

# 基于模型平均的青岛市空气质量分析

苗佳豪

青岛大学数学与统计学院, 山东 青岛

收稿日期: 2026年3月23日; 录用日期: 2026年4月18日; 发布日期: 2026年4月24日

## 摘要

近年来, 人民生活水平提高, 人民对美好生活的向往从过去的有没有转变成了好不好, 在物质文明和精神文明相协调的过程中, 空气质量问题成为众多学者关注的焦点。针对当前研究中存在着空气质量影响因素复杂并且传统单一的模型预测不稳定的情况, 本文以青岛市为研究对象, 基于模型平均方法对空气质量指数进行分析和预测, 主要工作如下: (1) 首先, 利用python爬虫获取2018~2023年青岛市空气质量指数AQI以及大气污染物浓度(PM2.5、PM10、CO、NO<sub>2</sub>、O<sub>3</sub>)数据, 并对原始数据进行标准化处理和描述性分析, 了解当前空气质量变化情况, 揭示青岛市空气质量的年际变化趋势和季节性特征。(2) 基于变量相关性分析的基础上, 构建多种回归模型, 并且利用AIC准则筛选代表性子模型, 并且引入四种模型平均方法, 如平滑AIC (S-AIC)、平滑BIC (S-BIC)、Mallows模型平均(MMA)和Jackknife模型平均(JMA), 通过赋予不同权重确定策略对多个回归模型进行融合, 并利用真实数据进行预测和比较, 采取绝对误差、均方误差等指标对模型性能进行评估。并且和传统机器学习模型进行比较, 以验证本次研究中所提出方法预测的有效性。本文研究结果表明, 近年来, 随着政府推出一系列绿色政策, 积极响应生态文明建设, 青岛市空气质量整体呈改善趋势。在预测性能方面, S-AIC和S-BIC模型平均方法在短期预测中具有较高精度和稳定性, 而MMA和JMA模型平均方法在长期预测中表现出更好的稳定性和适应性。综合研究表明, 模型平均方法能够有效降低单一模型预测过程中存在的不确定性, 同时, 本文的研究结果也为后续城市环境管理和空气质量污染防控提供了科学依据。

## 关键词

空气质量预测, 模型平均, S-AIC, S-BIC, MMA, JMA

# Air Quality Analysis of Qingdao Based on Model Averaging

Jiahao Miao

School of Mathematics and Statistics, Qingdao University, Qingdao Shandong

Received: March 23, 2026; accepted: April 18, 2026; published: April 24, 2026

## Abstract

In recent years, with the improvement of people's living standards, the pursuit of a better life has shifted from "whether it exists" to "whether it is of high quality". In the process of coordinating material and spiritual civilization, air quality has become a major concern among researchers. Considering the complexity of influencing factors and the instability of traditional single-model predictions, this study takes Qingdao as the research object and analyzes and predicts the Air Quality Index (AQI) based on model averaging methods. The main contributions are as follows: (1) First, Python-based web crawling techniques are employed to collect AQI data and atmospheric pollutant concentrations (PM<sub>2.5</sub>, PM<sub>10</sub>, CO, NO<sub>2</sub>, and O<sub>3</sub>) in Qingdao from 2018 to 2023. The raw data are standardized and subjected to descriptive statistical analysis to explore the overall air quality conditions, as well as the interannual trends and seasonal characteristics. (2) Based on correlation analysis of variables, multiple regression models are constructed. Representative sub-models are selected using the Akaike Information Criterion (AIC). Furthermore, four model averaging approaches, including smoothed AIC (S-AIC), smoothed BIC (S-BIC), Mallows Model Averaging (MMA), and Jackknife Model Averaging (JMA), are introduced. Different weighting strategies are applied to combine multiple regression models. The predictive performance is evaluated using real data with metrics such as mean absolute error (MAE) and mean squared error (MSE). And compare it with traditional machine learning models to verify the effectiveness of the method proposed in this study. The results show that, in recent years, with the implementation of a series of green policies and the promotion of ecological civilization, the air quality in Qingdao has shown an overall improving trend. In terms of predictive performance, S-AIC and S-BIC exhibit higher accuracy and stability in short-term forecasting, while MMA and JMA demonstrate better stability and adaptability in long-term forecasting. Overall, model averaging methods effectively reduce the uncertainty associated with single-model predictions. The findings of this study provide a scientific basis for urban environmental management and air pollution control in the future.

