

# 基于SARIMA模型和Holt-Winters平滑法的山西省进出口总额预测

王瑞敏

上海工程技术大学, 上海

收稿日期: 2022年8月15日; 录用日期: 2022年9月9日; 发布日期: 2022年9月19日

## 摘要

从国家宏观层面看, 进出口总额体现了一国参与对外贸易的程度, 是国民经济的关键性指标之一。而在我国的各个省份内, 由于发展力度的差异, 不同省份对对外贸易的依赖程度也不同。作为经济欠发达地区的山西, 其对外贸易水平也在缓步增长。新冠疫情的到来, 对进出口贸易也产生了一定的冲击。因此, 分析山西省的进出口总额变化情况以及未来发展趋势尤为重要。本文以山西省2000~2021年进出口总额的264期月度数据为研究对象, 通过建立、统计检验和拟合乘法SARIMA模型对其进行时间序列分析, 并分析比较该模型和Holt-Winters指数模型的预测效果。研究结果表明, 山西省月度进出口总额在时间上具有趋势性和季节性, 并且模型拟合效果好。通过预测, 为进一步制定贸易政策和优化贸易结构提供了理论依据。

## 关键词

进出口总额, ARIMA模型, Holt-Winters指数模型

## Prediction of the Total Import and Export Volume of Shanxi Province Based on SARIMA Model and Holt-Winters Smoothing Method

Ruimin Wang

Shanghai University of Engineering Science, Shanghai

Received: Aug. 15<sup>th</sup>, 2022; accepted: Sep. 9<sup>th</sup>, 2022; published: Sep. 19<sup>th</sup>, 2022

## Abstract

At the national macro level, the total import and export value reflects a country's degree of participation in foreign trade and is one of the key indicators of the national economy. In each province of our country, due to the difference in development intensity, the degree of dependence of different provinces on foreign trade is also different. As an economically underdeveloped area, Shanxi's foreign trade level is also growing slowly. The arrival of the new crown epidemic has also had a certain impact on import and export trade. Therefore, it is particularly important to analyze the changes in the total import and export volume of Shanxi Province and the future development trend. This paper takes the monthly import and export data of Shanxi Province from 2000 to 2021 as the research object, conducts time series analysis by establishing, statistical testing and fitting the multiplicative SARIMA model, and analyzes and compares the prediction effect of the model and the Holt-Winters index model. The research results show that the monthly import and export volume of Shanxi Province has a trend and seasonality in time, and the model fitting effect is good. The forecast provides a theoretical basis for further formulating trade policy and optimizing trade structure.

## Keywords

Total Import and Export, ARIMA Model, Holt-Winters Exponential Model

Copyright © 2022 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

如今对外贸易对经济全球化起到了关键的作用。自改革开放以来,我国对外贸易得到了很大程度的发展。由于经济发展程度和经济结构的差异,我国不同省份对对外贸易的依赖程度也不同。作为我国中部省份,山西省虽然地域广阔、资源丰富,但对外贸易上有着不具备区位优势、贸易规模小、增速低和进出口产品单一等问题。在经过十几年的发展后,外贸结构逐渐优化,呈现出渐进式发展的趋势。近年来,“一带一路”给山西省的贸易转型打开了新的大门,该省致力于创新开放对外平台和贸易渠道,积极拓宽国际市场。

进出口总额数据体现着一国或一个省份对外贸易的客观发展规律,通过运用一定的数学方法和理论模型分析就可以掌握进出口额变化发展规律,将有助于对对外贸易的未来发展进行预测、规划与风险管理等活动,促进经济向好发展。本文以山西省 2000 至 2021 年进出口总额的 264 期月度数据为研究对象,通过建立、统计检验和拟合乘法 SARIMA 模型对其进行时间序列分析,并分析比较该模型和 Holt-Winters 指数模型的预测效果。最后得出结论,山西省月度进出口总额具有逐年递增的趋势和季节性,并且模型拟合效果好。通过预测,为进一步制定贸易政策和优化贸易结构提供了理论依据。

