we collected sufficient data only for CMA and BMA.

Table 5. Detailed information about the data-collection objects.

Brand	Power Types	Car Types	Price Range (¥k)	Architecture	Sales Period
number one	gasoline, electric, hybrid	SUV, Sedan	69.9–88.9	CMA	2016.01–2023.08
number two	gasoline	SUV, Sedan	68.9–105.9	BMA	2016.01–2023.08
number three	gasoline	SUV	94.9–114.9	BMA	2016.01–2023.08

3.2. Analysis of Automobile Sales Data

As shown in Figure 3, all sales-volume data exhibit a certain degree of autocorrelation, trending, and non-stationarity, and none exhibits an obvious linear relationship, showing instead complex characteristics of multi-mode aliasing. Due to the impacts of the COVID-19 pandemic and the rise of new-energy vehicles on the gasoline-automobile market, sales of all brands showed a downward spike in the 50th month, which corresponds to February 2020, when the COVID-19 pandemic started. Sales of the number-two and number-three brands, which had not yet launched their own new-energy products, also show an overall downward trend in the following months. In addition, every time series shows a noticeable increase in the winters, which corresponds to the Chinese habit of purchasing automobiles before the spring festival; these trends show autocorrelation and seasonality.

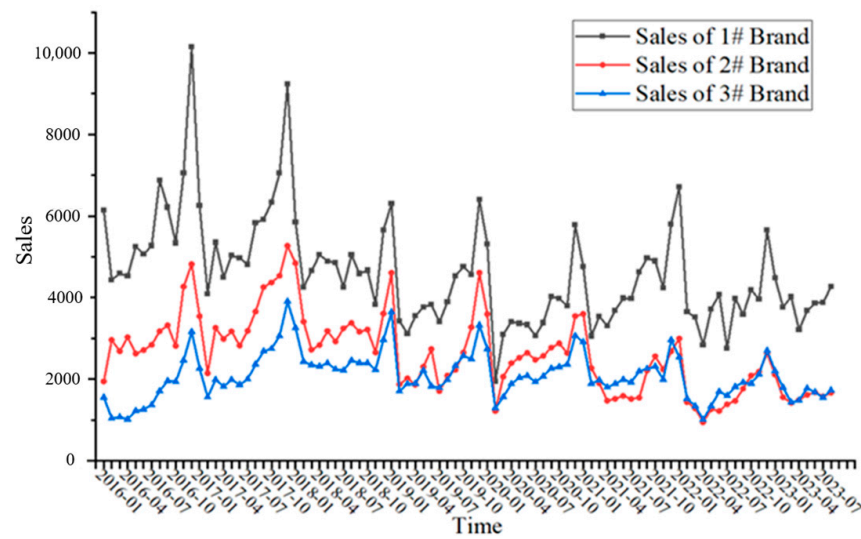


Figure 3. Automobile sales of three brands in mainland China.

3.2.1. Analysis of Timing Characteristics

To clarify the timing characteristics of monthly automobile sales, EMD was used to decompose sales sequences into several IMFs, and the stationarity and autocorrelation test methods were used to analyze these IMFs. Monthly sales of the number-one brand are introduced as an example below.

Monthly sales sequence $S(t)$ is decomposed into 4 IMFs, from $IMF_1(t)$ to $IMF_4(t)$, and one residual component, $R(t)$, by the EMD method. IMFs are arranged according to the order of fluctuation frequency, from high to low. $R(t)$ presents low volatility or monotonicity, and the decomposition results are shown in Figure 4. The results reflect the trend of change in monthly sales volume, and each IMF reflects the different nature of the associated monthly sales-volume series under the influence of various factors. ACF and PACF test results and ADF test results for the original sequence and each IMF sequence are shown in Figure 5 and Table 6, respectively.

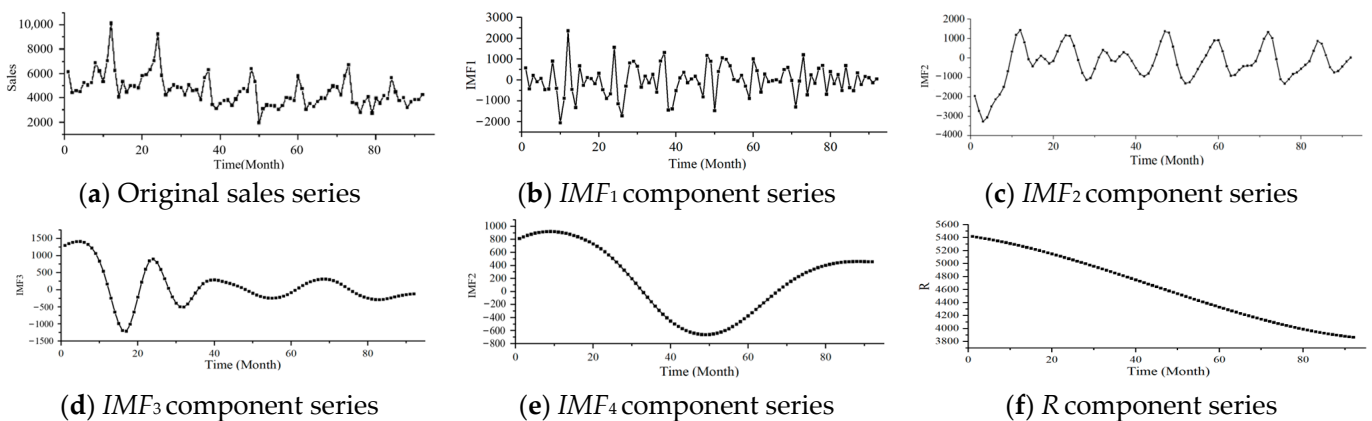


Figure 4. EMD-decomposition result of monthly automobile sales of the number-one brand.

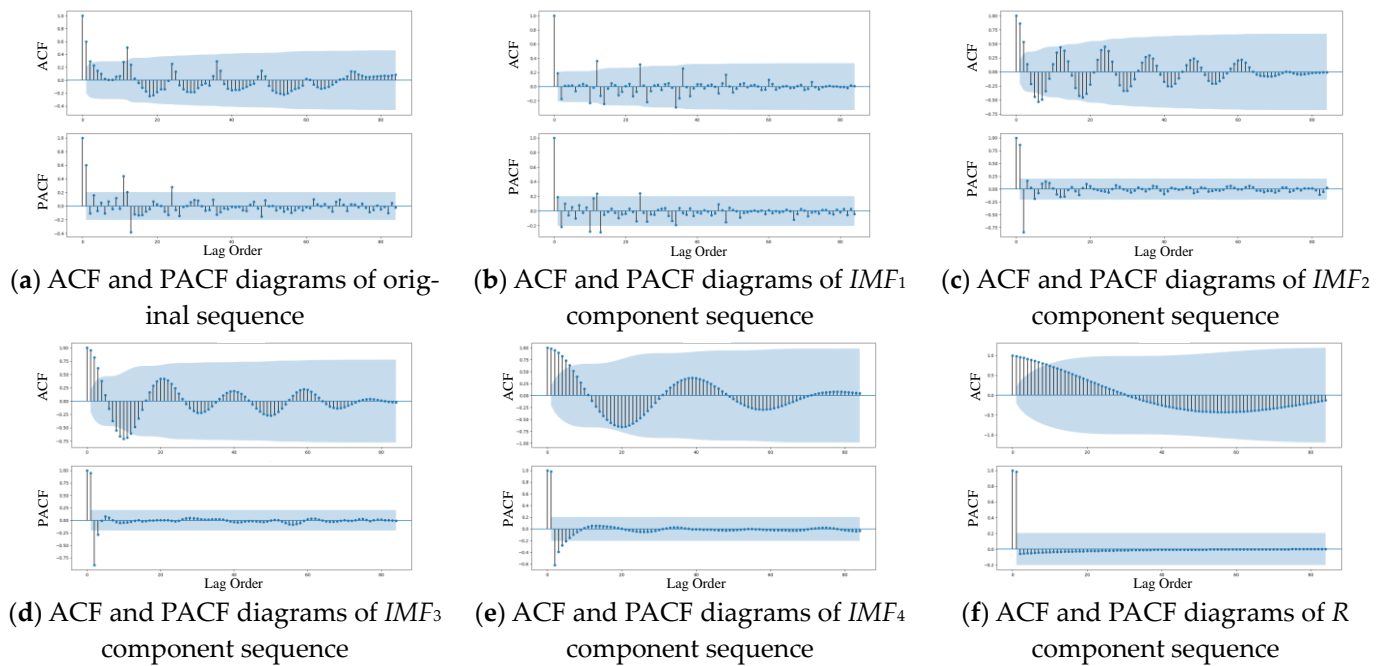


Figure 5. Diagrams of ACF and PACF of the original sequence, of each IMF, and of each residual sequence.

Table 6. ADF test results of original monthly sales sequence and each IMF component.

Series	T Value	p Value	Significance Level 1%	Significance Level 5%	Significance Level 10%
$S(t)$	-1.748	0.407	-3.516	-2.899	-2.587
$IMF_1(t)$	-2.902	0.045	-3.516	-2.899	-2.587
$IMF_2(t)$	-5.957	2.08×10^{-7}	-3.509	-2.896	-2.585
$IMF_3(t)$	-3.653	0.0048	-3.514	-2.895	-2.585
$IMF_4(t)$	-2.110	0.241	-3.509	-2.896	-2.585
$IMF_{12}(t)$	-4.232	5.82×10^{-4}	-3.516	-2.899	-2.587

According to the ADF result, the significance level corresponding to the T -value of is greater than 10% and the p -value does not approach 0, showing non-stationarity. According to the analysis of Figure 5a, ACF of $S(t)$ shows a certain degree of seasonality, and its PACF tends to converge with the increase of the lag order. Therefore, $S(t)$ is a non-stationary sequence with a certain degree of seasonality, making it difficult to obtain accurate predictions by using SARIMA modeling directly.