## Keywords

Air Quality Prediction, Model Averaging, S-AIC, S-BIC, MMA, JMA

Copyright © 2026 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

### 1.1. 研究背景

随着工业化发展进程加快, 环境污染越来越严重, 众多学者对空气质量做出研究。宋艳玲等通过分析北京市大气污染物二氧化硫以及其他细微颗粒物发现: 细微颗粒物是北京市最主要的污染物, 且污染集中于冬季[1]。刘丽等通过对青岛市6种污染物进行分析, 得出青岛市夏季空气质量优于冬季的结论[2]。武苗苗通过对区域进行划分, 研究山东省各区域污染物时空分布特征[3]。徐红霞等通过进行描述性统计和绘制污染物分布图, 对山东省各市的空气质量进行分析, 得出颗粒污染物德州最高的结论[4]。

空气中的污染物对人类健康有害, 如颗粒物、臭氧、二氧化硫等被证实与呼吸系统疾病、心血管疾病等健康问题相关联。此外, 空气污染还会影响生态系统的平衡, 损害植被、水体和土壤等环境要素。许多国家和地区都对空气质量制定了严格的监管要求和政策措施, 要求相关机构对空气质量进行监测和

分析,并根据分析结果制定相应的控制措施和政策。因此,空气质量分析成为了政府部门、环保组织和企业等重要的工作内容。随着科学技术的不断发展和进步,空气质量分析的技术和方法也在不断创新和完善,从传统的监测方法到基于模型的预测分析,空气质量分析工具和手段不断更新,为更准确、更全面地了解空气质量提供了新的途径和思路。

传统用来衡量空气质量的指标为空气污染指数(API),2012年上半年出台规定,将用空气质量指数(AQI)替代原有的空气污染指数(API)。AQI与原来发布的空气污染指数(API)有着很大的区别。AQI分级计算参考的标准是新的环境空气质量标准(GB3095-2012),参与评价的污染物分别为SO<sub>2</sub>、NO<sub>2</sub>、PM<sub>10</sub>、PM<sub>2.5</sub>、O<sub>3</sub>、CO;而API分级计算参考的标准是老的环境空气质量标准(GB3095-1996),评价的污染物仅包括SO<sub>2</sub>、NO<sub>2</sub>和PM<sub>10</sub>三项,且AQI采用分级限制标准更严。因此AQI较API监测的污染物指标更多,其评价结果更加客观。本次研究将以AQI及其相关指标作为研究对象。

当前对于空气质量分析的研究虽然很多,但是由于空气质量影响因素的多变性,大部分研究都是基于机器学习进行的单独讨论某个变量对空气质量的影响,或者以单一模型来对空气质量情况进行预测,这就需要对变量和模型进行选择。但是对于模型选择方法来说,我们只能选取待选模型之一来进行研究,同时我们并不能确定待选模型中一定包含正确的模型,因此在研究的过程中会存在很多不确定性。另外,在模型选择的过程中,研究者们往往会忽略模型存在的问题,大大降低预测的精度。因此学者开始尝试利用模型组合或者模型平均的方法去解决这些问题。

## 1.2. 国内外研究现状

模型平均方法起源于20世纪60年代,Bates和Granger将两种无偏的预测组合起来作为最终的预测结果,展示了组合预测的优势。近年来,对于模型平均的研究主要包括两大主流方向,其一为贝叶斯模型平均,其二为频率模型平均。

在早期贝叶斯模型平均的发展过程中,Min & Zellner, A.提出了贝叶斯模型平均方法,它的核心思想是通过每个候选模型及模型中的参数设置一个先验概率,然后用经典贝叶斯方法进行统计推断,计算不同候选模型的后验概率,将其作为权重进行加权组合[5]。随着研究的深入,研究者们发现贝叶斯模型平均在实际应用中存在许多缺陷,首先,不同的先验概率对模型预测的结果影响较大,其次,贝叶斯模型平均要求真实模型包含于被考虑的模型中,但是现实中真实模型往往要比一般模型复杂得多。对于频率模型平均来说,它更加侧重于基于数据模型给出合理的组合权重的选择。