## 2. 文献综述

通过梳理文献发现,历年来运用不同研究方法分析预测对外贸易进出口总额的文献不在少数。王红(2006) [1]借助逐步回归理论,通过三次逐步回归拟合出最佳的进出口总额预测模型。李菊梅(2008) [2]建立 ARMA 模型拟合我国的年度进出口总额数据,得出结论该模型更适合进行短期预测。程兰芳(2010) [3]

进一步运用 ARIMA 模型拟合和预测我国的服务贸易进出口数据, 并得出了 95% 的概率意义下的短期区间预测效果。张玥(2021) [4] 利用随机森林和支持向量机两种算法来构建回归模型, 通过对 GDP、CPI、企业商品价格指数和社会消费品零售总额等八个变量进行回归来预测我国进出口总额, 结果表明支持向量机对于小样本数据的预测更优。

同时也有不少学者研究对外贸易的影响因素。谭亮(2010) [5] 采用协整理论证实对外直接投资与进口、出口存在短期替代效应和长期互补效应, 并建立了误差修正模型。胡登(2018) [6] 基于多元线性回归模型分析影响进出口总额的因素, 得出生产总值对进出口总额的重要作用以及财政收入于进出口总额的微弱反相关关系。俞力宁(2022) [7] 通过建立双对数线性回归模型得出地区国民生产总值和居民消费价格指数是影响对外贸易总额的重要因素。李岚(2022) [8] 基于格兰杰因果检验法发现汇率变动能抑制贸易、促进贸易出口, 并且汇率变动对进口额表现为单向的因果关系。

最后, 也有不少学者借助时间序列 ARIMA 模型进行不同领域的分析和预测。赵喜仓(2010) [9] 通过对不同的 SARIMA 模型进行参数估计和 AIC 准则比较后发现 ARIMA(2,1,1) (1,1,1)<sub>4</sub> 能较好地拟合我国季度 GDP 数据。戴玉泉(2019) [10] 基于 ARIMA(p,d,q) 模型对江西省 2017 年至 2019 年的 CPI 指数进行分析与预测, 表明 ARIMA(0,3,1) 模型预测效果较好, 给出了一些建议。强添刚等学者(2021) [11] 通过构建 SARIMA 模型和 GA-BP 神经网络的组合模型来预测地铁进站客流量, 表明复合预测模型在长期预测过程中的平均精度高于单个模型。

### 3. 理论模型

#### 3.1. 两个基本模型 AR(p) 和 MA(q) 模型

$p$  阶自回归模型的结构如下, 简记为 AR(p):

$$X_t = \alpha_0 + \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \cdots + \alpha_p X_{t-p} + \varepsilon_t$$

三个限制条件如下:

$$\alpha_p \neq 0$$

$$E(\varepsilon_t) = 0, \text{Var}(\varepsilon_t) = \sigma_\varepsilon^2, E(\varepsilon_t \varepsilon_s) = 0, \forall s \neq t$$

$$E(x_t \varepsilon_s) = 0, s < t$$

$q$  阶移动平均模型的结构如下, 简记为 MA(q):

$$X_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q}$$

两个限制条件如下:

$$\theta_q \neq 0$$

$$E(\varepsilon_t) = 0, \text{Var}(\varepsilon_t) = \sigma_\varepsilon^2, E(\varepsilon_t \varepsilon_s) = 0, s \neq t$$

#### 3.2. SARIMA 模型

ARIMA(p,d,q) 模型可以看作差分运算和 AR(p)、MA(q) 的组合, 其中  $d$  是指对数据进行差分的阶数, 这表明经过适当阶数的差分后时间序列会变得平稳。针对无季节效应的非平稳序列时, 使用 ARIMA 模型拟合数据的效果较好。但当时间序列中出现明显的季节性变化时, ARIMA 模型将不再能很好的拟合数据。因此需要运用 SARIMA 模型较好的预测有季节效应的非平稳时间序列。SARIMA 模型分为季节乘法 and 季节相加两种, 能够提取序列的长期趋势、季节效应和随机波动之间的相互关系。

