

基于时间序列ARIMA模型预测新能源汽车销量

陈友珍 钟楚熙

中央民族大学, 中国·北京 10080

【摘要】近年来,在新时代的绿色发展的背景下,国家对新能源汽车的发展投入了大量的资金和人力,促进了新能源汽车行业的发展。新能源汽车的销量持续保持良好的势头,新能源汽车出口也呈现出了较快的增长趋势。2023年新能源汽车的销量预计将超过2950万辆,新能源汽车行业整体呈现强势的增长趋势,于此同时,不同品牌的汽车也不断推出新能源汽车产品,新能源汽车行业内卷也越演越烈。为了研究目前新能源汽车市场上不同品牌的市场占有率,本文采用了时间序列分析方法对新能源汽车总销量进行了研究和预测。经研究发现,新能源汽车销量新能源汽车市场正处于快速增长的阶段,未来有望成为汽车市场的主流。随着技术的不断进步和成本的不断下降,新能源汽车的市场份额将会进一步扩大。

【关键词】新能源汽车;销量;时间序列ARIMA模型

Prediction of New Energy Vehicle Sales Based on Time Series ARIMA Model

Youzhen Chen, Chuxi Zhong

Central University for Nationalities Beijing 10080

[Abstract] In recent years, against the background of green development in the new era, the country has invested a large amount of funds and manpower in the development of new energy vehicles, promoting the development of the new energy vehicle industry. The sales of new energy vehicles continue to maintain a good momentum, and the export of new energy vehicles has also shown a rapid growth trend. The expected sales of new energy vehicles in 2023 are expected to exceed 29.5 million units, and the overall growth trend of the new energy vehicle industry is strong. At the same time, different brands of cars are constantly launching new energy vehicle products, and the competition in the new energy vehicle industry is becoming increasingly fierce. In order to study the market share of different brands in the current new energy vehicle market, this paper adopts time series analysis method to study and predict the total sales volume of new energy vehicles. Through research, it has been found that the sales of new energy vehicles are in a stage of rapid growth, and the new energy vehicle market is expected to become the mainstream of the automotive market in the future. With the continuous progress of technology and the continuous decrease in costs, the market share of new energy vehicles will further expand.

[Keywords] New energy vehicles; Sales volume; Time series ARIMA model

1 ARIMA模型构建

构建ARMA(p, q)模型,零均值的平稳时间序列{x_t}满足:

$$x_t = \phi_1 x_{t-1} + \dots + \phi_p x_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}$$

其

中

$$\phi_p \neq 0, \theta_q \neq 0, E(\varepsilon_t) = 0, D(\varepsilon_t) = \sigma_\varepsilon^2, E(\varepsilon_t \varepsilon_s) = 0, s \neq t, E(x_s \varepsilon_t) =$$

$$0, \forall s < t$$

。用延迟算子来描述该模型为 $\Phi(B)x_t = \Theta(B)\varepsilon_t$ 。

对非平稳序列进行适当的阶数差分运算,得到平稳的序列,再对其进行ARMA建模,可以得到ARIMA(p, d, q)模型:

$$\Phi(B)\nabla^d x_t = \Theta(B)\varepsilon_t$$

$$E(\varepsilon_t) = 0, D(\varepsilon_t) = \sigma^2_\varepsilon, E(\varepsilon_t \varepsilon_s) = 0, s \neq t$$

$$E x_s \varepsilon_t = 0, \forall s < t$$

其中 $\nabla^d = (1 - B)^d = \sum_{i=0}^d (-1)^i C_d^i B^i$;

$$\Phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$$

$$\Theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$$

ARIMA(p, d, q)可以进一步化简为

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

2 总销量时间序列模型检验

本文选取自2015年1月到2023年9月的新能源汽车总的月季销售量作为研究数据，建立时间序列分析ARIMA模型，预测接下来几个月里新能源汽车销售量的走势。时间序列ARIMA模型要求序列平稳性，因此要根据ADF的检验结果，判断是否可以进行时间序列分析。

ADF检验的结果，包括变量、差分阶数、T检验结果、AIC值等，用于检验时间序列是否平稳。原假设为该序列为一个不平稳的时间序列，是否拒绝原假设的标准如下式所示：

$$\begin{cases} P < 0.05 & \text{拒绝原假设} \\ P \geq 0.05 & \text{接受原假设} \end{cases}$$