According to the ADF result, the significance level corresponding to the T of $IMF_1(t)$ is greater than 5% but less than 1%, which meets the condition of sequence stationarity to some degree. The significance level corresponding to the T of $IMF_2(t)$ is greater than 1% and the p -value approaches 0, so this test result meets the determination condition of sequence stationarity. The ACF and PACF of $IMF_1(t)$ and $IMF_2(t)$ are shown in Figure 5b,c respectively. The ACF of $IMF_1(t)$ is significant at the 12th-order and 24th-order lags, and both ACF and PACF fall into the error range quickly and are truncated at the 1st-order lags, indicating that this is an almost stationary sequence with significant seasonality. The ACF maximum of $IMF_2(t)$ is interrupted by 12 lag orders, showing significant seasonality, and its PACF rapidly falls within the error range at the third-order lag position, showing truncation, indicating that it is a stationary sequence with significant seasonality. Since components both show good stationary and have the same seasonality, they are reconstructed as the seasonal components, $IMF_{12}(t)$. The significance level corresponding

to the T of $IMF_{12}(t)$ is greater than 1% and the p -value approaches 0, indicating that it is a stationary sequence, which makes it suitable for analysis by SARIMA.

The ACF and PACF of $IMF_3(t)$ and $IMF_4(t)$ are shown in Figure 5d,e, respectively. Both the ACF and the PACF oscillate, and the period of oscillation changes as lag order grows, reflecting the nonlinear mode in the original sequence. All the standard seasonal models assume that there is a pattern in the fixed period in the sequence, that this non-fixed periodicity cannot be described by the standard seasonal model, and that nonlinear models are required to handle this situation.

3.2.2. Analysis of Influencing Factors

In addition to timing characteristics, it is equally important to understand external factors that can influence automobile sales. This paper extracts 31-dimensional feature data from the following three types of influencing factors and analyzes their correlation.

- (1) Date factors, including total days in the month, number of holidays, number of working days, number of working days within 30 days after a holiday, number of weekends within 30 days after a holiday, month of the year, whether January was the previous month, whether January is the next month, number of holidays after the Chinese New Year, number of working days after the Chinese New Year, number of days after the Chinese New Year, whether it is the end of the quarter, whether it is the end of the half-year, and the quarter in which it belongs. The above feature data can be obtained directly from the calendar.
- (2) Macroeconomic factors, including purchasing-managers' index, produce index, new-orders index, new-export-orders index, orders-in-hand index, finished-produce index, purchase-quantity index, import index, purchase index, employment index, production-expected index, business index, inventory index, input index, sales index, and expected-business-activity index. The above features are macroeconomic monitoring indicators issued by the China National Bureau of Statistics that can comprehensively reflect the monthly situation with regard to the various links among enterprise procurement, production, circulation, etc., and are internationally accepted macroeconomic monitoring indicators.
- (3) Consumption-promotion factors, including purchase-tax discount for the month.

The heatmap of the correlation matrix is shown in Figure 6.

First, the degree of correlation between the response variable (Sales) and each explanatory variable (Features) is presented in the heatmap. The correlation coefficients of most macroeconomic features are in the range [0.23, 0.4]. The correlation coefficient of purchase-tax discount in the current month is 0.43. The correlation coefficients of date factors such as month, quarter, Spring Festival, and working days are in the range [0.28, 0.59]. There is a certain degree of correlation between the extracted features and sales volume, which means that these features are useful for establishing regression models.

In addition, the similar degree of correlation among some explanatory variables shows multicollinearity, which is harmful when constructing a regression model. According to the heatmap, the correlation coefficients of most macroeconomic factors and some date factors are greater than 0.6, implying the existence of multicollinearity. For example, the correlation coefficient between factors of "days_count_before_springfestival" and "holidays_count_before_springfestival" is 0.97, which means the factors are almost linearly related and provide the same information to the regression model. The inclusion of both may reduce the stability and accuracy of the model, so further feature engineering before modeling is required.

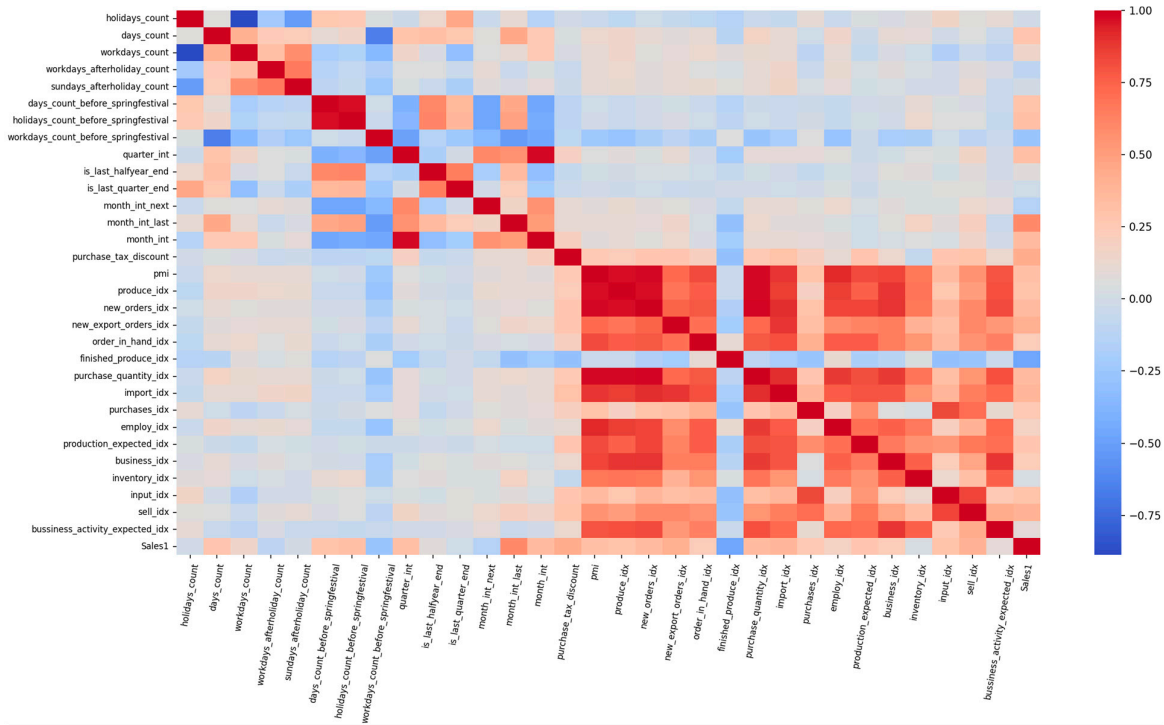


Figure 6. Heatmap of correlation matrix of features and sales data.

3.3. Trend-Extrapolation Submodel

According to the analysis of the monthly sales sequence, a trend-extrapolation model based on EMD-SARMI-PR is proposed. The proposed trend-extrapolation model consists of the following five steps, and its flow chart is shown in Figure 7.

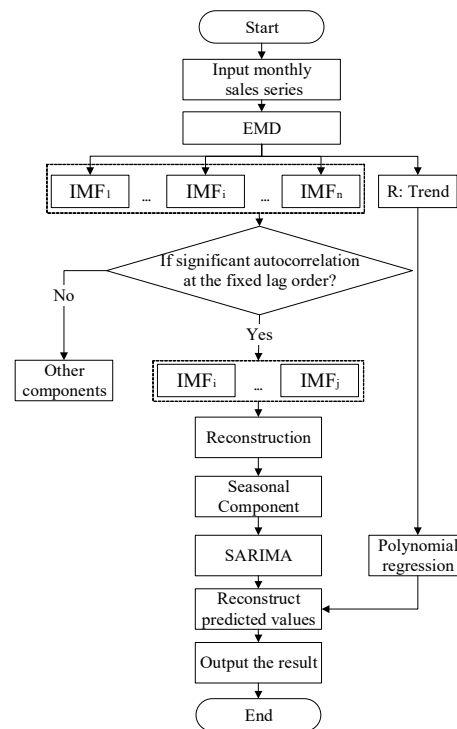


Figure 7. Flow chart of the trend-extrapolation submodel.

Step 1: EMD is used to decompose the monthly sales sequence into multiple IMFs and one residual component, the relationship between which can be expressed in Equation (1) [40], as follows:

$$S(t) = \sum_{i=1}^n IMF_i(t) + R(t) \quad (1)$$

where $S(t)$ is the monthly automobile-sales sequence; $IMF_i(t)$ is the i th IMF sequence; $R(t)$ is the residual sequence; and n is the total number of IMFs.

Step 2: Using ACF and PACF test results for each IMF, the IMF sequences are recomposed with both stationary and the same seasonality as a seasonality component.

Step 3: The seasonality component is forecast by SARIMA.

Step 4: Considering $R(t)$ as the tendency component, it is forecast by PR.

Step 5: The predicted values of components in Step 3 and Step 4 by sampling points are summed to obtain the predicted value of the trend-extrapolation submodel \hat{S}_{TE} .

In addition, the other IMFs do not have obvious seasonality or trending but are characterized by non-stationarity, nonlinearity, and minor amplitude. Such components are not suitable for modeling by linear methods such as SARIMA and PR, so they were not considered in the trend-extrapolator models.

3.4. Causality-Analysis Submodel

To compensate for the problem that the trend-extrapolator model has difficulty dealing with the nonlinear mode in $S(t)$, a causality-analysis submodel based on GBDT was constructed.

Feature fusion was necessary before the model could be trained because of the multicollinearity of features mentioned previously. Principal component analysis (PCA) [41] was used to reduce the dimensions of the features to obtain the feature set $\mathbf{x}(t)$. The result of PCA analysis is shown in Figure 8. The sum of the ratios of the first three principal components is greater than 0.9, and the first three principal components were used to train the GBDT model.