季节性时间序列模型  $ARIMA(p,d,q)(P,D,Q)_S$  表示为:

$$\Phi_p(B)\Phi_p(B^S)\nabla^d\nabla_S^D x_t = \Theta_q(B)\Theta_q(B^S)\varepsilon_t$$

其中:  $S$  是周期长度,  $P$ 、 $Q$  均为季节性的阶数。  $\Phi_p(B)$  表示  $p$  阶自回归多项式,  $\Theta_q(B)$  表示  $q$  阶移动平均多项式。  $\nabla^d = (1-B)^d$ ,  $\nabla_S^D = (1-B^S)^D$ 。  $\Phi_p(B^S)$  和  $\Theta_q(B^S)$  表示季节性的系数多项式。

### 3.3. SARIMA 模型的建模流程

第一, 对序列绘制时序图, 直观观察序列是否平稳。第二, 对于存在季节性的时间序列数据, 需要提取季节趋势, 确定周期长度  $S$ 。通过恰当阶数的差分, 使序列从非平稳转变为平稳, 并进行差分后 ADF 平稳性检验和白噪声检验。第三, 结合自相关图和偏自相关图, 基于短期相关性得到阶数  $p$ 、 $q$ , 基于季节自相关特征得到阶数  $P$ 、 $Q$ 。第四, 将可能的  $p$ 、 $q$ 、 $d$  和  $P$ 、 $Q$ 、 $D$  的取值代入 SARIMA 模型, 进行模型显著性检验。第五, 基于拟合模型对序列进行预测。

## 4. 基于 SARIMA 模型的时间序列分析

### 4.1. 数据预处理

本文从中经网统计数据库收集了 264 期从 2000 年 1 月份至 2021 年 12 月份的山西省进出口总额月度数据, 单位为千万美元。本文所使用的数据处理分析软件为 Rstudio。首先对序列数据绘制时序图, 见图 1。山西省的进出口总额呈现逐年稳步递增的趋势, 2021 年增长幅度最大, 表明“十四五”对外贸易规划取得良好开端以及“一带一路”在该省发挥的积极成效。

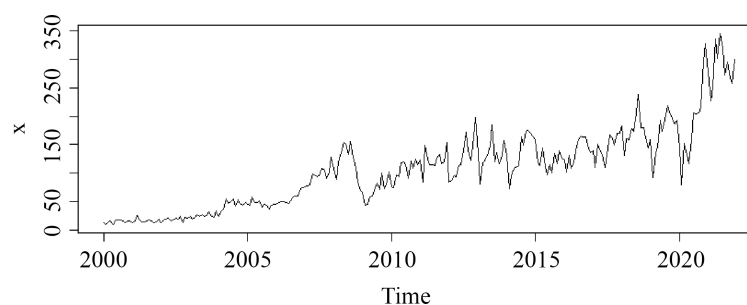


Figure 1. Timing diagram

图 1. 时序图

通过对其进行进一步的因素分解, 可以看出这个序列受到三个因素的影响: 长期趋势(trend)、季节效应(seasonal)和随机波动(random), 说明山西省进出口总额序列是一个季节性的非平稳时间序列, 见图 2。

### 4.2. 构建 SARIMA 模型

从因素分解图看出, 山西省进出口总额数据表现出明显的以年度为周期的特征。因此, 对进出口总额时间序列进行 1 阶差分提取趋势效应, 12 步差分提取季节效应, 得到 1 阶 12 步差分后的序列时序图, 见图 3。差分后数据围绕一定数值上下波动, 序列已经没有了明显的递增趋势和以年为周期的波动。