考虑到贝叶斯模型平均存在这些问题,频率模型平均也得到了研究者更多的关注。Buckland等在AIC与BIC模型选择准则的基础上提出了平滑AIC(S-AIC)与平滑BIC(S-BIC),通过给候选模型赋予不同的权重,进而得到最终的结果[6]。S-AIC与S-BIC由于使用简单,是目前应用最广泛的方法;Hansen提出了Mallows模型平均方法(MMA),通过极小化Mallows准则来选择权重组合各个模型参数的最小二乘估计量,这一准则是关于权重的二次函数,因此在计算时是二次规划的一个简单的应用[7]。同时,Hansen在2008年将MMA方法应用于时间序列数据,进一步验证了该方法的准确性和有效性。Zhang和Zou证明了MMA估计是渐进最优的,即当把权重限定在一定集合中时,MMA估计的二次损失依概率渐进等价于该集合中权重加权的模型的平均估计的二次损失的下确界[8]。但MMA方法在应用时要设定同方差的前提并假定待组合的模型是嵌套的,也就是说,第 $k+1$ 个模型中必须要包含第 $k$ 个模型中的所有解释变量,即MMA方法依赖于解释变量的次序,这在实际应用中是不现实的;基于MMA方法的缺陷,Hansen和Racine提出了Jackknife模型平均方法(JMA),其本质为留一交叉验证。在文章中,作者证明了其在异方差的情况下仍然具有渐进最优性,并且这种方法对于嵌套与非嵌套模型均适用,但此结论仅限于随机样本,存在一定的缺陷[9]。对于MMA方法和JMA方法来说,二者都是基于参数的最小二乘估计,在同

方差条件下, MMA 方法和 JMA 方法几乎是等价的, 但在异方差的条件下, JMA 方法具有更小的风险, 这极大推动了模型平均方法研究的进步。

## 2. 数据与预处理

### 2.1. 数据选取及处理

本次研究所用数据来自于国家气象中心以及天气后报网, 利用 Python 爬虫从网页中爬取青岛市 2018 年至 2023 年空气质量数据以及各项污染物的指标数据, 其中包括可入肺颗粒物(PM<sub>2.5</sub>), 可吸入颗粒物(PM<sub>10</sub>), 二氧化硫(SO<sub>2</sub>), 二氧化氮(NO<sub>2</sub>), 一氧化碳(CO), 臭氧(O<sub>3</sub>)。其中, 将 2018 年至 2022 年数据划分为训练集, 将 2023 年空气质量数据划分为测试集, 对部分缺失数据进行删除, 不会影响本次研究的后续进程。首先利用训练集对各个模型进行训练, 再将测试集中解释变量的值带入训练得到的模型以得到预测值, 并将其与真实值进行对比。

**Table 1.** Example of raw data

**表 1.** 原始数据示例

日期	AQI	PM <sub>2.5</sub>	PM <sub>10</sub>	SO <sub>2</sub>	NO <sub>2</sub>	CO	O <sub>3</sub>
2018-01-01	107	73	145	16	47	1.15	56
2018-01-02	78	43	103	14	49	1.01	39
2018-01-03	54	21	58	14	33	0.6	41
2018-01-04	54	28	61	13	39	0.69	34
2018-01-05	62	37	70	15	42	0.79	37
2018-01-06	111	82	132	31	72	1.37	25
2018-01-07	79	56	98	17	64	1.09	25
2018-01-08	122	74	145	14	40	1.01	39
2018-01-09	69	18	85	14	32	0.67	46
2018-01-10	57	13	65	12	24	0.44	51

对于表 1 所提到的解释变量和被解释变量, 我们对其做出具体解释如下(表 2):

**Table 2.** Description of pollutant indicators

**表 2.** 污染物各项指标说明

符号	含义	单位
PM <sub>2.5</sub>	可入肺颗粒物	mg/m <sup>3</sup>
PM <sub>10</sub>	可吸入颗粒物	μg/m <sup>3</sup>
SO <sub>2</sub>	二氧化硫	μg/m <sup>3</sup>
NO <sub>2</sub>	二氧化氮	μg/m <sup>3</sup>
CO	一氧化碳	mg/m <sup>3</sup>
O <sub>3</sub>	臭氧	μg/m <sup>3</sup>

在实际问题处理过程中, 由于变量间存在单位上的差异, 因此会导致拟合的模型中估计的系数存在问题, 为了消除量纲的影响, 一般处理方法为对原始数据进行标准化, 将数据限制在一个比较狭窄的区间内。常用的标准化方法有高斯法和最大最小值法。

