

# 我国新能源汽车销量的长短期预测研究

袁 铮\*, 谭 琦

云南财经大学统计与数学学院, 云南 昆明

收稿日期: 2023年3月26日; 录用日期: 2023年4月21日; 发布日期: 2023年4月29日

## 摘 要

近些年来, 随着人们的环保意识与能源意识的增强, 以及新能源汽车配套基础设施和相关鼓励政策的完善, 新能源汽车的发展正在迎来一个爆发期。在当前时代背景下, 预测和分析新能源汽车的产销情况已成为一个备受关注的研究方向, 该方向的研究成果有利于更好地辅助政府、企业及个人等多方的决策。本文针对新能源汽车整体销量的预测, 构建了年度销量灰色预测模型、季度销量ARIMA预测模型与月度销量SARIMA预测模型。长期和短期的预测结果均表明, 新能源汽车发展势头良好, 未来将占据更大的汽车市场份额, 成为广大消费者的不二之选。

## 关键词

新能源汽车销量预测, 灰色预测模型, SARIMA预测模型

# Research on Long-Term and Short-Term Forecasting of New Energy Vehicle Sales in China

Zheng Yuan\*, Qi Tan

School of Statistics and Mathematics, Yunnan University of Finance and Economics, Kunming Yunnan

Received: Mar. 26<sup>th</sup>, 2023; accepted: Apr. 21<sup>st</sup>, 2023; published: Apr. 29<sup>th</sup>, 2023

## Abstract

In recent years, with the increasing awareness of environmental protection and energy conservation, as well as the improvement of supporting infrastructure and related encouragement policies for new energy vehicles, the development of new energy vehicles is entering an explosive period.

\*通讯作者。

**In the current era, predicting and analyzing the production and sales of new energy vehicles has become a research direction that attracts much attention. The research results of this direction are conducive to better assisting the decision-making of governments, enterprises, individuals, and other parties. In order to predict the overall sales of new energy vehicles, this paper provides a grey forecasting model for the annual sales, an ARIMA forecasting model for the quarterly sales, and a SARIMA forecasting model for the monthly sales. The long-term and short-term forecasting results show that the development of new energy vehicles is promising and will occupy a larger share of the automobile market in the future, becoming the first choice for consumers.**

## Keywords

**New Energy Vehicle Sales Forecast, Grey Forecasting Model, SARIMA Forecasting Model**

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

随着经济社会的不断发展,我国人民日益提高的生活水平以及不断扩大的出行范围都促使汽车成为了不可或缺的“家庭成员”。然而,随之而来的汽车保有量的增长也带来了一系列负面影响,除了交通拥堵和车位紧缺外,环境与能源问题也日益严峻。为了贯彻新发展理念,实现经济又好又快的发展,协调人与自然和谐发展,发展新能源汽车已成为大势所趋。无论是在环境友好性方面,还是在动力性、舒适性等方面,新能源汽车都凭借绝对的优势向我们证明,大力发展和支持新能源汽车是正确可行的。相信在全球政策以及需求的共振下,新能源汽车将在发展的赛道上高速前进。

如今,对反映经济活动各方面情况的变量进行合理的预测已成为经济活动参与者进行更好地管理或最大化其利益的重要法宝。预测的结果能够很好地辅助相关经济主体制定有效的战略计划。同样地,对新能源汽车销量的预测也具有重要意义。对于企业而言,预测新能源汽车的销量有利于把握市场的发展动态,为生产、库存、销售等环节制定合理战略,识别其竞争优势。对于政府而言,把握销量的发展情况有利于更好地制定相关政策以及建设基础设施。对于消费者而言,销量数据是了解新能源汽车行业发展状况的重要信息,向消费者提供新能源汽车行业的前景预测有利于新能源汽车市场的推广。因此,对于各方而言,对新能源汽车销量的预测都具有重要的意义。

目前,绝大多数的研究都是在某一个特定的时间尺度(月度、季度或者年度)下进行的,而本文将结合长期和短期不同的时间尺度进行研究。

## 2. 新能源汽车销量预测的研究现状与总结

相较于预测新能源汽车销量的研究,对传统燃油车销量进行预测的研究更加成熟。早在 20 世纪 50 年代, Nerlove [1]就基于汽车新增需求、上期报废数量和保有量之间的关系构建回归模型,来对美国汽车的保有量进行短期预测。陈道平[2]收集了 2001 年 1 月到 2011 年 6 月的我国汽车销量月度数据,构建 ARIMA 模型对我国汽车整体销量进行了预测分析。胡彦君[3]通过预测汽车销量帮助政府及相关部门制定交通布局、停车位增设和城市绿化等方面的政策。章旭[4]提出基于网络大数据和传统时间序列分析的汽车销售预测 BOAR 模型,将汽车销量的预测细化至单一汽车品牌的粒度,并开创性地提出一种基于 MARS 变量选择过程和 BP 神经网络的 MISF 模型。有关传统燃油汽车销量预测的研究成果为新能源汽车

销量的预测研究提供了丰富的研究思路与研究方法。

对于预测新能源汽车销量的研究,目前大多数的研究都是从固定周期的角度进行的。常用到的预测模型包括自回归移动平均模型、线性回归模型、灰色理论模型和 BP 神经网络模型等。周彦福[5]等根据销量影响因素的影响能力构建基于果蝇算法优化的灰色神经网络模型进行月度预测。刘颖琦[6]等以丰田普锐斯混合动力汽车的历史全球销量数据作为基础数据,由 Bass 模型预测中国新能源汽车产业的总体走势、典型新能源车型的未來市场表现。

Andrawis [7]等指出短期预测与长期预测相结合是提高预测准确性的有效方法,月度数据、季度数据和年度数据等不同时间尺度的数据体现出不同的动态特征,其研究证实长短期预测相结合的方法优于单一模型的预测。刘媛[8]收集了 2014 年至 2019 年 6 年 24 个季度 72 个月的新能源汽车销量数据,通过构建 SARMA 模型和奇异谱分析模型两个模型,对新能源汽车销量进行了短期月度预测,由 ARMA 模型和基于数据分组方法的灰色预测模型预测长期季度销量。通过比较不同预测周期下,不同模型的预测效果,刘媛指出,由于新能源汽车的销量数据有限,相比 ARMA 模型与 SARMA 模型,灰色系统预测模型和奇异谱分析模型等小样本预测模型对新能源汽车销量的预测更加准确。白一凡[9]的研究结果表明 SARIMA 模型与 BP 神经网络的组合模型对新能源汽车月度销量的预测结果在多方面都优于单一的 SARIMA 模型。不过,不可否认的是,随着我国新能源汽车销量样本数据的增加,ARIMA 模型等时间序列分析模型的预测效果依然会是理想的。

经过对国内外相关研究的分析,可以看出大部分研究的预测对象都是我国新能源汽车的整体销售量,而对某一特定品牌的新能源汽车销量或是某一特定新能源汽车车型的销量的预测研究仍较少。此外,大多数的研究局限于应用单一类型的模型进行预测,而且很少有研究结合长期和短期两个不同的视角进行研究。