#### 4.2.1. 平稳性检验和纯随机性检验

##### 1) ADF 检验

原假设: 序列非平稳。

Decomposition of multiplicative time series

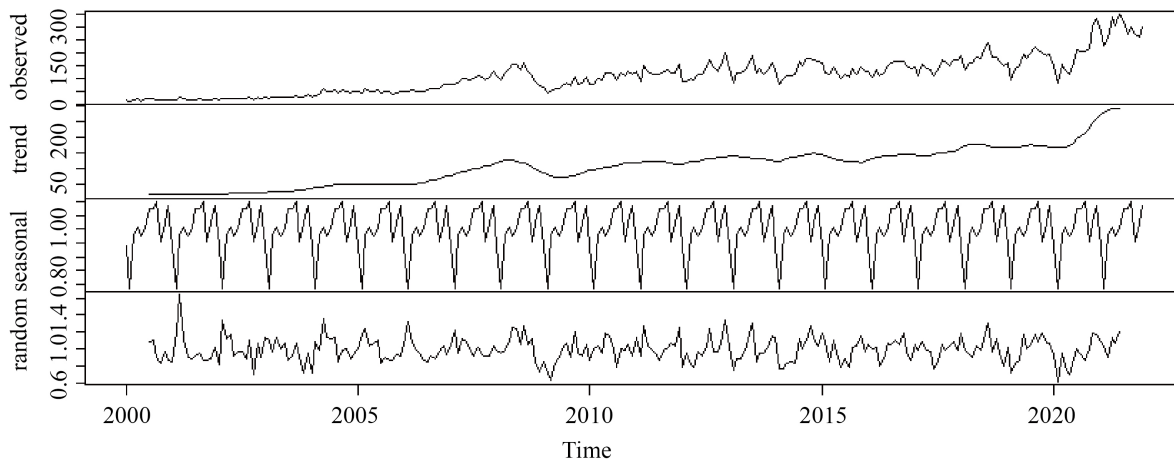


Figure 2. Factorization diagram  
图 2. 因素分解图

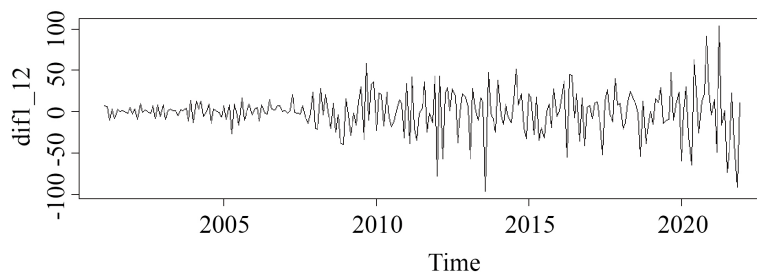


Figure 3. Sequence diagram after 1st-order 12-step difference  
图 3. 1 阶 12 步差分后序列图

备择假设：序列平稳。假设条件用参数  $\rho$  表达，即为

$$H_0 : \rho \geq 0, H_1 : \rho < 0$$

差分后序列的 ADF 统计量的 P 值为 0.01，小于显著性水平 0.05，因此拒绝原假设，认为经过 1 阶 12 步差分后的山西省进出口总额序列平稳。

2) 纯随机性检验

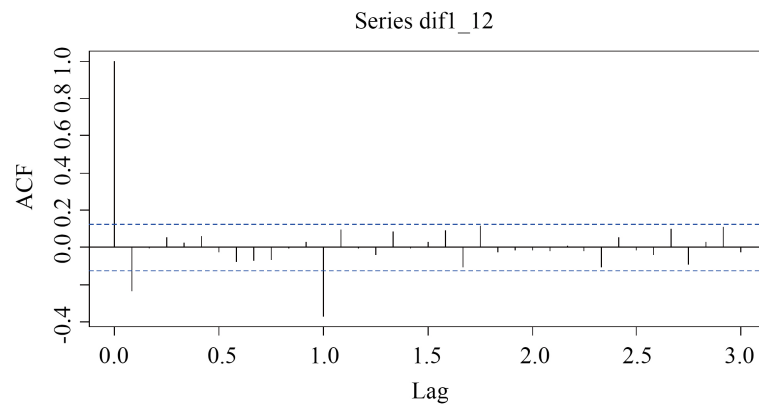
LB 统计量的 P 值远小于显著性水平 0.05，所以差分后的山西省进出口总额序列不是白噪声序列。