对于检验原假设的统计值和ADF测试结果，我们通过比较它们与1%、5%和10%显著性水平下的临界值来进行评估。若ADF测试结果同时小于这三个显著性水平下的临界值，表示我们非常有效地拒绝了原假设。差分阶数的主要目的是平稳化时间序列数据，通过消除一些波动使其趋于平稳。对于非平稳序列，我们可以通过差分变换来转化为平稳序列。AIC值是一种衡量统计模型拟合优良性的标准，其数值越小越好。在比较不同模型时，我们通常选择具有最小AIC值的模型，以确保其在拟合数据方面的优越性。临界值是针对给定显著性水平的固定值。在我们的分析中，我们根据1%、5%和10%的显著性水平来确定相应的临界值，以便更准确地评估原假设的拒绝情况。

表：ADF检验表

变量	差分阶数	t	P	AIC	临界值		
					1%	5%	10%
新能源汽车总销售量	0	1.671	0.098	436.348	-3.548	-2.913	-2.594
	1	-1.103	0.014	195.122	-3.553	-2.915	-2.595
	2	-7.896	0.000***	69.442	-3.553	-2.915	-2.595

注：***、**、*分别代表1%、5%、10%的显著性水平

从ADF检验表来看，当差分阶数为2阶时，显著性P=0<0.5，且AIC值为69.44小于差分为1阶时的AIC值为195.122。所以差分阶数为2是最佳差分，水平上呈显著性，拒绝原假设，该序列为平稳时间序列，可以继续建模分析。做出2阶差分如下：

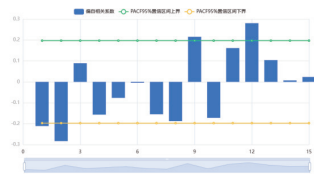


图：2阶差分序列图

由图，2015年到2018年新能源汽车销售量平稳，未发生较大波动，而2019年之后发生了较大的波动，这也符合了当时的社会背景，疫情时代经济波动变化较大，疫情结束之后经济回暖，新能源汽车销量也上升。自相关(ACF)图在q阶进行截尾，偏自相关(PACF)图拖尾，ARMA模型可简化为MA(q)模型；偏自相关(PACF)图在p阶进行截尾，自相关(ACF)图拖尾，ARMA模型可简化为AR(P)模型。倘若自相关与偏自相关图均拖尾，可结合PACF、ACF图中最显著的阶数(最小值)作为p、q值。结合以下偏自相关图和自相关图，最终确定p=1，q=2，ARMA(1,2)模型参数更优。



图：自相关图(ACF)



图：偏自相关图(PACF)

最终得到的模型如下：

表：模型参数表

ARIMA模型(1,1,2)检验表		
项	符号	值
	Df Residuals	95
样本数量	N	100
Q统计量	Q6(P值)	0.039(0.044)
	Q12(P值)	2.856(0.027)
	Q18(P值)	28.378(0.005***)
	Q24(P值)	31.322(0.026**)
	Q30(P值)	38.294(0.032**)
信息准则	AIC	48.613
	BIC	61.588
拟合优度	R ²	0.924

在本次模型检验中，我们考察了多个关键指标，包括样本数、自由度、Q统计量以及信息准则模型的拟合水平。ARIMA模型的关键要求之一是其残差应呈现无自相关性，也就是说，残差应当表现为白噪声。我们通过观察模型检验表中Q统计量的P值来判断残差是否符合白噪声特

性 (P值大于0.1表明残差为白噪声)。AIC和BIC值被用于多次比较模型,其数值越低表示模型的拟合效果越好。另外, R²用于衡量时间序列的拟合程度,数值越接近1代表效果越好。从模型的检验结果来看, Q统计量的P值均小于0.5,说明无需进一步进行白噪声检验,因为模型已经符合白噪声序列的条件。此外,模型的信息准则AIC和BIC的值分别为48.613和61.588,均相当小。拟合优度R²为0.924,非常接近1,表明模型在拟合时间序列方面表现出色。综合研究结果显示,该模型的性能出色,满足我们的需求,因此可以得出建模是成功的结论。

3 模型结果

基于变量新能源汽车总销售量,系统基于AIC信息准则自动寻找最优参数,模型结果为ARIMA模型(3,1,2)检验表且基于2差分数据,模型参数结果,包括模型的系数、标准差, T检验结果等,用于分析模型公式,根据模型参数表,最终确定模型公式为:

$$y(t) = 11734.632 - 0.407y_{t-1} + 0.386y_{t-2} + 0.099y_{t-3} + 0.207\varepsilon_{t-1}$$