### 3. 相关理论及方法

#### 3.1. ARIMA 模型

非平稳序列在差分后往往会显示平稳序列的特征,ARIMA 模型可拟合差分平稳序列的发展。

##### 3.1.1. ARIMA 模型的结构

求和自回归移动平均(autoregressive integrated moving average)模型,简记为 ARIMA ( $p, d, q$ )模型:

$$\begin{cases} \Phi(B)\nabla^d x_t = \Theta(B)\varepsilon_t \\ E(\varepsilon_t) = 0, \text{Var}(\varepsilon_t) = \sigma_\varepsilon^2, E(\varepsilon_t \varepsilon_s) = 0, s \neq t \\ E(x_s \varepsilon_t) = 0, \forall s < t \end{cases} \quad (1)$$

其中,  $\nabla^d = (1-B)^d$ 。  $\Phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$ ,  $\Theta(B) = 1 - \theta_1 B - \dots - \theta_q B^q$  分别为平稳可逆 ARMA( $p, q$ )模型的自回归系数多项式和移动平均系数多项式。

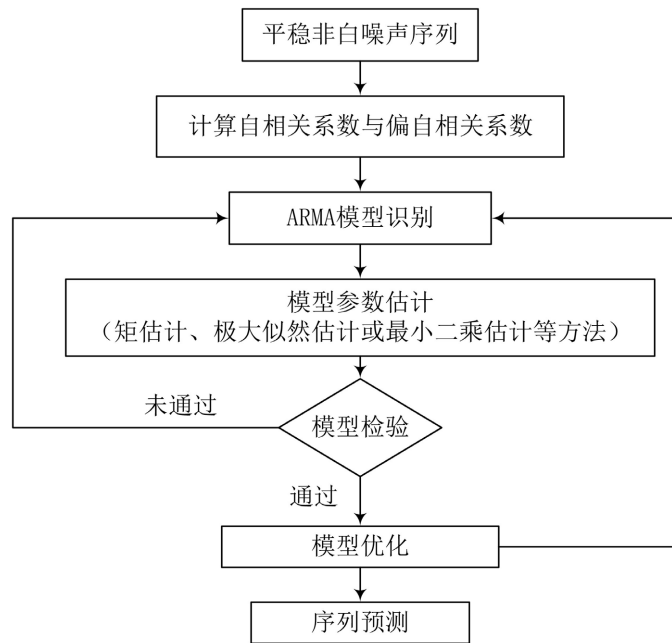
$d$  阶差分后的序列可用下式表示:

$$\nabla^d x_t = \sum_{i=0}^d (-1)^i C_d^i x_{t-i} \quad (2)$$

其中,  $C_d^i = \frac{d!}{i!(d-i)!}$ 。ARIMA 模型实质就是差分运算与 ARMA 模型的组合[10]。

##### 3.1.2. ARIMA 模型建模

对于一个非平稳序列,若其能在适当阶数的差分后实现平稳,那么可由 ARMA 模型拟合该差分平稳非白噪声序列。ARMA 模型的建模步骤见图 1。



**Figure 1.** Modeling procedure  
**图 1.** 建模步骤

### 3.2. SARIMA 模型

在现实生活中, 很多的时间序列都带有季节效应, 从而呈现出周期性波动的规律。ARIMA 模型也可对这些具有季节效应的时间序列进行建模, 相应地, 可将这类考虑季节效应的 ARIMA 模型称为 SARIMA (Seasonal Autoregressive Integrated Moving Average)模型。根据季节效应的提取方式不同, 考虑序列季节效应的 ARIMA 模型可分为 ARIMA 加法模型与 ARIMA 乘法模型。

对于新能源汽车月度销量的预测, 本文采用了 ARIMA 加法模型。

ARIMA 加法模型是指序列的季节效应与其他效应之间是加法关系, 这意味着序列中的季节信息可由周期步长差分提取, 趋势信息可由低阶差分提取。

ARIMA 加法模型的结构如下:

$$\nabla_s \nabla^d x_t = \frac{\Theta(B)}{\Phi(B)} \varepsilon_t \tag{3}$$

其中,  $S$  表示周期步长。 $d$  为提取趋势信息所用的差分阶数。 $\{\varepsilon_t\}$  为白噪声序列, 且  $E(\varepsilon_t) = 0, Var(\varepsilon_t) = \sigma_\varepsilon^2$ 。 $\Phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$  是  $p$  阶自回归系数多项式。 $\Theta(B) = 1 - \theta_1 B - \dots - \theta_q B^q$  是  $q$  阶移动平均系数多项式。ARIMA 加法模型简记为 ARIMA  $(p, (d, S), q)$  或 ARIMA  $(p, d, q) \times (0, 1, 0)_S$ 。

### 3.3. 灰色预测模型 GM (1, 1)

灰色预测模型具备小样本建模的突出优点, 利用微分方程充分挖掘系统本质, 可将无规律的原始数据进行生成得到规律性较强的生成序列[11]。GM (1, 1)表示模型是一阶微分方程, 且只含 1 个变量的灰色模型。运用 GM (1, 1)模型进行预测的步骤如下。

#### 3.3.1. 数据的检验与处理

首先对原始序列  $x^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n))$  的级比进行计算, 以确保建模方法的可行性, 级比

计算公式如下:

$$\lambda(k) = \frac{x^{(0)}(k-1)}{x^{(0)}(k)}, k = 2, 3, \dots, n \quad (4)$$

只有在所有的级比都位于可容覆盖内  $\Theta = \left( e^{-\frac{2}{n+1}}, e^{\frac{2}{n+2}} \right)$  时, 序列才可作为 GM (1, 1) 模型的数据, 否则需通过平移变换使新序列的级比落在可容覆盖内。平移变换公式如下:

$$y^{(0)}(k) = x^{(0)}(k) + c, k = 1, 2, \dots, n \quad (5)$$

### 3.3.2. 模型的建立

如果序列  $x^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n))$  通过了级比检验, 则可进一步建立灰色模型。首先, 对原始序列进行 1 次累加操作:

$$\begin{aligned} x^{(1)} &= (x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n)) \\ &= (x^{(0)}(1), x^{(0)}(1) + x^{(0)}(2), \dots, x^{(0)}(1) + \dots + x^{(0)}(n)) \end{aligned} \quad (6)$$

再对原始序列  $x^{(0)}$  的 1 次累加生成序列  $x^{(1)}$  进行均值生成操作:

$$z^{(1)} = (z^{(1)}(2), z^{(1)}(3), \dots, z^{(1)}(n)) \quad (7)$$

$$z^{(1)}(k) = 0.5x^{(1)}(k-1) + 0.5x^{(1)}(k), k = 2, 3, \dots, n \quad (8)$$

进一步可建立灰微分方程:

$$x^{(0)}(k) + ax^{(1)}(k) = b, k = 2, 3, \dots, n \quad (9)$$

相应的白化微分方程为:

$$\frac{dx^{(1)}}{dt} + ax^{(1)}(t) = b \quad (10)$$

记  $u = [a, b]^T$ ,  $Y = [x^{(0)}(2), x^{(0)}(3), \dots, x^{(0)}(n)]^T$ ,  $B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix}$ 。由最小二乘法求得使