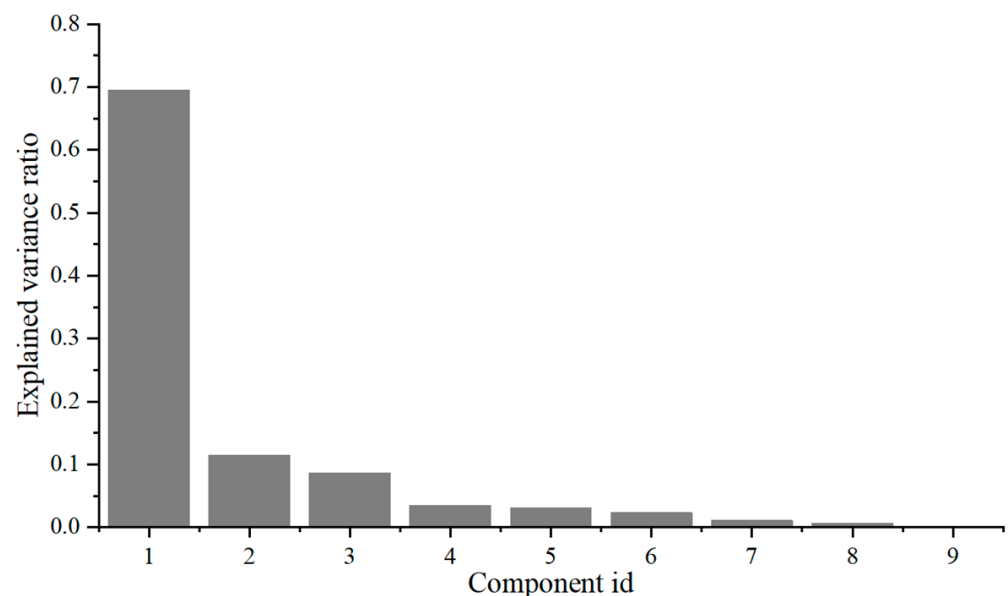


Figure 8. Explained variance ratio of each principal element of PCA.

After the feature fusion was complete, the relationship between $\mathbf{x}(t)$ and $S(t)$ was obtained based on GBDT, and the root mean square error between model output \hat{S}_{CA} and actual sales volume S was reduced as a loss function, the causality-analysis submodel $f_{GBDT}(\mathbf{x}(t))$ was finally obtained. The LightGBM framework was used for GBDT training.

The main GBDT parameters are as follows:

- (1) num_boost_round decides the number of the tree, the larger the number, the weaker the generalization ability; the smaller the number, the weaker the learning ability.
- (2) learning_rate decides the contribution of each tree to the result; a lower learning rate usually leads to better training results.
- (3) max_depth, num_leaves, min_data_in_leaf, and min_sum_hessian_in_leaf determine the complexity of each decision tree, appropriate values of these values can prevent overfitting.
- (4) feature_fraction and bagging_fraction are used in the selection of training samples and training features, introducing randomness at each iteration to prevent overfitting.

GBDT model parameters were selected by grid search, and the range of parameters and the values selected are shown in Table 7.

Table 7. Parameter details of GBDT.

Parameters	Range of Parameters	Selected Value
learning_rate	0.01, 0.05, 0.1	0.01
num_boost_round	600, 1000, 1400, 1800	1400
max_depth	−1	−1
num_leaves	20, 30, 40, 50	30
min_data_in_leaf	10, 12, 14, 16, 18, 20	14
min_sum_hessian_in_leaf	0.001, 0.1, 1, 10	0.001
feature_fraction	0.6, 0.8, 1.0, 1.2, 1.4	1.0
bagging_fraction	0.5, 0.6, 0.7, 0.8, 0.9, 1.0	0.7

3.5. Combination Strategy of Two Submodels

The combination strategies of the submodels will affect the final predictions. Models using different combination strategies were designed, and a comparative experiment, conducted to select the best combination strategy, is described in Section 4. Combination strategies can be sorted into two types: the weighted-average strategy and the ensemble-learning strategy.

The weighted-average strategy is a linear combination method to allocate different weights to different sub-modules by their historical performance. The inverse variance weighting method is always used to determine the weights.

The ensemble-learning strategy is another commonly used method which can capture the nonlinearity of the outputs of submodels [42]. There are three usual methods: boosting, stacking, and bagging [43]. Boosting uses a series structure, where each model learns from the previous model and fixes the errors of the previous model. Stacking uses a parallel structure to take the output of each submodel as the input of the meta-learner, and the meta-learner determines the proportion of the predictions of each submodel. Bagging generates multiple datasets by bootstrapping samples to train different copies of the same model and then obtains the result by averaging or voting. The exact combination method should be determined by the specific problem and data; bagging is not suitable for combining two different models.

Therefore, the combination strategies tested in the experiment include the boosting series strategy, the weighted-average parallel strategy, and the stacking parallel strategy. The methods used to implement these three combination strategies are shown in Figure 9a–c, respectively.

- (1) Boosting series combination

First, the predictions $\hat{S}_{TE}(t)$ of the trend fader model were obtained. Furthermore, the fitting results of the trend-extrapolation submodel and the data on other influencing factors were used as the features of the causal analysis submodel for training, as shown in Equation (2), as follows:

$$S(t) = f_{GBDT}(\mathbf{x}(t), \hat{S}_{TE}(t)) \quad (2)$$

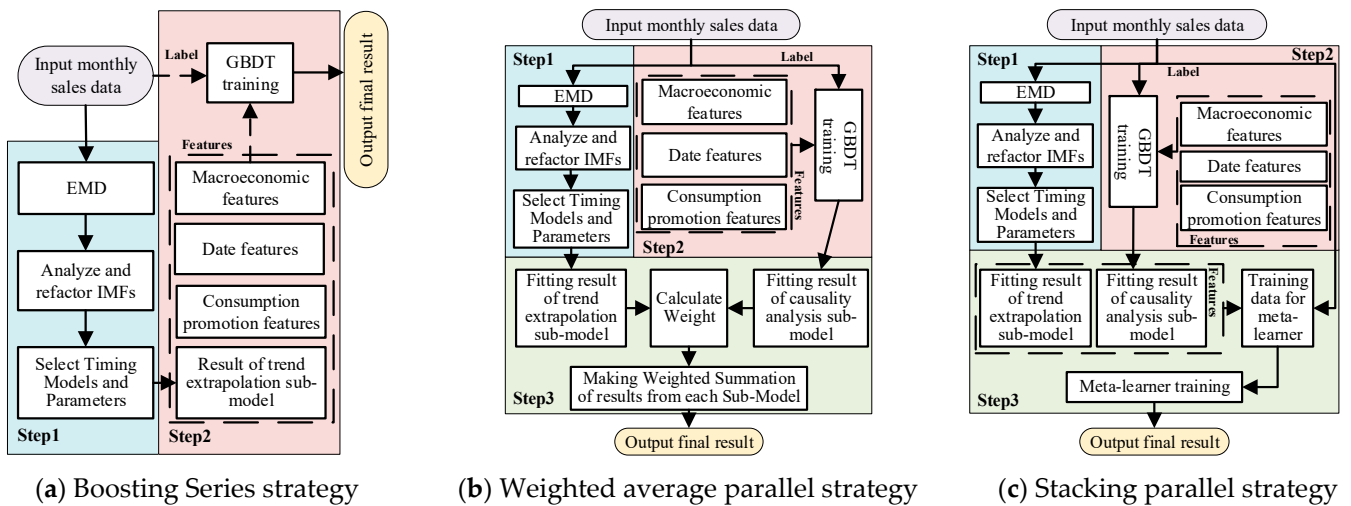


Figure 9. Forecasting models using different combination strategies.

Finally, the forecasts $\hat{S}_{TE}(t)$ of the trend-extrapolation submodel and the features data $\mathbf{x}(t)$ of the month were input into the trained causality-analysis submodel to obtain the final predictions.

(2) Weighted-average parallel-combination model

The first step was to obtain the forecast of the trend-extrapolation submodel $\hat{S}_{TE}(t)$ and the causality analysis result $\hat{S}_{CA}(t)$. Next, the variance of the forecast error and the weight of submodels for each month in the past year were calculated using Equations (3) and (4). Finally, the forecasting result was output via Equation (5) [44], as follows:

$$\lambda_i = \frac{\sum_{j=1}^n (e_j^i - \bar{e}^i)}{n} \tag{3}$$

$$\eta_i = \frac{1}{\lambda_i \sum_{i=1}^2 \frac{1}{\lambda_i}} \tag{4}$$

$$S = \eta_1 S_{TE} + \eta_2 S_{CA} \tag{5}$$

where S , S_{TE} , and S_{CA} are the forecasting results of the combination model, the trend-extrapolation submodel, and the causality-analysis submodel, respectively; η_i is the weight for each submodel; λ_i is the variance of the forecast error; e_j^i is the forecast result of the i th submodel for the j th month; and \bar{e}^i is the mean error of the i th submodel.

(3) Stacking parallel combination

The first step was to obtain the forecast from the trend-extrapolation submodel $\hat{S}_{TE}(t)$ and the causality-analysis result $\hat{S}_{CA}(t)$. Next, the meta-learner was trained based on GBDT, and the mapping between the forecast results of the two submodels and the actual sales data was constructed as shown in Equation (6), as follows:

$$S(t) = f_{GBDT}(\hat{S}_{TE}(t), \hat{S}_{CA}(t)) \tag{6}$$

Finally, the forecast was output from the meta-learner.

4. Experiment and Discussion

In this section, proposed submodels and combination strategies are evaluated and analyzed. The dataset used in this paper is described, and the experimental design is provided in detail. Additionally, several ablation experiments and comparison experiments were conducted, and their results are discussed.