高斯法利用了数据内部的标准差和均值, 即通过标准正态分布的属性进行标准化处理, 将数据限制

在 $[-1, 1]$ 之间。高斯法的具体公式如下：

$$X_j^i = \frac{X_j^i - \mu^i}{\sigma^i} \quad (1)$$

公式中 $i$ 表示数据属性的编号, $j$ 表示观测值的编号, $X_j^i$ 就是第 $j$ 行 $i$ 列样本值, $\sigma^i$ 是第 $i$ 列标准差, $\mu^i$ 是第 $i$ 列的均值。

最大最小值方法是将数据限制在 $[0, 1]$ 的区间内,其中数据的最大值对应的是1,最小值对应的是0,具体公式如下：

$$X_j^i = \frac{X_j^i - \min(X^i)}{\max(X^i) - \min(X^i)} \quad (2)$$

在本次研究中,我们选用高斯法对数据进行预处理,即通过标准正态分布的属性完成数据的标准化处理。

## 2.2. 数据描述性统计

在这一部分,我们将对预处理后的青岛市空气质量数据进行描述性统计分析,以提供对数据的基本了解和总结。其中包括AQI以及各种污染物的均值变化,青岛市空气质量的季节性差异以及相关图表的展示。在这一部分的研究中,对于季节的划分我们采用季节年,即将3月至5月认定为春季,6月至8月认定为夏季,9月至11月认定为秋季,12月至次年2月认定为冬季,使用季节年的优势在于它可以保留完整的冬季。

为了研究青岛市六种空气污染物2018年至2022年总体变化情况,将六种污染物每年的均值绘制成折线图,由于CO的平均浓度小于1,与其余五种污染物均值差距过大,无法用同一坐标轴刻度衡量,因此将CO的年均值变化图单独绘制,如下图1所示：

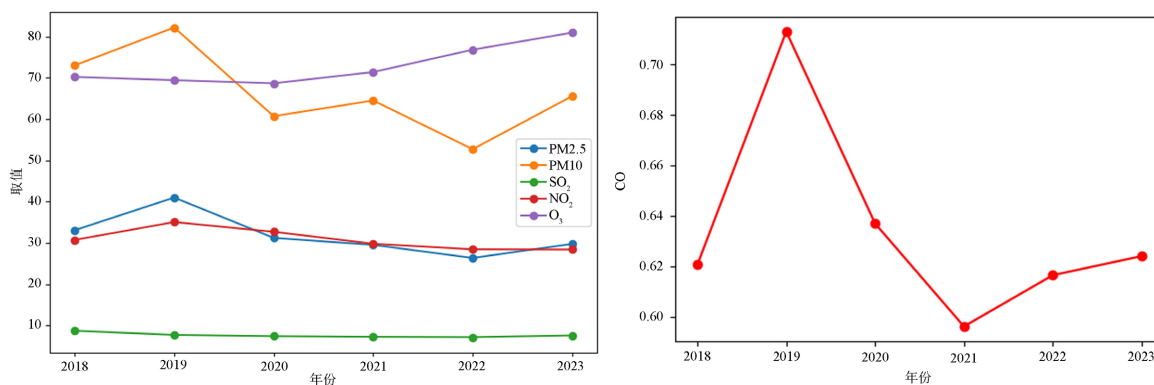
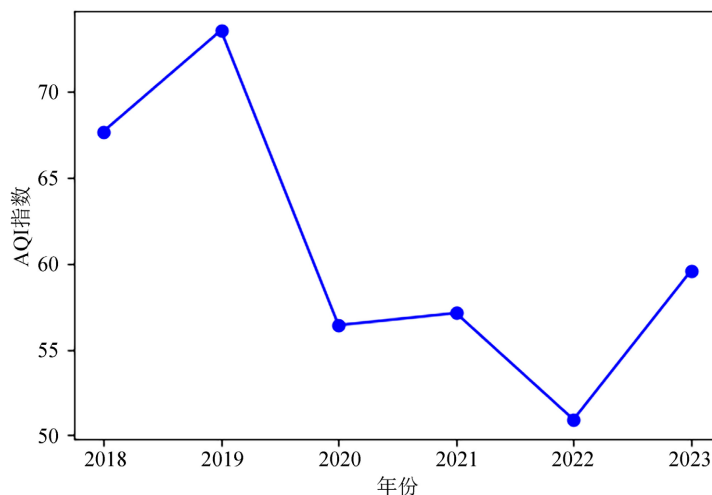