$J(u) = (Y - Bu)^T (Y - Bu)$  取最小值的  $u$  的估计值为:

$$\hat{u} = (B^T B)^{-1} B^T Y \quad (11)$$

最后可得到预测值为:

$$\hat{x}^{(1)}(k+1) = \left( x^{(0)}(1) - \frac{\hat{b}}{\hat{a}} \right) e^{-\hat{a}k} + \frac{\hat{b}}{\hat{a}}, k = 0, 1, \dots, n-1, \dots \quad (12)$$

$$\hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k), k = 1, 2, \dots, n-1, \dots \quad (13)$$

### 3.3.3. 预测值的检验

对模型预测结果的检验主要有残差检验与级比偏差值检验。其中, 残差检验的公式如下:

$$\varepsilon(k) = \frac{x^{(0)}(k) - \hat{x}^{(0)}(k)}{x^{(0)}(k)}, k = 1, 2, \dots, n \quad (14)$$

注意  $\hat{x}^{(0)}(1) = x^{(0)}(1)$ 。如果  $\varepsilon(k) < 0.2$ ，则可认为达到一般要求；而  $\varepsilon(k) < 0.1$ ，则意味着预测达到较高的要求。

级比偏差计算公式如下：

$$\rho(k) = 1 - \left( \frac{1 - 0.5a}{1 + 0.5a} \right) \lambda(k), k = 2, 3, \dots, n \quad (15)$$

若  $\rho(k) < 0.1$ ，则可认为达到较高的要求。

### 3.3.4. 预测预报

根据实际问题的需要，由 GM(1, 1) 预测模型给出相应的预测值。

## 4. 新能源汽车年度销量预测研究

### 4.1. 数据采集

本文使用 2011~2020 年共 10 年的新能源汽车年度销量数据进行预测模型的搭建，数据来源于中国汽车工业协会。原始数据时序图如图 2 所示。

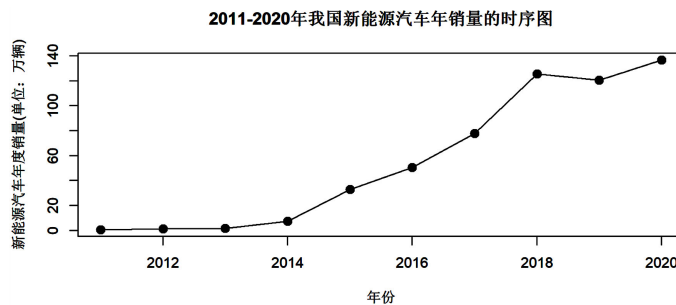


Figure 2. Time series diagram of the annual sales of new energy vehicles in China  
图 2. 我国新能源汽车年度销量序列时序图

### 4.2. 时间序列数据的预处理

原始序列的时序图呈现显著的上升趋势，这是典型的非平稳序列特征。本文用差分方法充分提取确定性信息，原始序列差分运算后的序列时序图见图 3。

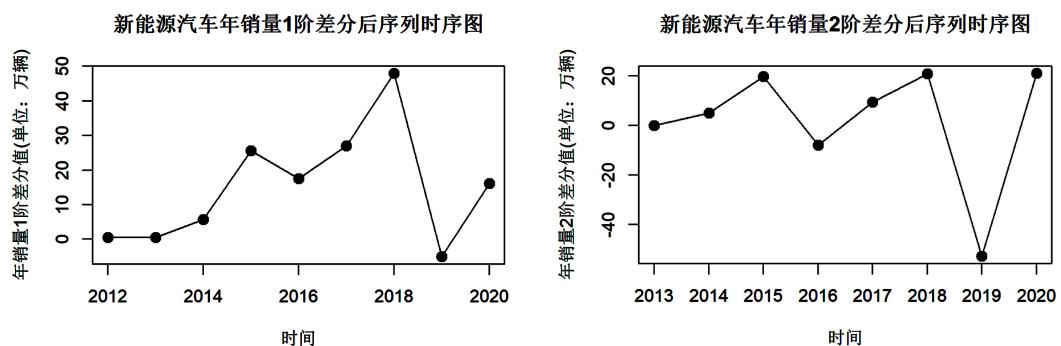


Figure 3. Sequence diagram after difference operation  
图 3. 差分运算后序列的时序图

1 阶差分后的序列仍呈现长期递增的趋势, 而且 ADF 检验结果表明 1 阶差分序列是非平稳的。通过对原始序列进行二阶差分运算, 对应时序图显示二阶差分运算较充分地提取了原序列所蕴含的长期趋势。进一步地, 图 4 的 ADF 检验结果表明二阶差分后的序列是平稳的。

#### Augmented Dickey-Fuller Test

alternative: stationary

Type 1: no drift no trend				Type 2: with drift no trend				Type 3: with drift and trend			
	lag	ADF	p. value		lag	ADF	p. value		lag	ADF	p. value
[1,]	0	-4.297	0.0100	[1,]	0	-3.929	0.010	[1,]	0	-4.40	0.0100
[2,]	1	-2.236	0.0269	[2,]	1	-2.563	0.128	[2,]	1	-3.74	0.0401
[3,]	2	-0.667	0.4303	[3,]	2	0.942	0.990	[3,]	2	NaN	NaN

Note: in fact, p. value = 0.01 means p. value  $\leq$  0.01

Figure 4. Results of ADF unit root test

图 4. ADF 单位根检验结果

纯随机性检验结果见图 5。

#### Box-Ljung test

X-squared = 2.9781, df = 2, p-value = 0.2256

X-squared = 5.4337, df = 4, p-value = 0.2456

Figure 5. Results of pure randomness test for stationary sequence

图 5. 平稳序列纯随机性检验结果

纯随机性检验结果表明二阶差分后的序列为白噪声序列, 故可停止相关统计分析。基于上述时间序列预处理的结果, 本文采用 GM(1, 1) 预测模型分析未来新能源汽车年度销量的发展趋势。