表: 模型参数表

	系数	标准差	t	P> t	0.025	0.975
常数	11734.632	4082.7	2.874	0.004	3732.687	19736.577
ar.L1.D.新能源汽车总销售量	-0.407	0.214	-1.904	0.057	-0.826	0.012
ar.L2.D.新能源汽车总销售量	0.386	0.244	1.583	0.113	-0.092	0.863
ar.L3.D.新能源汽车总销售量	0.099	0.153	0.648	0.517	-0.2	0.398
ma.L1.D.新能源汽车总销售量	0.207	0.192	1.077	0.281	-0.169	0.583
ma.L2.D.新能源汽车总销售量	-0.752	0.167	-4.506	0	-1.078	-0.425

注: ***, **, *分别代表1%、5%、10%的显著性水平

时间序列预测和拟合图如下:

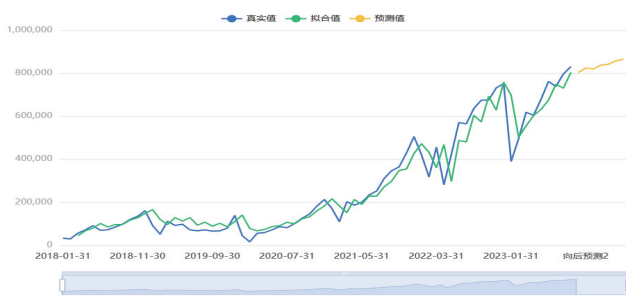


图: 时间序列图

4 预测未来五个月的销量

近年来国家大力支持新能源汽车的发展,出台了很优惠的政策以及制造鼓励。我国新能源汽车市场逐步提升,自2021年开始迅速增长,预测在2023年12月开始,新能源汽车月度销量将会突破80万。说明消费者对环保和节能意识的提高,以及政府对环保产业的支持力度增加。同时新能源汽车的技术逐步发展成熟,成本逐渐降低,新能源汽车的普及指日可待。人们可以根据新能源汽车市场的走向

和口碑选择自己心仪的新能源汽车。同时政府和企业还需继续努力,改善相关政策,加大投资,改进和完善技术,提高新能源汽车的电池续航、技能、安全性等。同时积极建设基础配套设施,如充电桩的建设。只有各方努力,全面建设新能源汽车的环境,新能源汽车的发展才会越来越好。

未来七个月的新能源汽车总销量预测如下:

时间	预测结果
2023.11	802924.2156682032
2023.12	823810.8353650767
2024.1	819058.4047370129
2024.2	837191.126931352
2024.3	840869.835218366
2024.4	856718.6309396633
2024.5	864305.5342074601

5 结语

新能源汽车市场的前景非常乐观。随着全球对环境保护和气候变化的关注不断增加,政府和企业对新能源汽车的支持和投资也在不断增加。以下是新能源汽车市场前景的一些方面:

政策支持: 许多国家和地区都出台了鼓励新能源汽车发展的政策,包括购车补贴、免费牌照、减少排放标准等。这些政策将进一步刺激新能源汽车的销售和推广。

技术进步: 随着新能源汽车的电动驱动、充电、电池等技术不断变好,新能源汽车的续航里程和性能将会不断提升,使得消费者更愿意购买新能源汽车。

成本下降: 随着新能源汽车的生产规模不断扩大,以及技术的进步,新能源汽车的成本将会逐渐下降,使得更多消费者能够接受和购买新能源汽车。

消费者意识: 随着环保意识的增强,越来越多的消费者愿意选择环保和节能的新能源汽车。

参考文献:

[1] 翟小伟, et al. 基于ARIMA模型的矿井灾害预测方法研究. 煤炭技术.

[2] 张思刘叶, and 金浩. 基于改进Ratio统计量的重尾AR(p)时间序列均值变点检验. 统计与决策.

[3] 叶光辉, 李松辉, 谢晟等. 新能源汽车产业开源科技情报多维视角分析[J/OL]. 情报科学.

[4] 郑爱琳, 蓝海林. 多重制度逻辑、地方政府响应行为与产业升级——以新能源汽车推广应用政策为例[J/OL]. 科学学研究.

作者简介:

陈友珍(1999--), 女, 彝族, 四川省凉山人, 中央民族大学, 研究方向: 信息与计算科学。

钟楚熙(2003--), 女, 汉族, 贵州省遵义市人, 中央民族大学, 研究方向: 信息与计算科学。