通过上述检验，可以确定差分后的山西省进出口总额序列为平稳非白噪声序列，进而建立 SARIMA 模型。

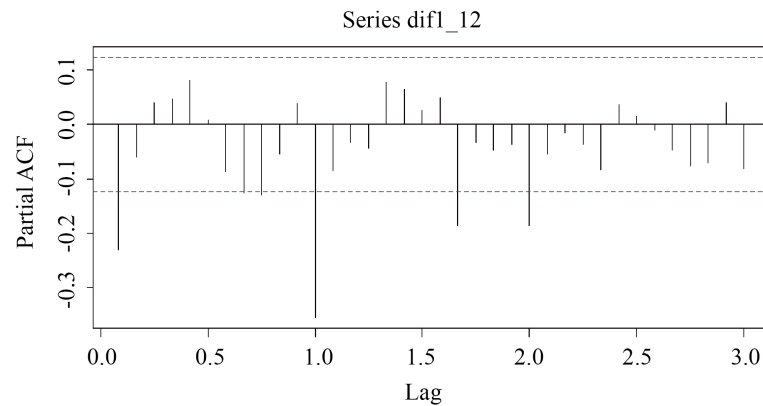
4.2.2. SARIMA 模型定阶

因为对进出口总额序列进行了 1 阶 12 步的差分，所以  $d, D$  取值都为 1。见图 4 和图 5，通过 ACF 图和 PACF 图可以看出，1 阶自相关系数落在 2 倍标准差外，偏自相关系数拖尾，所以  $p = 0, q = 1$ 。接着以 12 阶为一个周期分析季节趋势，12 阶(1 周期)自相关系数大于 2 倍标准差，其余系数接近零数值。延迟 1 个周期、2 个周期的偏自相关系数都大于 2 倍标准差，考虑拖尾。所以  $P = 0, Q = 1, S = 12$ 。

经过上述定阶，建立季节乘法模型为 SARIMA(0,1,1)(0,1,1)<sub>12</sub>。



**Figure 4.** ACF plot after differencing  
**图 4.** 差分后 ACF 图



**Figure 5.** PACF plot after differencing  
**图 5.** 差分后 PACF 图

### 4.3. SARIMA 模型检验

对拟合的模型进行显著性检验。见表 1，参数的估计值均大于 2 倍标准差，所以参数估计值有效。见图 6，残差序列的检验统计量的 P 值都显著大于 0.05，右下图显示近似服从正态分布，说明趋势效应和季节效应提取充分后的残差序列为白噪声序列，拟合的模型 SARIMA(0,1,1)(0,1,1)<sub>12</sub> 显著有效。

### 4.4. 模型拟合口径以及预测

使用条件最小二乘法与极大似然混合估计方法，拟合 SARIMA(0,1,1)(0,1,1)<sub>12</sub> 模型的表达式：

$$\nabla \nabla_{12} x_t = (1 - \theta_1 B)(1 - \theta_{12} B^{12}) \varepsilon_t$$