4.1. Experiment Setup

Due to the large fluctuations in vehicle sales between different quarters, the effectiveness of the model cannot be objectively evaluated by relying only on the monthly sales predictions of a specific quarter, so the training set and test set of each submodel were divided as shown in Figure 10. Monthly automobile sales for the four quarters from September 2022 to August 2023 were forecasted on a rolling basis.

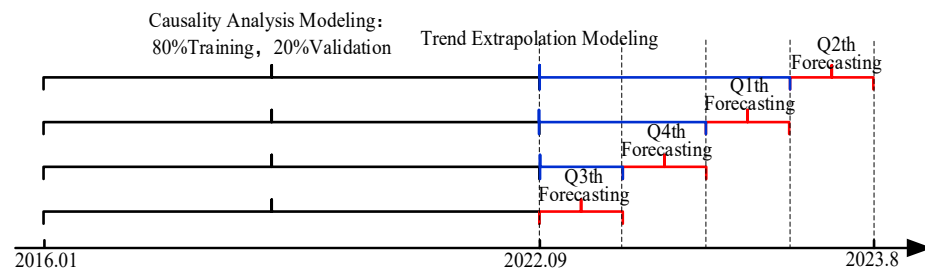


Figure 10. Dataset-partitioning diagram.

The sales data from January 2016 to August 2022 (black part of Figure 10) and the sales data from the past quarter (blue part of Figure 10) were used to train the trend-extrapolation submodel and forecast the monthly sales of the next quarter (red part of Figure 10). The sales data from January 2016 to August 2022 (black part of Figure 10) were divided into 80% training set and 20% validation set to train the causality-analysis submodel and predict the sales of each month. This kind of dataset partition is more in line with the actual needs of enterprises because it allows the forecast model to learn the recent trend, thereby reducing error accumulation.

Mean Absolute Percent Error (MAPE) of monthly sales forecast results for each quarter was used as the evaluation index. The smaller the MAPE value of the model, the higher the model accuracy. The formula used to calculate MAPE is shown in Equation (7), as follows:

$$\text{MAPE} = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{\hat{S}_i - S_i}{S_i} \right| \quad (7)$$

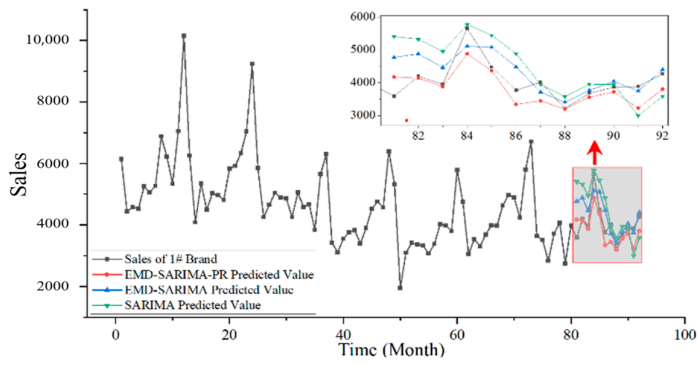
where \hat{S}_i is the monthly sales forecast for the i th month of each quarter and S_i is the actual sales value of the i th month.

The following experiments were programmed using Python on a computer with 8 cores, 16 GB of RAM, 6 GB of VRAM, and 512 GB of hard-disk space.

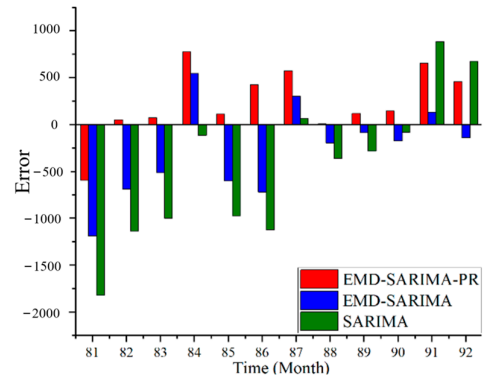
4.2. Ablation Experiment of the Trend-Extrapolation Submodel

To validate the effectiveness of the proposed trend-extrapolation model, the forecasting performance of EMD-SARIMA-PR, EMD-SARIMA, and SARIMA were compared. In EMD-SARIMA, SARIMA is used to forecast each IMF without distinction, and the predicted values of each IMF component are summed to obtain the result. In SARIMA, SARIMA is directly used to forecast the original sales sequence. The EMD algorithm and SARIMA algorithm were implemented based on the PyEMD library and the pmdarima library, respectively.

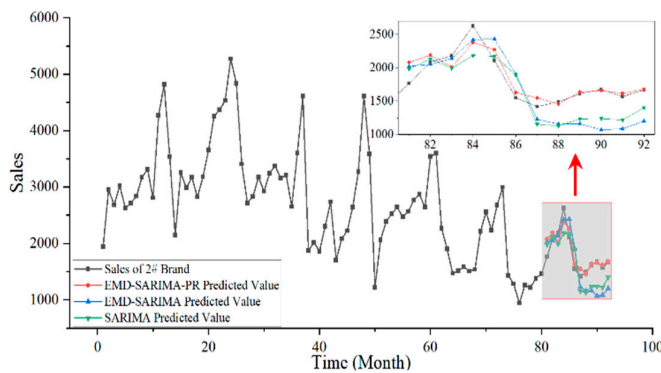
Ablation experiments were carried out on the three groups of sales data. The comparison images and prediction residuals of the predicted sales values and actual sales values of the three models are shown in Figure 11, and the MAPE values of the predicted results are shown in Table 8.



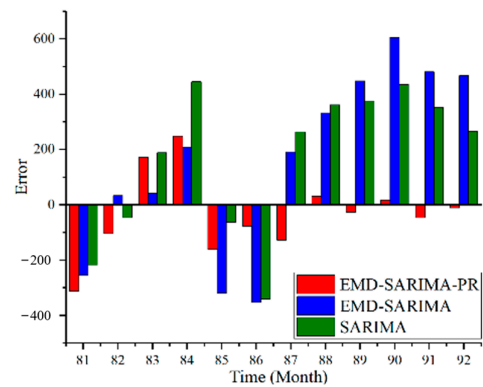
(a) Performance comparison of different models in sales forecasting of the number-one brand



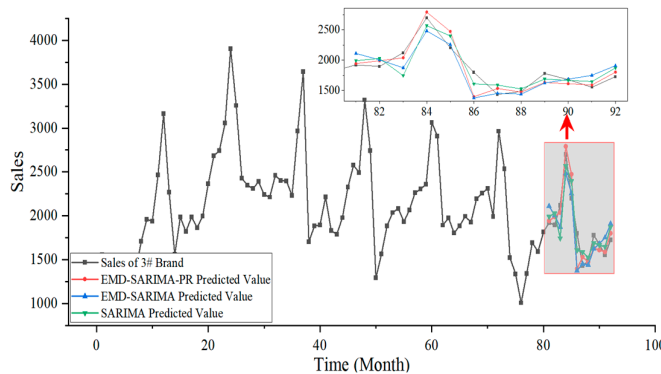
(b) Comparison of errors from different models in sales forecasting of the number-one brand



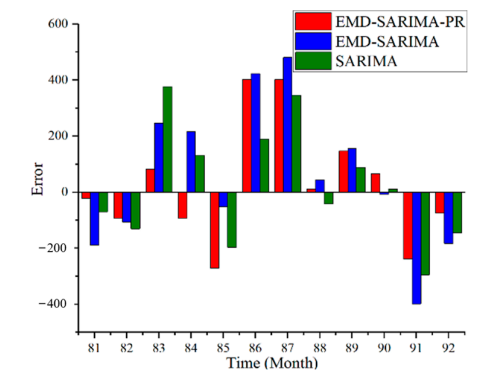
(c) Performance comparison of different models in sales forecasting of the number-two brand



(d) Comparison of errors from different models in sales forecasting of the number-two brand



(e) Performance comparison of different models in sales forecasting of the number-three brand



(f) Comparison of errors from different models in sales forecasting of the number-three brand

Figure 11. Performance and Comparison of errors from different trend-extrapolation models used for sales forecasting.

Table 8. Performance comparison of different models via MAPE.

Model	MAPE		
	Number-One Brand	Number-Two Brand	Number-Three Brand
SARIMA	0.1814	0.1629	0.0911
EMD-SARIMA	0.1098	0.1851	0.1151
EMD-SARIMA-PR	0.0804	0.0581	0.0864

According to Figure 11a,c,e, the sales sequence predicted by EMD-SARIMA-PR is most similar to the actual sales sequence. According to Figure 10b,d,f, the residual deviation of the result from EMD-SARIMA-PR has a low degree of deviation from the 0 line and a low range of variation, lower than those of the other two models in addition.

According to Table 8, the forecast results of EMD-SARMI-PR have the smallest MAPE value, which means they have the highest forecast accuracy. Compared with EMD-SARIMA and SARIMA, the MAPE of EMD-SARMI-PR decreased by 26.78% and 55.65%, respectively, for the number-one-brand data; the MAPE decreased by 68.61% and 64.33%, respectively, for the number-two-brand data; and the MAPE decreased by 24.93% and 5.16%, respectively, for the number-three-brand data.

In summary, the trend-extrapolation model based on EMD-SARIMA-PR shows the best forecasting performance, and this model can effectively capture the linear law in the monthly-car-sales series.

4.3. Comparison Experiment Among Three Combination Strategies

To select the best combination strategy, the forecasting performance of the models in Section 3 using the three combination strategies were compared. The comparison images of the forecast sales values and the residuals of the forecast models using different combination strategies are shown in Figure 12, and the MAPE values of the predictions are shown in Table 9.