### 4.3. 灰色预测模型 GM(1, 1) 的建立

#### 4.3.1. 数据的检验与处理

由于这 10 个年度的销量数据的级比并未全部处于可容覆盖(0.8338, 1.1814)内, 故本文选取常数  $C = 165$  对原始数据进行平移变换。

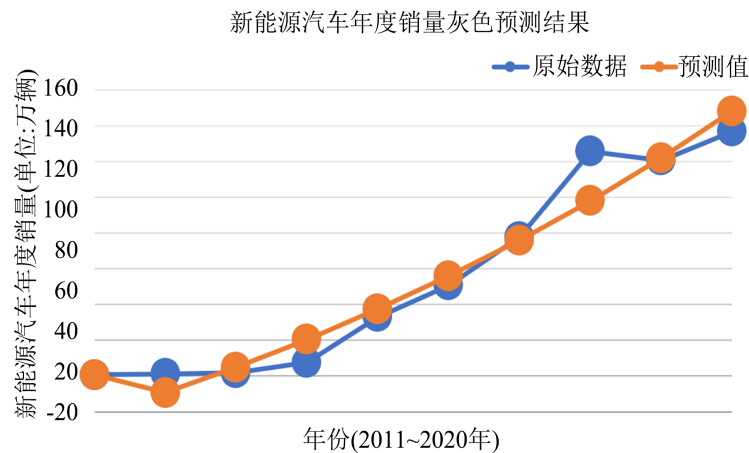
#### 4.3.2. 模型建立与检验

利用变换后的数据建立预测模型, 可得到预测值见图 6 和图 7, 预测模型中两个灰参数  $a$ 、 $b$  的值分别为:  $-0.08723151$ ,  $134.6529$ 。

年份	原始数据	平移变换后的数据	变换后数据的预测值	预测值	残差	相对误差	级比偏差
2011	0.8	165.8	165.8000	0.8000	0.0000	0.0000	
2012	1.3	166.3	155.8130	-9.1870	10.4870	0.0631	-0.0879
2013	1.8	166.8	170.0153	5.0153	-3.2153	0.0193	-0.0879
2014	7.5	172.5	185.5120	20.5120	-13.0120	0.0754	-0.0552
2015	33.1	198.1	202.4213	37.4213	-4.3213	0.0218	0.0498
2016	50.7	215.7	220.8719	55.8719	-5.1719	0.0240	-0.0022
2017	77.7	242.7	241.0042	76.0042	1.6958	0.0070	0.0302
2018	125.6	290.6	262.9715	97.9715	27.6285	0.0951	0.0887
2019	120.6	285.6	286.9412	121.9412	-1.3412	0.0047	-0.1103
2020	136.7	301.7	313.0957	148.0957	-11.3957	0.0378	-0.0330

Figure 6. Prediction and test results of the GM(1, 1) model

图 6. GM(1, 1) 模型的预测及检验结果



**Figure 7.** Comparison between the prediction of GM (1, 1) model with the actual values  
**图 7.** GM (1, 1)模型的预测结果与实际值的比较

图 6 中的相对误差与级比偏差两个检验值均是由平移变换后的数据计算得到的, 相对误差值都小于 0.1, 则可认为预测达到较高的要求。同样地, 由级比偏差值也可验证该模型的精度较高, 可进一步进行预测。

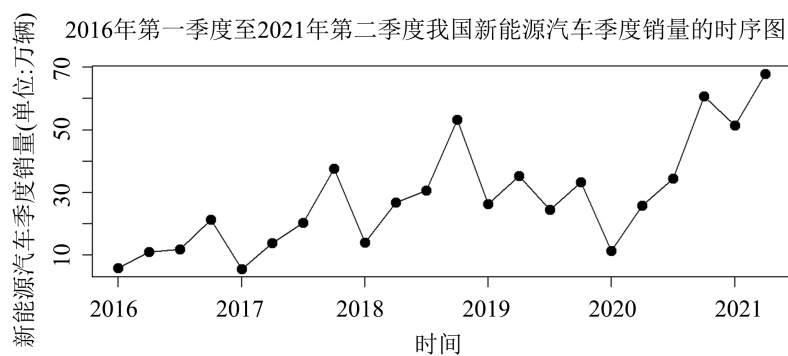
### 4.3.3. 模型预测

结合 GM (1, 1)模型对 2021 年新能源汽车年度销量的预测值(176.63 万辆)以及当前新能源汽车发展的良好势头, 本文认为新能源汽车在 2021 年度一定会取得不错的销量成绩。

## 5. 新能源汽车季度销量预测研究

### 5.1. 数据采集

本文使用 2016 年第一季度至 2021 年第三季度共 23 个季度的新能源汽车季度销量数据进行预测模型的搭建(其中, 前 22 个季度的数据用作训练集而第 23 个季度的数据用作测试), 数据来源于中国汽车工业协会。原始数据时序图如图 8 所示。

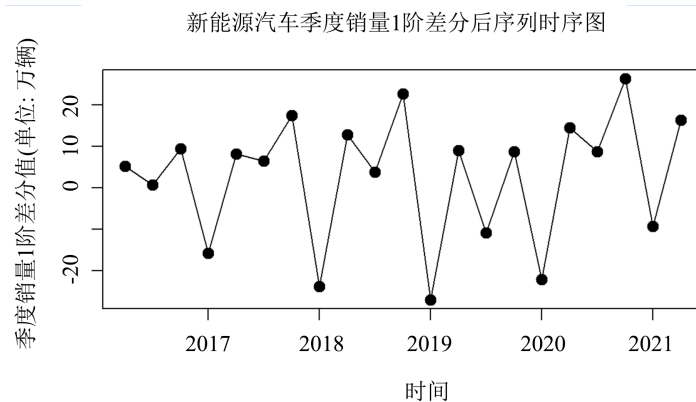


**Figure 8.** Time series diagram of the quarterly sales of new energy vehicles in China  
**图 8.** 新能源汽车季度销量序列时序图

### 5.2. 时间序列数据的预处理

原始序列的时序图显示该序列呈现上升趋势且蕴含固定的周期, 可认为原始序列是非平稳的。本文考虑用差分运算充分提取原始序列中的确定性信息, 原始序列一阶差分后的序列时序图如图 9 所示。





**Figure 9.** Sequence diagram of Chinese new energy vehicles quarterly sales after first-order difference operation  
**图 9.** 我国新能源汽车季度销量 1 阶差分序列的时序图

1 阶差分有效地提取了原序列中递增的趋势, 进一步地, 图 10 的 ADF 检验结果表明了 1 阶差分后的序列是平稳的。

#### Augmented Dickey-Fuller Test

alternative: stationary

Type 1: no drift no trend	Type 2: with drift no trend	Type 3: with drift and trend
lag ADF p. value	lag ADF p. value	lag ADF p. value
[1,] 0 -7.28 0.01	[1,] 0 -7.53 0.0100	[1,] 0 -7.43 0.010
[2,] 1 -2.87 0.01	[2,] 1 -3.05 0.0455	[2,] 1 -3.04 0.171
[3,] 2 -3.87 0.01	[3,] 2 -4.30 0.0100	[3,] 2 -4.38 0.010

Note: in fact, p. value = 0.01 means p. value  $\leq$  0.01

**Figure 10.** Results of ADF unit root test

**图 10.** ADF 单位根检验结果

对 1 阶差分后的序列进行纯随机性检验:

#### Box-Ljung test

X-squared = 42.201, df = 6, p-value = 1.678e-07

X-squared = 72.747, df = 12, p-value = 9.773e-11

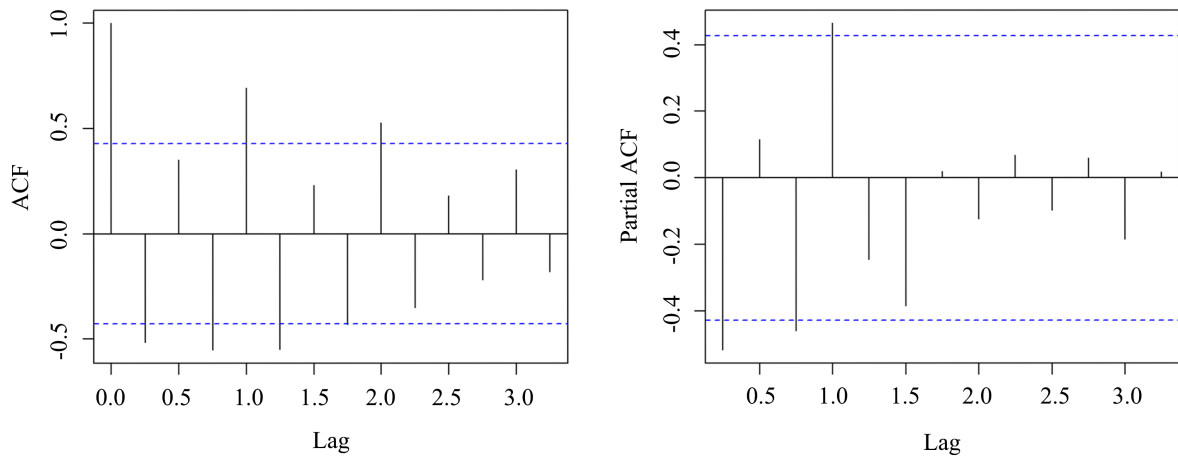
**Figure 11.** Results of pure randomness test for the quarterly sales after first-order difference operation

**图 11.** 季度销量序列 1 阶差分后序列的纯随机性检验

如图 11 所示, 纯随机性检验结果表明, 6 阶和 12 阶延迟的 LB 统计量的 P 值都远小于 0.05 的显著性水平, 故可采用 ARMA 模型拟合该 1 阶差分序列的发展。

### 5.3. 平稳序列的拟合与预测

1 阶差分后序列的自相关图和偏自相关图见图 12。自相关系数图显示 1 阶、3 阶、4 阶、5 阶和 8 阶的自相关系数都显著大于 2 倍标准差, 可认为自相关系数具有拖尾特征。偏自相关系数图表明 1 阶、3 阶与 4 阶偏自相关系数显著大于 2 倍标准差, 也可认为偏自相关系数具有拖尾特征。综上所述, 可考虑拟合以下 ARIMA 模型:

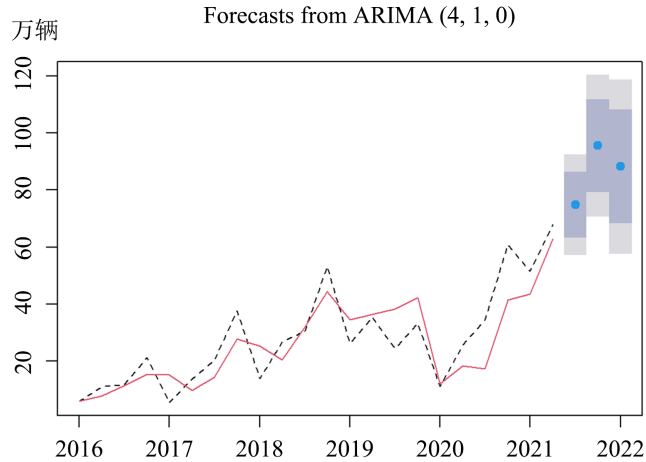


**Figure 12.** Autocorrelation and partial autocorrelation coefficients  
**图 12.** 自相关与偏自相关系数图

- 1) ARIMA (4, 1, 8);
- 2) ARIMA ((4), 1, 0)。

根据 AIC 信息准则, 进行模型优化, 最终选择 ARIMA ((4), 1, 0)模型进行新能源汽车季度销量的预测, 该 ARIMA 模型的估计结果及显著性检验结果详见附件。

新能源汽车季度销量的 ARIMA 预测模型对原始序列的拟合情况见图 13, 而对 2021 年第三季度销量数据的预测结果与实际值的比较见表 1。



**Figure 13.** Predictions of the quarterly sales of new energy vehicles by ARIMA ((4), 1, 0) model  
**图 13.** ARIMA ((4), 1, 0)模型对新能源汽车季度销量的拟合预测

**Table 1.** The predictions of the ARIMA quarterly sales forecast model  
**表 1.** ARIMA 季度销量预测模型的预测情况

时间	实际值	预测值	95%的置信区间	预测误差
2021 年第三季度	94.9	74.76	(57.16, 92.36)	20.14
2021 年第四季度	-	95.51	(70.62, 120.40)	-
2022 年第一季度	-	88.17	(57.69, 118.65)	-

## 6. 新能源汽车月度销量预测研究

### 6.1. 数据采集

本文使用 2016 年 1 月至 2021 年 10 月共 70 个月的新能源汽车月度销量数据进行预测模型的搭建(其中, 2016 年 1 月至 2021 年 6 月共 66 个月的数据用作训练集而 2021 年 7 月至 2021 年 10 月这四个月的数据用作测试), 数据来源于中国汽车工业协会。原始数据时序图如图 14 所示。

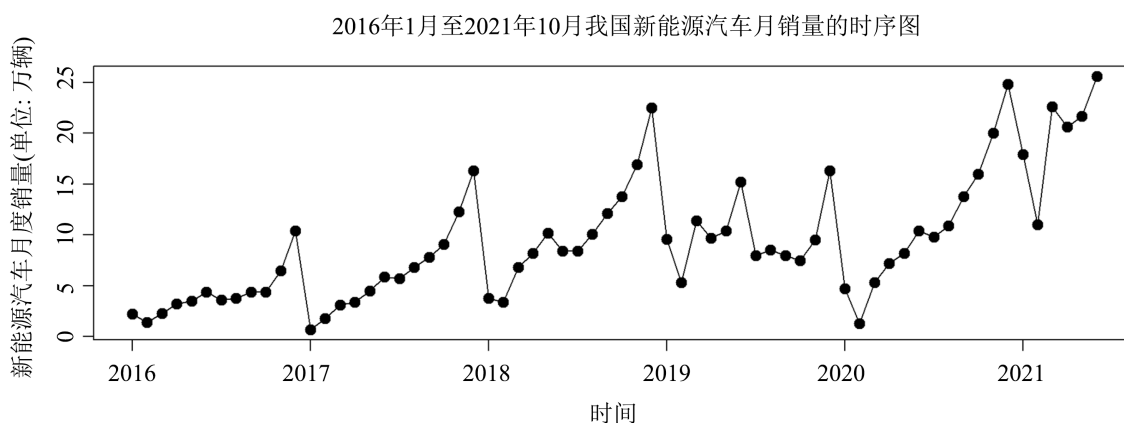


Figure 14. Time series diagram of the monthly sales of new energy vehicles in China