**Table 1.** Parametric significance test  
**表 1.** 参数显著性检验

	ma1	sma1
估计值	-0.3136	-0.8016
标准差 s.e.	0.0617	0.0468

Residual Diagnostics Plots

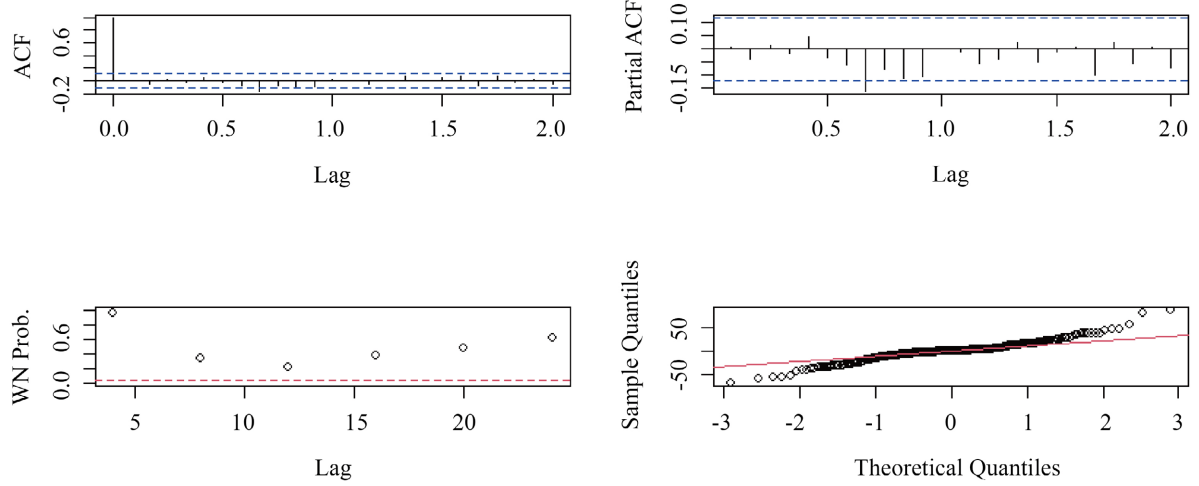


Figure 6. Model significance test

图 6. 模型显著性检验

代入参数估计值，得出：

$$\nabla \nabla_{12} x_t = (1 - 0.3136B)(1 - 0.8016B^{12})\varepsilon_t$$
$$\text{Var}(\varepsilon_t) = 417.3$$

通过拟合的 SARIMA(0,1,1)(0,1,1)<sub>12</sub>，对 2022 年的月度山西省进出口总额数据进行预测。由图 7 可知，拟合效果良好。