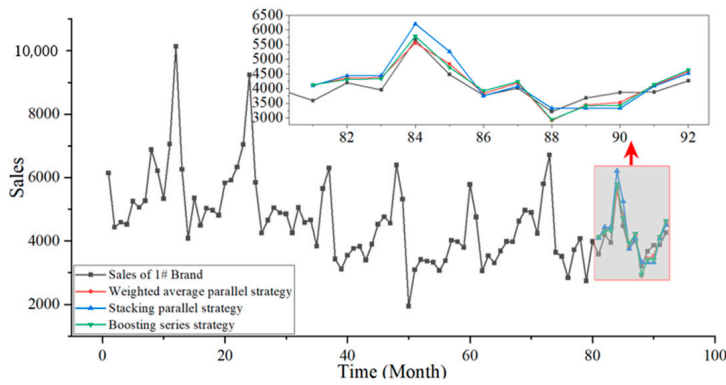
According to Figure 12a,c,e, the change tendency of the forecast sales sequence of the weighted-average parallel-combination model is most similar to the actual sales sequence. According to Figure 12b,d,f, the weighted-average parallel-combination model has a lower variance and deviation compared to the other models.

According to Table 9, the forecast results of the weighted-average parallel-combination model have the lowest MAPE value, which means this model has the highest prediction accuracy. For the number-one brand data, the forecast MAPE values of the weighted-average parallel-combination model were reduced by 16.81% and 4.68%, respectively, compared with the stacking parallel-combination model and the boosting series-combination model. For the number-two-brand data, the forecast MAPE values of the weighted-average parallel-combination model were reduced by 25.60% and 2.79%, respectively, compared with the Stacking parallel combination model and the Boosting series-combination model. For the number-three-brand data, the forecast MAPE values of the weighted-average parallel-combination model were reduced by 46.26% and 14.37%, respectively, compared with the stacking parallel-combination model and the boosting series-combination model.

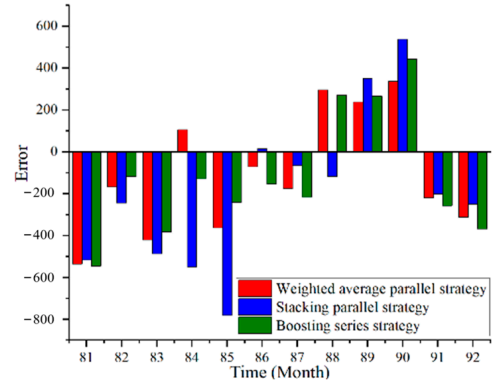
In summary, the weighted-average parallel-combination model shows the best forecasting performance.

Table 9. Performance comparison of different combination strategies in MAPE.

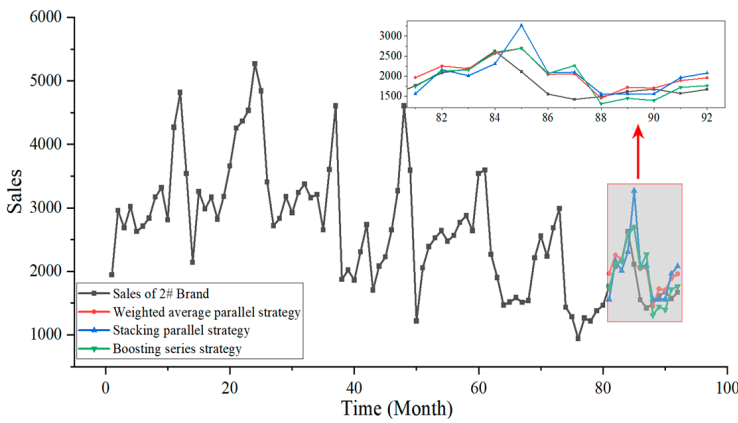
Combination Strategy	MAPE		
	Number-One Brand	Number-Two Brand	Number-Three Brand
Boosting series strategy	0.0727	0.1507	0.0696
Stacking parallel strategy	0.0833	0.1969	0.1109
Weighted average parallel strategy	0.0693	0.1465	0.0596



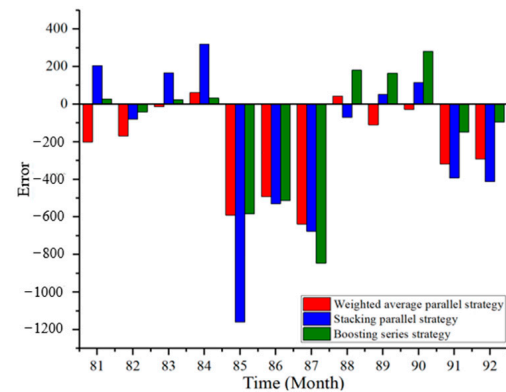
(a) Performance comparison of different combination strategies in sales forecasting of the number-one brand



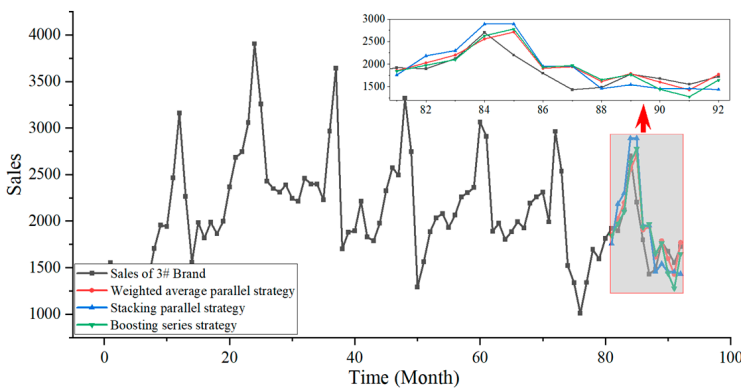
(b) Comparison of errors from different combination strategies in sales forecasting of the number-one brand



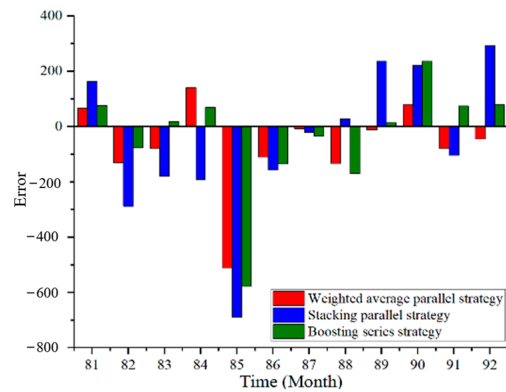
(c) Performance comparison of different combination strategies in sales forecasting of the number-two brand



(d) Comparison of errors from different combination strategies in sales forecasting of the number-two brand



(e) Performance comparison of different combination strategies in sales forecasting of the number-three brand



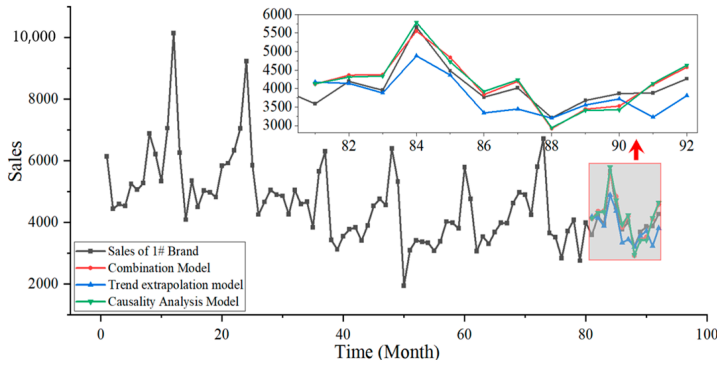
(f) Comparison of errors from different combination strategies in sales forecasting of the number-three brand

Figure 12. Performance and comparison of errors from different combination models used for sales forecasting.

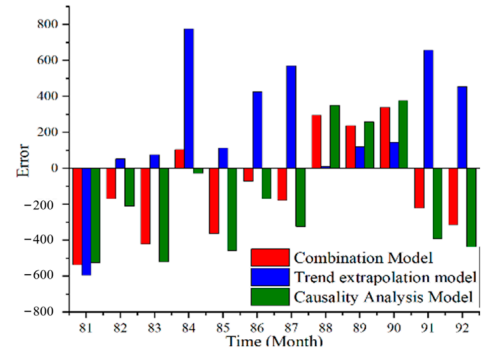
4.4. Ablation Experiment Conducted on the Proposed Combination Model

Through the comparison experiments in Section 4.3, the weighted-average parallel-combination strategy was selected to construct the forecast model. In this section, ablation experiments carried out on the proposed combination model are described and the rela-

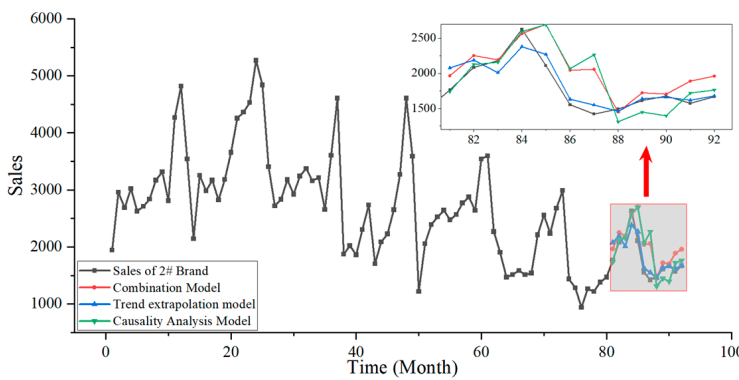
tionship between the composite model and its submodels is discussed. Images comparing the sales-forecast values and prediction residuals from the trend-extrapolation submodel, the causality-analysis submodel, and the weighted-average parallel-combination model are shown in Figure 13, and the MAPEs of the forecast results are shown in Table 10.