图 14. 我国新能源汽车月度销量序列时序图

### 6.2. 时间序列数据的预处理

原始序列的时序图显示该序列有着显著的上升趋势且蕴含固定的周期, 这是典型的非平稳序列特征。本文考虑用差分运算充分提取序列中的确定性信息, 原始序列差分处理后的时序图如图 15 所示。

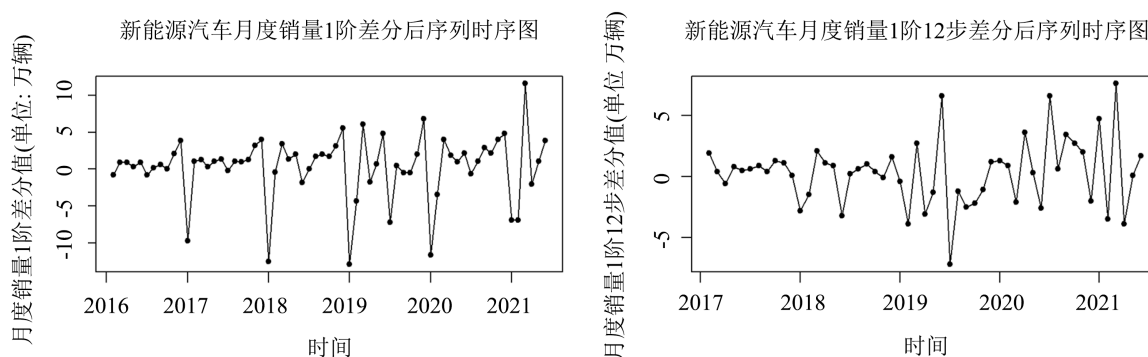


Figure 15. Sequence diagram of the monthly sales of Chinese new energy vehicles after difference operation

图 15. 我国新能源汽车月度销量差分处理后的时序图

1 阶差分有效提取了原序列中的递增趋势, 但仍残留季节变动和随机波动。对比 1 阶差分序列时序图, 1 阶 12 步差分很好地提取了周期信息。进一步地, ADF 检验结果表明 1 阶 12 步差分后的序列是平稳的, 见图 16。

纯随机性检验如图 17 所示, 纯随机性检验结果表明, 12 阶延迟的 LB 统计量的 P 值小于 0.05 的显著性水平, 故 1 阶 12 步差分后的序列为平稳非白噪声序列, 可用 ARMA 模型拟合该序列的发展。

**Augmented Dickey-Fuller Test**

alternative: stationary

Type 1: no drift no trend				Type 2: with drift no trend				Type 3: with drift and trend			
	lag	ADF	p. value		lag	ADF	p. value		lag	ADF	p. value
[1,]	0	-10.29	0.01	[1,]	0	-10.37	0.0100	[1,]	0	-10.44	0.010
[2,]	1	-5.45	0.01	[2,]	1	-5.53	0.0100	[2,]	1	-5.67	0.010
[3,]	2	-3.20	0.01	[3,]	2	-3.26	0.0235	[3,]	2	-3.35	0.073
[4,]	3	-2.81	0.01	[4,]	3	-2.88	0.0579	[4,]	3	-2.98	0.179

Note: in fact, p. value = 0.01 means p. value &lt;= 0.01

**Figure 16.** Results of ADF unit root test**图 16.** ADF 单位根检验结果**Box-Ljung test**

X-squared = 12.205, df = 6, p-value = 0.05756

X-squared = 29.253, df = 12, p-value = 0.003613

**Figure 17.** Results of pure randomness test**图 17.** 纯随机性检验结果

此外,考虑到异方差问题对模型拟合精度的影响,对上述1阶12步差分后的序列进行方差齐性检验。如图18所示,LM检验结果表明不存在异方差问题。

**ARCH LM-test; Null hypothesis: no ARCH effects**

Chi-squared = 3.3072, df = 1, p-value = 0.06898

Chi-squared = 3.1854, df = 2, p-value = 0.2034

Chi-squared = 3.1527, df = 3, p-value = 0.3687

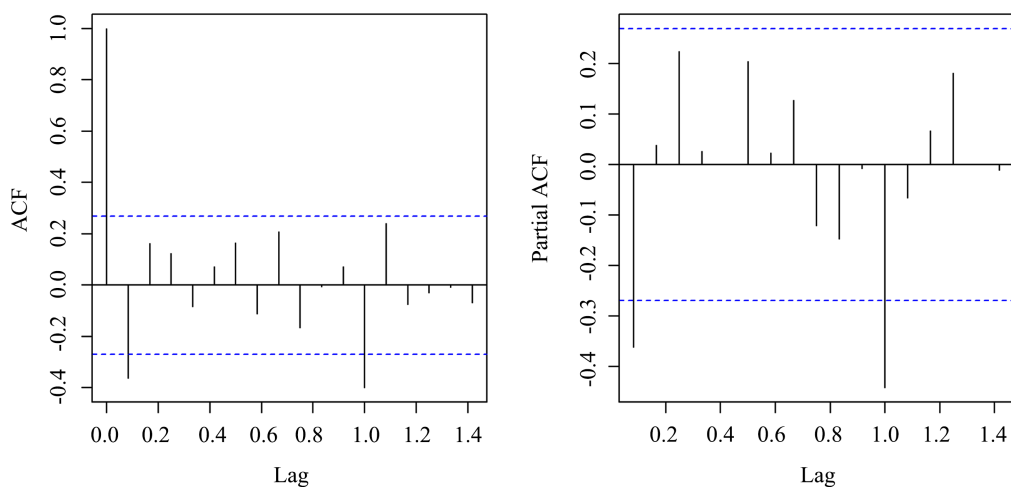
**Figure 18.** Results of the LM test for heteroscedasticity**图 18.** 异方差的 LM 检验结果**6.3. 平稳序列的拟合与预测**

1阶12步差分后序列的自相关与偏自相关图见图19。自相关系数图显示仅1阶和12阶的自相关系数显著大于2倍标准差。偏自相关系数图表明仅1阶与12阶偏自相关系数显著大于2倍标准差。可考虑拟合以下疏系数的季节加法模型:

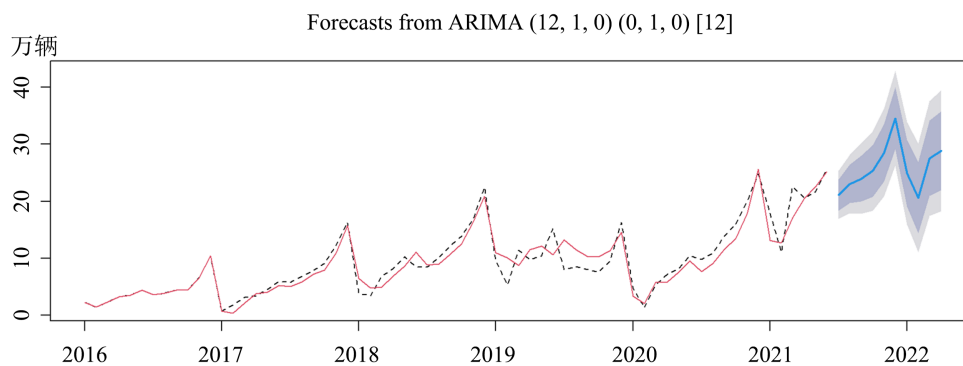
1) ARIMA ((1, 12), 1, (1, 12)) × (0, 1, 0)<sub>12</sub>;2) ARIMA (0, 1, (1, 12)) × (0, 1, 0)<sub>12</sub>;3) ARIMA ((1, 12), 1, 0) × (0, 1, 0)<sub>12</sub>。

根据AIC信息准则,进行模型优化,最终选择ARIMA ((1, 12), 1, 0) × (0, 1, 0)<sub>12</sub>模型进行新能源汽车月度销量的预测,该模型的估计结果及显著性检验结果详见附录。

新能源汽车月度销量的SARIMA预测模型对原始序列的拟合情况见图20,对2021年7月至10月的月销量数据的预测结果与实际值的比较见表2。



**Figure 19.** Autocorrelation and partial autocorrelation coefficients  
**图 19.** 自相关与偏自相关系数图



**Figure 20.** Predictions of the monthly sales of new energy vehicles by SARIMA model  
**图 20.** SARIMA 模型对新能源汽车月度销量的拟合预测

**Table 2.** The predictions of the SARIMA monthly sales forecast model  
**表 2.** SARIMA 月度销量预测模型的预测结果

时间	实际值	预测值	95%的置信区间	预测误差
2021 年 7 月	27.1	21.12	(16.91, 25.32)	5.98
2021 年 8 月	32.1	23.02	(17.86, 28.19)	9.08
2021 年 9 月	35.7	23.94	(17.79, 30.10)	11.76
2021 年 10 月	38.3	25.32	(18.36, 32.28)	12.98

此外, 本文的月度销量预测模型预测 2021 年 11 月的新能源汽车月销量为 28.53 万辆, 12 月预计销售 34.58 万辆, 而到 2022 年的 1 月, 销售量预计下降为 24.91 万辆, 2 月的销售量预计为 20.6 万辆, 3 月和 4 月的新能源汽车销售量分别为 27.55 万辆和 28.89 万辆。

## 7. 新能源汽车销量的长短期预测比较及分析

新能源汽车季度销量的 ARIMA 预测模型与新能源汽车月度销量的 SARIMA 预测模型对 2021 年第三和第四季度与 2022 年第一季度的销量的预测结果见表 3。

**Table 3.** Comparison of the results of different models in predicting the quarterly sales of new energy vehicles  
**表 3.** 不同模型对新能源汽车季度销量预测的结果比较

时间	季度销量 ARIMA 预测模型	月度销量 SARIMA 预测模型
2021 年第三季度	74.76	68.08
2021 年第四季度	95.51	88.43
2022 年第一季度	88.17	73.06

由表 3 中的结果可见, 对于新能源汽车季度销量的预测, 由月度销量预测数据转换得到的季度销量预测值相比于季度销量预测模型的直接预测结果更加低估了新能源汽车市场的发展。尽管如此, 本文构建的长短期预测模型都能很好地反映出新能源汽车销量序列的变化趋势。每年的 12 月都是创造该年月度销量记录的时候, 因为考虑到来年上市的新款车型将带来冲击, 商家在每年底都会给予消费者最大额度的购车优惠, 故大部分消费者都会选择在 12 月购入新车。

## 8. 结论与展望

本文对新能源汽车的整体销量进行了长期预测与短期预测。长短期预测模型均表明新能源汽车正以越来越快的速度成为汽车行业的主要角色。如今, 人工智能、云计算和大数据等关键前沿技术的发展为新能源汽车领域的诸多挑战与难题提供了有效的解决路径, 广大消费者们所关心的电池质量、车辆续航里程、充电效率、用车安全性、车辆智能化配置以及售后服务等问题得到逐步的解决, 我们有理由相信新能源汽车行业面临的发展机遇是前所未有的。

考虑到本研究在数据采集、模型方法及分析手段等方面存在的不足, 后续研究可从以下方面进行改进:

1) 考虑到汽车在设计与技术特性方面的不同而确定的用途分类: 乘用车与商用车, 后续研究可分别对新能源乘用车与新能源商用车的销量进行预测, 也可更进一步地比较两种预测方法(方法一: 基于新能源汽车整体的销量数据建立单一模型, 直接预测新能源汽车整体销量即本文的预测方法; 方法二: 由新能源乘用车的销量数据建立一个乘用车销量预测模型, 再由新能源商用车的销量数据建立一个商用车销量预测模型, 基于这两个模型的预测结果加总得到新能源汽车的整体销量预测)的效果差异。

2) 如今, 小鹏汽车、蔚来汽车及理想汽车等自主品牌已在新能源汽车领域取得了显著的成绩, 并不断地扩大各自的品牌影响力, 后续研究可关注特定品牌的新能源车型的产销发展状况。

## 参考文献

- [1] Nerlove, M. (1957) A Note on Long-Run Automobile Demand. *Journal of Marketing*, **22**, 57-64. <https://doi.org/10.1177/002224295702200109>
- [2] Chen, D. (2011) Chinese Automobile Demand Prediction Based on ARIMA Model. 2011 4th International Conference on Biomedical Engineering and Informatics (BMEI), Shanghai, 15-17 October 2011, 2197-2201. <https://doi.org/10.1109/BMEI.2011.6098744>
- [3] 胡彦君. ARIMA 模型在汽车销量预测中的应用及 SAS 实现[J]. 河北企业, 2012(4): 11-12.
- [4] 章旭. 基于时间序列分析的汽车销量预测研究[D]: [硕士学位论文]. 合肥: 合肥工业大学, 2017.
- [5] 周彦福, 王红蕾. 我国新能源汽车月度销售量预测模型研究[J]. 软件导刊, 2019, 18(8): 149-153.
- [6] 刘颖琦, 王萌, 王静宇. 中国新能源汽车市场预测研究[J]. 经济与管理研究, 2016, 37(4): 86-91.
- [7] Andrawis, R.R., Atiya, A.F. and El-Shishiny, H. (2011) Combination of Long Term and Short Term Forecasts, with Application to Tourism Demand Forecasting. *International Journal of Forecasting*, **27**, 870-886. <https://doi.org/10.1016/j.ijforecast.2010.05.019>

- [8] 刘媛. 新能源汽车销量的长期与短期预测[D]: [硕士学位论文]. 长春: 吉林大学, 2020.
- [9] 白一凡. 基于 SARIMA 和 BP 神经网络的新能源汽车销售预测[D]: [硕士学位论文]. 湘潭: 湘潭大学, 2019.
- [10] 王燕. 时间序列分析——基于 R [M]. 第 2 版. 北京: 中国人民大学出版社, 2020: 145-146.
- [11] 司守奎, 孙兆亮. 数学建模算法与应用[M]. 第 2 版. 北京: 国防工业出版社, 2015: 399-402.