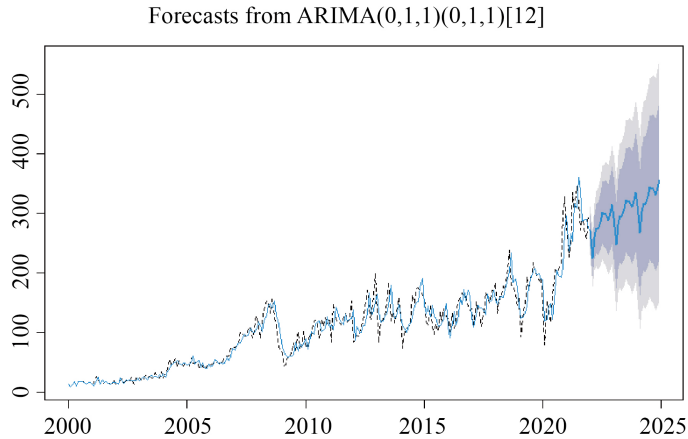


Figure 7. SARIMA model prediction effect chart

图 7. SARIMA 模型预测效果图

### 5. Holt-Winters 三参数指数平滑法

除了运用 SARIMA 模型预测带有季节效应的时间序列外，还可以采用 Holt-Winters 指数法。R 基于最优拟合原则计算出平滑系数：

$$\alpha = 0.66015, \quad \beta = 0, \quad \gamma = 0.4011472$$

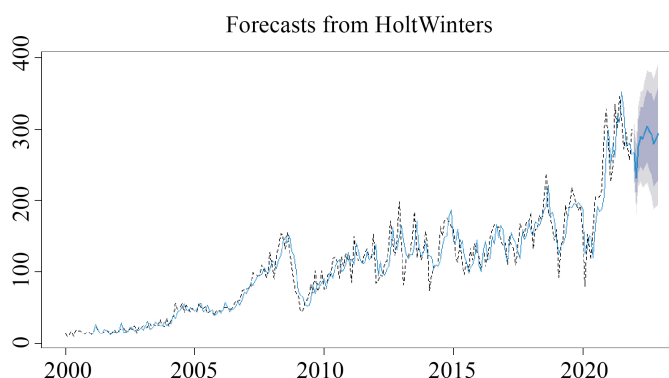
通过 Holt-Winters 三参数指数平滑加法迭代公式，得到三参数的最后迭代值为：

$$a(t) = 278.2183621, \quad b(t) = 0.1299257$$

所以，该序列的向前任意  $k$  期的预测值等于：

$$x_{t+k} = 278.22 + 0.13k + S_j, \quad \forall k \geq 1$$

由 Holt-Winters 三参数指数平滑预测效果图看出，拟合效果较好，见图 8。



**Figure 8.** Holt-Winters three-parameter exponential smoothing  
**图 8.** Holt-Winters 三参数指数平滑法

## 两种模型预测结果对比

采用两种模型对山西省 2022 年下半年的进出口总额进行预测，并与上半年实际进出口总额对比，结果如下，见表 2：

**Table 2.** System resulting data of standard experiment  
**表 2.** 标准试验系统结果数据

月份	实际值 (千万美元)	SARIMA(0,1,1) × (0,1,1) <sub>12</sub> 预测值	误差/%	Holt-Winters 三参 数指数模型预测值	误差/%
2022-1	253.63	271.57	7.07%	266.85	5.21%
2022-2	191.31	225.31	17.77%	230.28	20.37%
2022-3	268.20	260.13	-0.1%	275.9	2.87%
2022-4	239.67	274.49	14.53%	289.56	20.82%
2022-5	203.10	273.17	34.50%	285.83	40.74%
2022-6	260.07	285.70	9.85%	295.27	13.53%
2022-7	—	301.43	—	304.07	—
2022-8	—	298.50	—	296.66	—
2022-9	—	298.65	—	292.70	—
2022-10	—	288.91	—	279.12	—
2022-11	—	300.15	—	286.40	—
2022-12	—	314.04	—	293.68	—



通过对比发现, SARIMA 模型的预测效果与 Holt-Winters 三参数指数平滑法的预测效果大体一致, 但与实际值相比, 有些年份的预测值会有些偏离, 可能受到当月宏观环境的影响。

SARIMA 模型和 Holt-Winters 三参数指数平滑法均可以对山西省的进出口总额进行拟合和预测。对于 SARIMA 模型来说, 更适合拟合和预测线性的时间序列, 不需要考虑外界的因素, 通过序列内部的关系进行自回归, 找出合适的  $p$  和  $q$  值, 就可以灵活的拟合数据。而 Holt-Winters 三参数指数平滑法利用三个平滑系数来有效预测未来的趋势, 但却难以找到最佳的平滑系数值, 且只适合预测 2~3 期的数据。由于山西省进出口总额还会受到外界宏观因素的影响而出现某一期的跳跃, 就会导致偏差变大。因此, 通过调整  $p$  和  $q$  的取值, SARIMA 模型能更好地拟合序列。

## 6. 研究结论

本文使用 SARIMA(0,1,1)(0,1,1)<sub>12</sub> 模型拟合山西省 2000 年至 2021 年的进出口总额数据, 结果显示山西省的进出口总额呈现出稳步增长的长期趋势和周期性波动。同时使用该模型和 Holt-Winters 三参数指数平滑法分别预测山西省 2022 年下半年的进出口总额, 结果显示两种方法的拟合效果大体一致, 在未来半年将会保持稳步增长态势。但由于受到宏观环境的影响, 会与实际值之间存在误差。