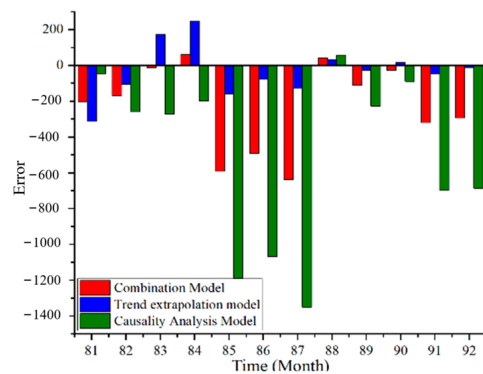
(a) Comparison of results from the proposed model and submodels in sales forecasting of the number-one brand



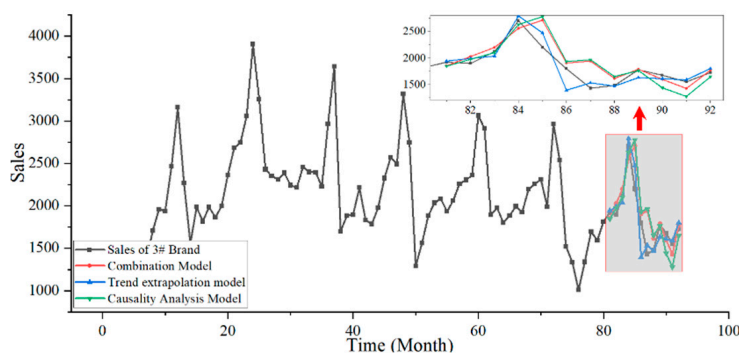
(b) Comparison of errors from the proposed model and submodels in sales forecasting of the number-one brand



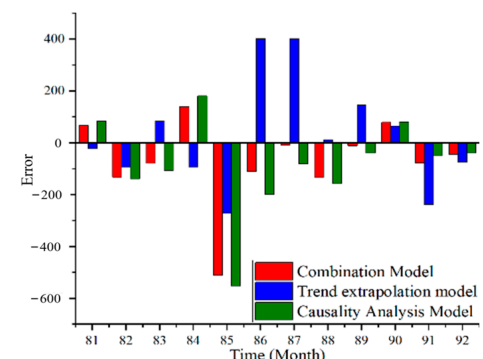
(c) Comparison of results from the proposed model and submodels in sales forecasting of the number-two brand



(d) Comparison of errors from the proposed model and submodels in sales forecasting of the number-two brand



(e) Comparison of results from the proposed model and submodels in sales forecasting of the number-three brand



(f) Comparison of errors from the proposed model and submodels in sales forecasting of the number-three brand

Figure 13. Comparisons of results and errors from the proposed model and submodels for sales forecasting.

Table 10. Performance comparison of the proposed combination model and its submodels according to MAPE.

Model	MAPE		
	Number-One Brand	Number-Two Brand	Number-Three Brand
Trend-extrapolation model	0.0804	0.0581	0.0864
Causality-analysis model	0.0871	0.3029	0.0726
Weighted-average parallel-combination model	0.0693	0.1465	0.0596

For the number-one-brand data, in the 84th, 86th, 87th, 91st and 92nd months, the combination model effectively reduced the forecast error of the trend-extrapolation model; in the 82nd, 83rd, 88th, 89th and 90th months, although the forecast errors of the combination model are higher than that of the trend-extrapolation model, the forecast sales sequence still shows a similar change tendency with the actual sales sequence. From the overall forecast results, the MAPE values of the combination model were 13.81% and 20.44% lower than those of the trend-extrapolation model and the causality-analysis model.

For number-two-brand data, the forecasting performance of the combination model was worse than that of the trend-extrapolation model. Compared with those of the number-one brand and the number-three brand, the sales sequence of the number-two brand has a more obvious downward trend that is not captured by the causality-analysis model, which leads to an obvious deviation in the prediction of the causal-analysis model. The trend-extrapolation model can effectively capture the trend in the historical data, so its prediction accuracy is much higher than that of causality-analysis model. The combination model synthesizes the predictions of the two models and corrects the bias of the causality-analysis model from the 83rd month to the 92nd month to some extent.

For the number-three-brand data, in the 86th, 87th, 91st and 92nd months, the combined model effectively reduced the prediction error of the trend-extrapolation model; in the 85th month, the causality-analysis model and the combination model incorrectly predicted the peak sales volume, resulting in an error. Except in the 85th month, the change trend of the forecast sales sequence of the combination model is more similar to that of the actual sales sequence and the error distribution of the combination model is the most uniform. From the overall forecast results, the MAPE of the combination model is 31.02% and 17.91% lower than those of the trend-extrapolation model and the causal-analysis model, respectively.

In summary, the trend-extrapolation model can effectively capture and extrapolate the law of historical sales data, but there is a problem of large forecast variance. The forecast variance of the causality-analysis model is smaller than that of the trend-extrapolation model, but it cannot capture the sales-series trend, which may be the cause of the deviation between the forecast and the data. The proposed combination model can effectively integrate the forecast results of the two submodels and improve the forecast accuracy.

4.5. Comparison Experiment between Proposed Combination Model and Other Common Models

The proposed combination model was compared with common time-series models and a deep learning model. The models involved in the comparison are as follows:

- (1) ES: Forecast automobile sales based on a triple-exponential smoothing method.
- (2) GM: Forecast automobile sales based on the first-order unary grey model GM (1,1).
- (3) XGBoost: Forecast automobile sales based on the XGBoost model.
- (4) SVR: Forecast automobile sales based on the SVR model with a linear kernel.
- (5) CNN: Forecast automobile sales based on a convolutional neural network that includes two Conv1D layers, two Maxpooling1D layers, and one FC layer. The number and the size of convolution kernels in each Conv1D layer are 32 and 3, the activation function is Relu, and the loss function is the mean square error.

- (6) LSTM: Forecast automobile sales based on a long short-term-memory neural network, which includes two LSTM layers and one FC layer. The number of hidden neurons in each LSTM layer is 50, the activation function is Relu, and the loss function is the mean square error.

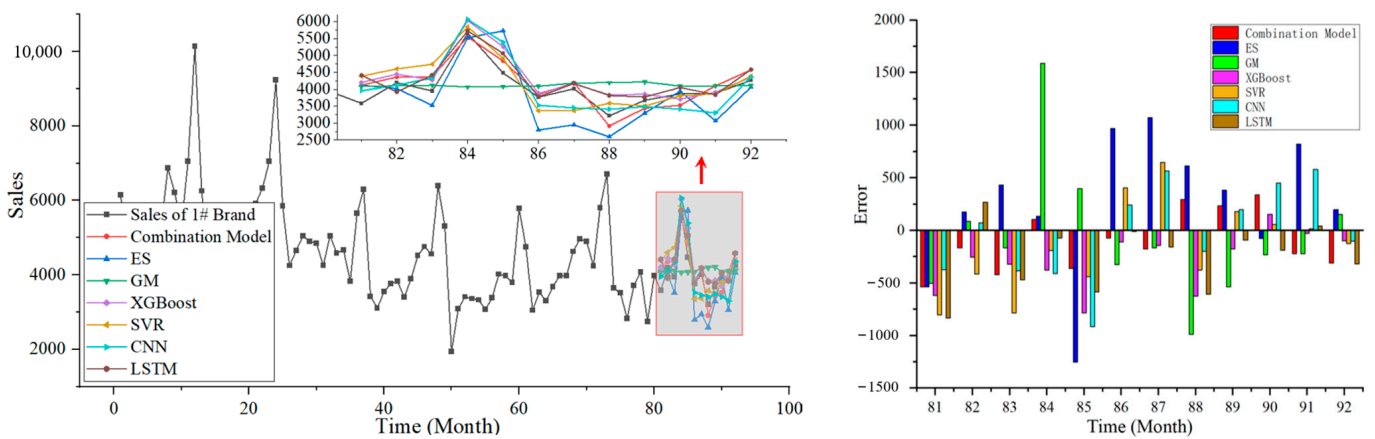
ES and GM use only historical sales data for forecasting, and their datasets are divided in the same way as are those used in the trend-extrapolation model. XGBoost, SVR, CNN, and LSTM use features data for forecasting, and their datasets are divided in the same way as that used in the causal analysis model. The comparison images of sales forecast values and forecast residuals of each model are shown in Figure 14, and the MAPE values of forecast results are shown in Table 11. XGBoost, SVR, ES, and LSTM forecast the seasonal characteristics in the sales sequence to some extent, but the forecast results of GM reflect only the trend.

For the number-one-brand data, the MAPE of the forecast results from the proposed combination model is the lowest, being 51.2%, 36.42%, 11.15%, 26.74%, 24.84%, and 12.28% lower than those of ES, GM, XGBoost, SVR, CNN, and LSTM, respectively. In addition, the forecast residuals of the combination model are the most balanced, which proves that its forecast results have the lowest variance and deviation.

For the number-two-brand data, the MAPE of the forecast results from the proposed combination model is lower than those of ES, XGBoost, SVR, CNN, and LSTM, being lower by 21.32%, 69.99%, 69.93%, 68.91%, and 63.03%, respectively, and is close to that of the result of the GM prediction, with an increase of 2.59%. In addition, the forecast residuals of four models from months 84 to 91 are biased to one side of the 0 line, showing different degrees of bias, with the GM having the lowest bias, followed by the combined model.

For the number-three-brand data, the MAPE of the forecast results from the proposed combination model is the lowest, being 63.07%, 56.78%, 74.2%, 71.51%, 71.71%, and 49.36% lower than those of ES, GM, XGBoost, SVR, CNN, and LSTM, respectively. In addition, the forecast residuals of the combination model are the most balanced, which proves that its forecast results have the lowest variance and deviation.

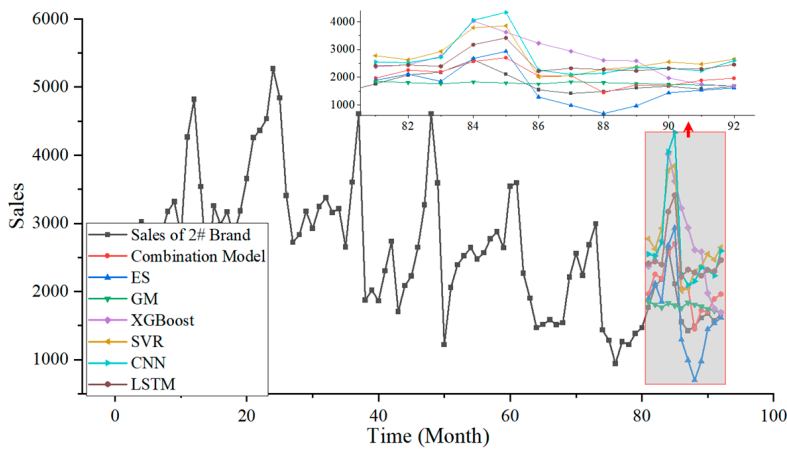
In summary, the proposed combination model shows the best forecasting performance.