## 附录

**Table A1.** Monthly sales data of Chinese new energy vehicles from January 2016 to October 2021 (Unit: 10,000)  
**表 A1.** 2016 年 1 月至 2021 年 10 月我国新能源汽车月度销量统计(单位: 万辆)

时间	销量	时间	销量
2016 年 1 月	2.2	2018 年 12 月	22.5
2016 年 2 月	1.4	2019 年 1 月	9.6
2016 年 3 月	2.3	2019 年 2 月	5.3
2016 年 4 月	3.2	2019 年 3 月	11.4
2016 年 5 月	3.5	2019 年 4 月	9.7
2016 年 6 月	4.4	2019 年 5 月	10.4
2016 年 7 月	3.6	2019 年 6 月	15.2
2016 年 8 月	3.8	2019 年 7 月	8
2016 年 9 月	4.4	2019 年 8 月	8.5
2016 年 10 月	4.4	2019 年 9 月	8
2016 年 11 月	6.5	2019 年 10 月	7.5
2016 年 12 月	10.4	2019 年 11 月	9.5
2017 年 1 月	0.7	2019 年 12 月	16.3
2017 年 2 月	1.8	2020 年 1 月	4.7
2017 年 3 月	3.1	2020 年 2 月	1.3
2017 年 4 月	3.4	2020 年 3 月	5.3
2017 年 5 月	4.5	2020 年 4 月	7.2
2017 年 6 月	5.9	2020 年 5 月	8.2
2017 年 7 月	5.7	2020 年 6 月	10.4
2017 年 8 月	6.8	2020 年 7 月	9.8
2017 年 9 月	7.8	2020 年 8 月	10.9
2017 年 10 月	9.1	2020 年 9 月	13.8
2017 年 11 月	12.3	2020 年 10 月	16
2017 年 12 月	16.3	2020 年 11 月	20
2018 年 1 月	3.8	2020 年 12 月	24.8
2018 年 2 月	3.4	2021 年 1 月	17.9
2018 年 3 月	6.8	2021 年 2 月	11
2018 年 4 月	8.2	2021 年 3 月	22.6
2018 年 5 月	10.2	2021 年 4 月	20.6
2018 年 6 月	8.4	2021 年 5 月	21.7
2018 年 7 月	8.4	2021 年 6 月	25.6
2018 年 8 月	10.1	2021 年 7 月	27.1

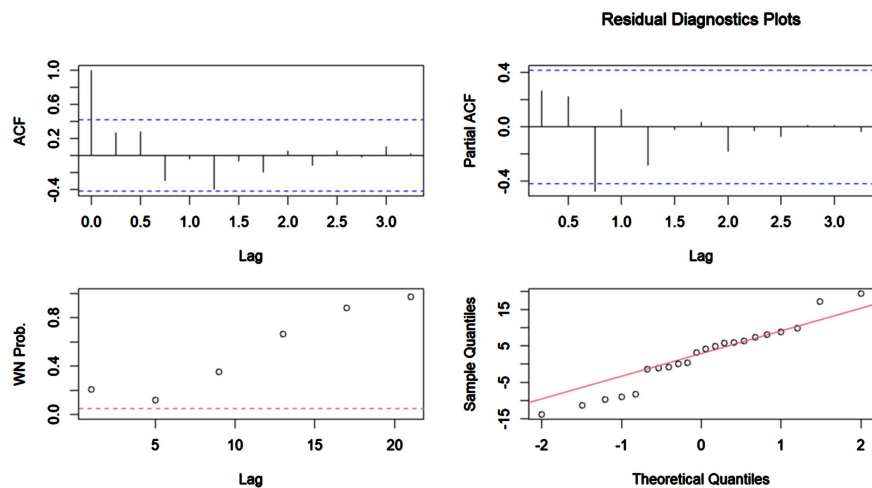


Continued

2018年9月	12.1	2021年8月	32.1
2018年10月	13.8	2021年9月	35.7
2018年11月	16.9	2021年10月	38.3

Coefficients:

	ar1	ar2	ar3	ar4
	0	0	0	0.7888
s. e.	0	0	0	0.1115

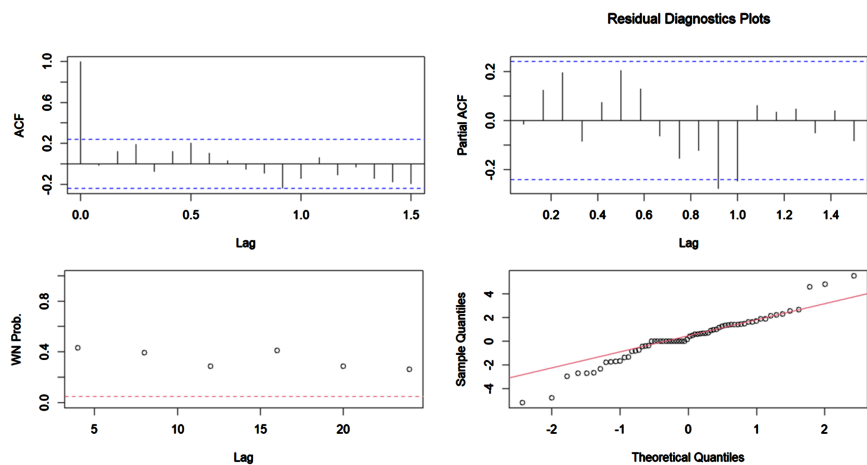
 $\sigma^2$  estimated as 80.63: log likelihood = -77.84, aic = 159.67


**Figure A1.** Results of the parameter estimation and model significance test of the ARIMA ((4), 1, 0) model

**图 A1.** ARIMA ((4), 1, 0)模型的参数估计结果与模型显著性检验结果

Coefficients:

	ar1	ar2	ar3	ar4	ar5	ar6	ar7	ar8	ar9	ar10	ar11	ar12
	-0.2868	0	0	0	0	0	0	0	0	0	0	-0.5144
s. e.	0.1022	0	0	0	0	0	0	0	0	0	0	0.1125

 $\sigma^2$  estimated as 4.604: log likelihood = -117.87, aic = 241.73


**Figure A2.** Results of the parameter estimation and model significance test of the ARIMA ((1, 12), 1, 0)  $\times$  (0, 1, 0) 12 model

**图 A2.** ARIMA ((1, 12), 1, 0)  $\times$  (0, 1, 0) 12 模型的参数估计结果与模型显著性检验结果