通过序列趋势可知, 2021 年山西省的进出口总额相较 2020 年有了很大的提升, “十四五”对外贸易规划取得了良好的开端, 同时“一带一路”在该省发挥的成效显现。2021 年, 山西省外贸企业数量增加, 进出口业务达历史新高, 民营企业活力增强。从贸易结构看, 制造业贸易和一般贸易进出口额明显增加, 外商投资企业发挥主导作用。同年, 全省增加了对“一带一路”沿线国家的进出口业务。

2022 年进出口贸易总额的实际值略低于预测值, 是因为疫情的反复对经济增长有一定的抑制作用。据太原海关统计, 2022 年上半年全省货物贸易同比下降了 18.8%。面对该省进出口贸易总额的下降, 为应对疫情对贸易的冲击, 山西省金融机构通过加大信贷投放、降低汇率风险和创新跨境金融服务等措施保障外贸行业的资金需求, 促进该省外贸进出口业务的恢复。下半年预测值保持稳定, 意味着该省需要保持现有的外贸状态, 并不断优化外贸结构, 充分挖掘外贸企业的潜力和地域资源, 实现新的突破。因此, 一方面山西省要积极推进“一带一路”建设, 拓宽贸易渠道, 创新对外贸易产品, 提升产品核心技术。另一方面要寻求发展服务型贸易, 提高服务贸易占比, 将产业升级、服务型贸易与数字化融合, 推动高质量发展。最后, 山西省应该继续聚焦于“十四五”外贸目标, 培育新型外贸主体, 对现有外贸业务转型升级, 同时提升对外开放平台功能、积极开拓国际市场, 进一步推动外贸经济的转型发展。

## 参考文献

- [1] 王红, 童恒庆, 魏平. 进出口贸易额预测的逐步回归建模研究[J]. 统计与决策, 2006(2): 80-82.
- [2] 李菊梅, 樊亮, 崔俊峰. 用时间序列分析方法预测我国的进出口总额[J]. 重庆工学院学报(自然科学版), 2008, 22(12): 136-138.
- [3] 程兰芳, 陆敏. 中国服务贸易进出口规模的 ARIMA 模型和预测[J]. 黑龙江对外经贸, 2010(12): 25-27.
- [4] 张玥, 杜宝军. 基于机器学习对我国进出口总额的研究[J]. 质量与市场, 2021(12): 150-152.
- [5] 谭亮, 万丽娟. 中国对外直接投资与进出口贸易关系的实证分析[J]. 重庆大学学报(社会科学版), 2010, 16(1): 59-64.
- [6] 胡登. 陕西进出口总额影响因素研究——基于多元线性回归模型[J]. 当代经济, 2018(14): 88-89.
- [7] 俞力宁. 宏观经济因素对进出口贸易总额影响的研究——基于江苏省的实证分析[J]. 现代商业, 2022(9): 36-40. <https://doi.org/10.14097/j.cnki.5392/2022.09.016>
- [8] 李岚. 人民币汇率变动对我国进出口贸易影响的实证分析[J]. 现代营销(下旬刊), 2022(5): 4-7. <https://doi.org/10.19932/j.cnki.22-1256/F.2022.05.004>
- [9] 赵喜仓, 周作杰. 基于 SARIMA 模型的我国季度 GDP 时间序列分析与预测[J]. 统计与决策, 2010(22): 18-20.

---

<https://doi.org/10.13546/j.cnki.tjyj.2010.22.023>

- [10] 戴玉泉. 基于 ARIMA 模型对江西省 CPI 的时间序列分析与预测[J]. 科学技术创新, 2019(35): 18-20.
- [11] 强添纲, 刘涛, 裴玉龙. 地铁进站客流量 SARIMA 与 GA-BP 神经网络组合预测模型[J]. 铁道运输与经济, 2021, 43(12): 134-142. <https://doi.org/10.16668/j.cnki.issn.1003-1421.2021.12.20>