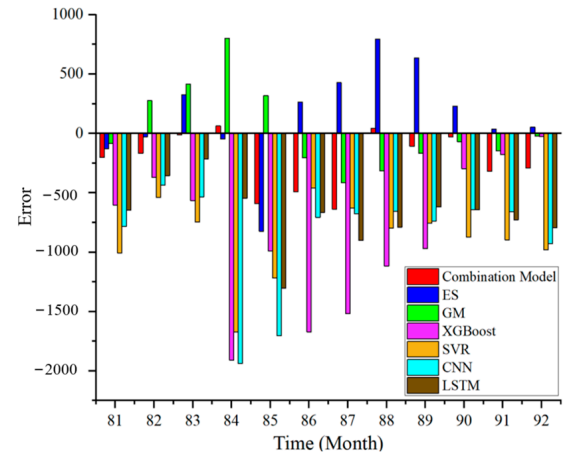
(a) Comparison of results from the proposed model and other models in sales forecasting of the number-one brand

(b) Comparison of errors from the proposed model and other models in sales forecasting of the number-one brand

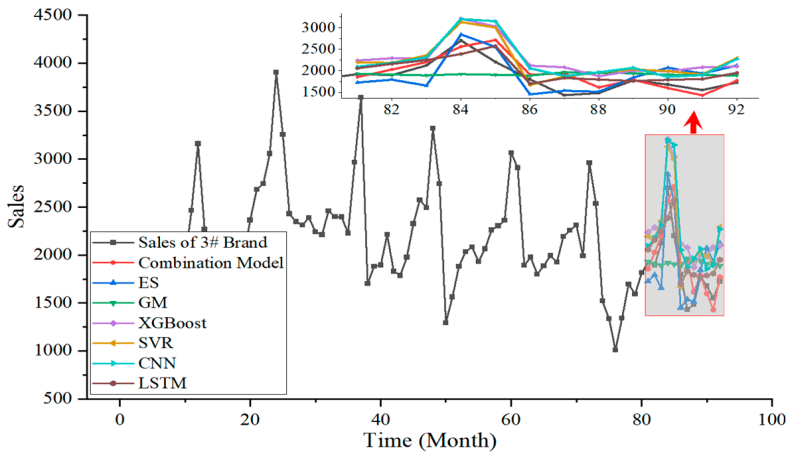
Figure 14. Cont.



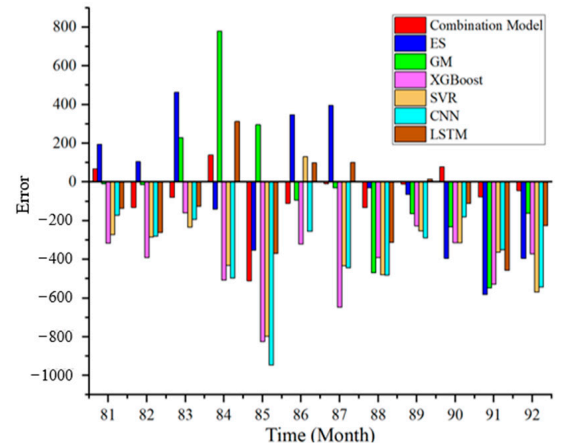
(c) Comparison of results from the proposed model and other models in sales forecasting of the number-two brand



(d) Comparison of errors from the proposed model and other models in sales forecasting of the number-two brand



(e) Comparison of results from the proposed model and other models in sales forecasting of the number-three brand



(f) Comparison of errors from the proposed model and other models in sales forecasting of the number-three brand

Figure 14. Comparisons of results and errors between the proposed model and other models in sales forecasting.

Table 11. Comparison of performance between the proposed combination model and other models.

Combination Strategy	MAPE		
	Number-One Brand	Number-Two Brand	Number-Three Brand
ES	0.1420	0.1862	0.1614
GM	0.1090	0.1428	0.1379
XGBoost	0.0780	0.4882	0.2310
SVR	0.0946	0.4872	0.2092
CNN	0.0922	0.4712	0.2107
LSTM	0.0790	0.3963	0.1177
Weighted-average parallel-combination model	0.0693	0.1465	0.0596

5. Conclusions

To evaluate the future demand for automobiles via sales forecasting, this paper analyzes the characteristics of the sales data by mining the historical predictors and external influences. A combination model based on trend extrapolation and causality analysis is proposed, and its effectiveness with automobile sales data with strong seasonality, strong trending, non-stationarity and nonlinearity was demonstrated through experiments. The conclusions are summarized as follows:

- (1) In view of the complex time-series characteristics of seasonality, trending, and non-stationarity of automobile sales data, a trend-extrapolation model based on SARIMA and PR is proposed to capture the historical predictors of sales series. By introducing time-series-analysis methods such as EMD, ADF, ACF and PACF, the sales series were reconstructed into a trend component and a broadly stationary and strongly seasonal component, which effectively solves the problem that SARIMA has difficulty modeling nonlinear and non-stationary series. In the ablation experiment, compared with EMD-SARIMA and SARIMA, the MAPE of the EMD-SARIMA-PR forecast results was reduced by 26.78% and 55.65%, respectively, for the number-one-brand data. For the number-two-brand data, the reductions were 68.61% and 64.33%, respectively. For the number-three-brand data, the reductions were 24.93% and 5.16%, respectively.
- (2) Since the trend-extrapolation submodel can establish only a linear model, the GBDT was used to construct a causality-analysis submodel. Compared with the trend-extrapolation submodel, the variance in the predictions of the causal-analysis submodel is smaller, but it is not sensitive to the change trend of the sales series, and the predictions may be biased.
- (3) The weighted-average parallel-combination model has the best forecast ability and can reduce the variance or deviation of the predictions of the two submodels. Compared with the stacking parallel-combination strategy and the boosting series-combination strategy, for the number-one-brand data, the MAPE value of the prediction result was reduced by 16.81% and 4.68%, respectively. For the number-two-brand data, the MAPE values of the predictions were reduced by 25.60% and 2.79%, respectively. For the number-three-brand data, the MAPE values of the predictions were reduced by 46.26% and 14.37%, respectively.
- (4) The proposed combination model has the best prediction effect according to the comparison experiments carried out with six basic models: ES, GM, XGBoost, SVR, CNN, and LSTM. For the number-one-brand data, the MAPE of the forecast results was reduced by 51.2%, 36.42%, 11.15%, 26.74%, 24.84% and 12.28%, respectively. For the number-two-brand data, compared with the predictions of ES, XGBoost, SVR, CNN, and LSTM, the MAPE was reduced by 21.32%, 69.99%, 69.93%, 68.91%, and 63.03%, respectively, although it was increased by 2.59% compared with GM. GM cannot predict the seasonality of the series. For the number-three-brand data, the MAPE was reduced by 63.07%, 56.78%, 74.2%, 71.51%, 71.71%, and 49.36%, respectively.

Author Contributions: Z.Z.: software, resources, data curation; H.C.: software, writing-original draft, visualization; L.W.: software, writing—review & editing; N.Z.: validation; F.W.: conceptualization, methodology. All authors have read and agreed to the published version of the manuscript.

Funding: This study is funded by National Nature Science Foundation of China (grant numbers 92266203) and Key Projects of Shijiazhuang Basic Research Program (241791077A).

Data Availability Statement: The data is available from the corresponding author.

Acknowledgments: We would like to express our gratitude to the National Natural Science Foundation of China and Shijiazhuang Science and Technology Bureau for funding.

Conflicts of Interest: Author Liyan Wu was employed by the company Xinfra Group Co. LTD. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Wu, Q. Product demand forecasts using wavelet kernel support vector machine and particle swarm optimization in manufacture system. *J. Comput. Appl. Math.* **2010**, *233*, 2481–2491. [[CrossRef](#)]
2. Danese, P.; Kalchschmidt, M. The role of the forecasting process in improving forecast accuracy and operational performance. *Int. J. Prod. Econ.* **2011**, *131*, 204–214. [[CrossRef](#)]
3. Mitra, R.; Saha, P.; Kumar Tiwari, M. Sales forecasting of a food and beverage company using deep clustering frameworks. *Int. J. Prod. Res.* **2023**, *62*, 3320–3332. [[CrossRef](#)]
4. Fu, W.; Jing, S.; Liu, Q.; Zhang, H. Resilient Supply Chain Framework for Semiconductor Distribution and an Empirical Study of Demand Risk Inference. *Sustainability.* **2023**, *15*, 7382. [[CrossRef](#)]
5. Van Belle, J.; Guns, T.; Verbeke, W. Using shared sell-through data to forecast wholesaler demand in multi-echelon supply chains. *Eur. J. Oper. Res.* **2021**, *288*, 466–479. [[CrossRef](#)]
6. Liu, R.; Vakharia, V. Optimizing Supply Chain Management through BO-CNN-LSTM for Demand Forecasting and Inventory Management. *J. Organ. End. User Com.* **2024**, *36*, 1–25. [[CrossRef](#)]
7. Moalem, S.; Ahari, R.M.; Shahgholian, G.; Moazzami, M.; Kazemi, S.M. Long-Term Electricity Demand Forecasting in the Steel Complex Micro-Grid Electricity Supply Chain—A Coupled Approach. *Energies* **2022**, *15*, 7972. [[CrossRef](#)]
8. Matsumoto, M.; Komatsu, S. Demand forecasting for production planning in remanufacturing. *Int. J. Adv. Manuf. Technol.* **2015**, *79*, 161–175. [[CrossRef](#)]
9. Gao, J.; Xie, Y.; Gu, F.; Xiao, W.; Hu, J.; Yu, W. A hybrid optimization approach to forecast automobile sales of China. *Adv. Mech. Eng.* **2017**, *9*. [[CrossRef](#)]
10. Rožanec, J.M.; Kažič, B.; Škrjanc, M.; Fortuna, B.; Mladenčić, D. Automotive OEM Demand Forecasting: A Comparative Study of Forecasting Algorithms and Strategies. *Appl. Sci.* **2021**, *11*, 6787. [[CrossRef](#)]
11. Vijayakumar, S.; Suresh, P. Sales Demand Forecasting in Car Industry Using Seagull Optimization Based Holt Winter and Quantile Regression Neural Network. *Wireless Pers. Commun.* **2023**, *133*, 49–72. [[CrossRef](#)]
12. Ou-Yang, C.; Chou, S.; Juan, Y. Improving the Forecasting Performance of Taiwan Car Sales Movement Direction Using Online Sentiment Data and CNN-LSTM Model. *Appl. Sci.* **2022**, *12*, 1550. [[CrossRef](#)]
13. Evangelos, S.; Vassilios, A.; Konstantinos, N. Forecasting with a hybrid method utilizing data smoothing, a variation of the Theta method and shrinkage of seasonal factors. *Int. J. Prod. Econ.* **2019**, *209*, 92–102. [[CrossRef](#)]
14. Doornik, J.; Castle, J.; Hendry, D. Short-term forecasting of the coronavirus pandemic. *Int. J. Forecast.* **2022**, *38*, 453–466. [[CrossRef](#)] [[PubMed](#)]
15. Deina, C.; dos Santos, J.; Biuk, L.; Lizot, M.; Converti, A.; Siqueira, H.; Trojan, F. Forecasting Electricity Demand by Neural Networks and Definition of Inputs by Multi-Criteria Analysis. *Energies* **2023**, *16*, 1712. [[CrossRef](#)]
16. Jun, P.; Jun, D. Predicting Chinese tourists' B&B preferences through a method of online reviews causality analytic. *Inform. Process Manag.* **2024**, *61*, 103634. [[CrossRef](#)]
17. Aljarrah, O.; Li, J.; Huang, W.; Heryudono, A.; Bi, J. ARIMA-GMDH: A low-order integrated approach for predicting and optimizing the additive manufacturing process parameters. *Int. J. Adv. Manuf. Technol.* **2020**, *106*, 701–717. [[CrossRef](#)]
18. Sumer, K.; Goktas, O.; Hepsag, A. The application of seasonal latent variable in forecasting electricity demand as an alternative method. *Energy Policy* **2009**, *37*, 1317–1322. [[CrossRef](#)]
19. Anh, N.; Anh, N.; Tang, T.; Solanki, V.; Crespo, R.; Dat, N. Online SARIMA applied for short-term electricity load forecasting. *Appl. Intell.* **2024**, *54*, 1003–1019. [[CrossRef](#)]
20. Pritularga, K.; Svetunkov, I.; Kourentzes, N. Shrinkage estimator for exponential smoothing models. *Int. J. Forecast.* **2023**, *39*, 1351–1365. [[CrossRef](#)]
21. Wang, C.H.; Chen, T.Y. Combining biased regression with machine learning to conduct supply chain forecasting and analytics for printing circuit board. *Int. J. Syst. Sci.-Oper.* **2022**, *9*, 143–154. [[CrossRef](#)]
22. Chen, Y.; Liu, H.; Hsieh, H. Time series interval forecast using GM(1,1) and NGBM(1,1) models. *Soft Comput.* **2019**, *23*, 1541–1555. [[CrossRef](#)]
23. Andueza, A.; Del Arco-Osuna, M.; Fornés, B.; González-Crespo, R.; Martín-Alvarez, J. Using the Statistical Machine Learning Models ARIMA and SARIMA to Measure the Impact of COVID-19 on Official Provincial Sales of Cigarettes in Spain. *Int. J. Interact. Multi* **2023**, *8*, 73–87. [[CrossRef](#)]
24. Choi, T.; Yu, Y.; Au, K. A hybrid SARIMA wavelet transform method for sales forecasting. *Decis. Support. Syst.* **2011**, *51*, 130–140. [[CrossRef](#)]
25. Williams, A.; Sperl, R.; Chung, S. Anomaly Detection in Multi-Seasonal Time Series Data. *IEEE Access* **2023**, *11*, 106456–106464. [[CrossRef](#)]
26. Zhou, B.; Liu, C.; Li, J.; Sun, B.; Yang, J. A Hybrid Method for Ultrashort-Term Wind Power Prediction considering Meteorological Features and Seasonal Information. *Math. Probl. Eng.* **2020**, *2020*, 1795486. [[CrossRef](#)]
27. Lippi, M.; Bertini, M.; Frasconi, P. Short-Term Traffic Flow Forecasting: An Experimental Comparison of Time-Series Analysis and Supervised Learning. *IEEE trans. Intell. Transp. Syst.* **2013**, *14*, 871–882. [[CrossRef](#)]
28. Zhang, W.; Lin, Z.; Liu, X. Short-term offshore wind power forecasting—A hybrid model based on Discrete Wavelet Transform (DWT), Seasonal Autoregressive Integrated Moving Average (SARIMA), and deep-learning-based Long Short-Term Memory (LSTM). *Renew. Energy* **2022**, *185*, 611–628. [[CrossRef](#)]

29. Sriram, L.; Gilanifar, M.; Zhou, Y.; Ozguven, E.; Arghandeh, R. Causal Markov Elman Network for Load Forecasting in Multi-network Systems. *IEEE Trans. Ind. Electron.* **2019**, *66*, 1434–1442. [[CrossRef](#)]
30. Cheng, X.; Wu, P.; Liao, S.; Wang, X. An integrated model for crude oil forecasting: Causality assessment and technical efficiency. *Energy Econ.* **2023**, *117*, 106467. [[CrossRef](#)]
31. Zhang, M.; Xu, H.; Ma, N.; Pan, X. Intelligent Vehicle Sales Prediction Based on Online Public Opinion and Online Search Index. *Sustainability* **2022**, *14*, 10344. [[CrossRef](#)]
32. Xia, Z.; Xue, S.; Wu, L.; Sun, J.; Chen, Y.; Zhang, R. ForeXGBoost: Passenger car sales prediction based on XGBoost. *Distrib. Parallel Dat* **2020**, *38*, 713–738. [[CrossRef](#)]
33. Dai, D.; Fang, Y.; Wang, S.; Zhao, M. Prediction of China Automobile Market Evolution Based on Univariate and Multivariate Perspectives. *Systems* **2023**, *11*, 431. [[CrossRef](#)]
34. Li, F.; Sun, L.; Kong, N.; Zhang, H.; Mo, L. Sales Forecasting Method for Inventory Replenishment Systems of Vehicle Energy Stations Without Stockouts. *IEEE Trans. Eng. Manag.* **2023**, *71*, 6568–6580. [[CrossRef](#)]
35. Liu, B.; Song, C.; Liang, X.; Lai, M.; Yu, Z.; Ji, J. 2023 Regional differences in China's electric vehicle sales forecasting: Under supply-demand policy scenarios. *Energy Policy* **2023**, *177*, 113554. [[CrossRef](#)]
36. Wang, X.; Hyndman, R.; Li, F.; Kang, Y. Forecast combinations: An over 50-year review. *Int. J. Forecast.* **2023**, *39*, 1518–1547. [[CrossRef](#)]
37. Muhammed Anvar, P.; Balakrishna, N. Some weighted mixed portmanteau tests for diagnostic checking in linear time series models. *J. Stat. Comput. Simul.* **2018**, *88*, 3000–3017. [[CrossRef](#)]
38. Ganaie, M.A.; Hu, M.H.; Malik, A.K.; Tanveer, M.; Suganthan, P.N. Ensemble deep learning: A review. *Eng. Appl. Artif. Intell.* **2022**, *115*, 105151. [[CrossRef](#)]
39. Wang, Q.; Song, L.; Zhao, J.; Wang, H.; Dong, L.; Wang, X.; Yang, Q. Application of the gradient boosting decision tree in the online prediction of rolling force in hot rolling. *Int. J. Adv. Manuf. Technol.* **2023**, *125*, 387–397. [[CrossRef](#)]
40. Parisouj, P.; Jun, C.; Bateni, S.M.; Heggy, E.; Band, S.S. Machine learning models coupled with empirical mode decomposition for simulating monthly and yearly streamflows: A case study of three watersheds in Ontario, Canada. *Eng. Appl. Comp. Fluid* **2023**, *17*, 2242445. [[CrossRef](#)]
41. Zhang, S.M.; Wang, S.J. Spectral radius-based interval principal component analysis (SR-IPCA) for fault detection in industrial processes with imprecise data. *J. Process Control* **2022**, *114*, 105–119. [[CrossRef](#)]
42. You, W.; Guo, D.; Wu, Y.; Li, W. Multiple Load Forecasting of Integrated Energy System Based on Sequential-Parallel Hybrid Ensemble Learning. *Energies* **2023**, *16*, 3268. [[CrossRef](#)]
43. Hajirahimi, Z.; Khashei, M. Hybrid structures in time series modeling and forecasting: A review. *Eng. Appl. Artif. Intell.* **2019**, *86*, 83–106. [[CrossRef](#)]
44. Yang, Y.; Chen, Y.; Wang, Y.; Li, C.; Li, L. Modelling a combined method based on ANFIS and neural network improved by DE algorithm: A case study for short-term electricity demand forecasting. *Appl. Soft Comput.* **2016**, *49*, 663–675. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.