Figure 1. Trends in annual mean concentrations of pollutants

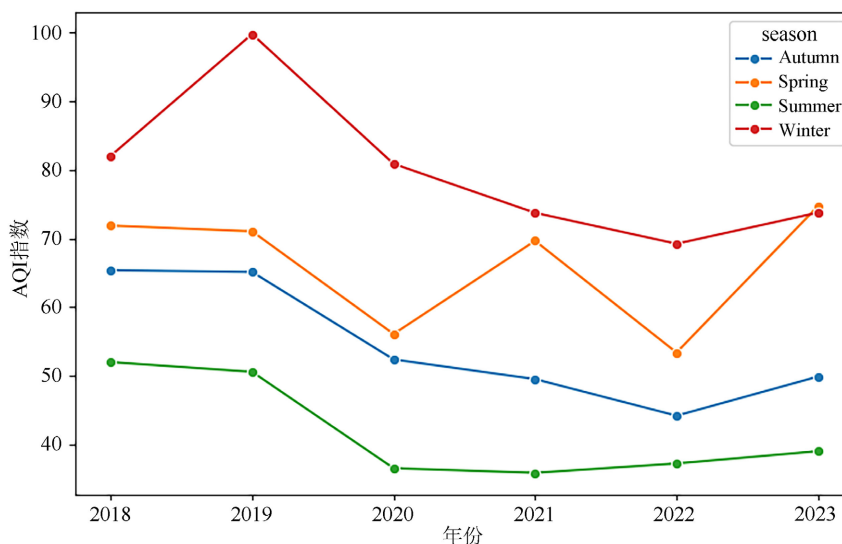
图 1. 污染物年均值变化图

从图中曲线可以看出,青岛市六种污染物中有三种大致呈现平稳下降趋势,但 $O_3$ 在2018年至2023年间年均值呈现上升趋势,主要原因可能是近年来,人类活动,尤其是工业和交通运输的增加,排放了大量的氮氧化物和挥发性有机化合物,这些物质在阳光的照射下发生光化学反应,生成 $O_3$ 。其次,从图中可以看到,PM2.5以及PM10的年均值逐年呈现下降趋势,使得大气中的颗粒物减少,从而为 $O_3$ 的产生提供了较强的光照条件,也是 $O_3$ 浓度逐年上升的主要原因之一。建议青岛市相关部门加强对 $O_3$ 的治理,比如减少 $O_3$ 前体物质的排放,改进工业和交通技术,优化城市规划等。



**Figure 2.** Trends in annual mean AQI  
**图 2.** AQI 指数年均值变化图

由青岛市 AQI 指数年均值变化图可以看出(图 2), 2018 年至 2023 年之间, AQI 指数总体呈现下降趋势, 说明近年来, 青岛市空气质量得到明显的改善。但在 2019 年以及 2021 年出现短暂的上升趋势, 可能是因为由于政策调整或者相关部门执行力度不够, 导致污染控制措施不到位; 也可能由于在数据收集过程中, 增加了新的空气质量监测点或改变了原有空气质量监测点的位置导致测量结果出现波动, 因此 AQI 指数出现局部峰值。



**Figure 3.** Seasonal variation of AQI  
**图 3.** AQI 指数季节性变化图

由 AQI 指数季节性变化图可以看出(图 3), AQI 指数与 PM<sub>2.5</sub>、PM<sub>10</sub> 以及 SO<sub>2</sub> 的变化规律相似, 冬季 AQI 指数最高, 平均值在 80 以上, 其次为春季和秋季, 夏季最低, 平均值均在 50 以下。出现这种变化趋势的原因主要有以下几点: 首先, 夏季高温使得大气对流增强, 有利于污染物的垂直扩散, 另外, 夏季降水较多, 可以清洗空气中的污染物, 减少颗粒物浓度。而冬季取暖需求增加, 尤其在北方地区, 燃煤和其他取暖方式导致大量污染物的排放; 此外, 冬季农作物收割后, 一些地区会进行秸秆焚烧, 增

加空气污染；这些因素共同导致了以上现象的产生。

### 3. 方法与模型

模型平均方法的提出是为了克服模型选择的一些不足而产生的方法，其基本原理可以理解为：假设有  $N$  个模型，每个模型对应一个预测结果  $\mu_i$ ,  $i=1,2,\dots,N$ ，模型平均方法对每个模型的预测结果都产生一个权重  $w_i$ ,  $i=1,2,\dots,N$ 。通过模型平均得到的最终预测值为： $\tilde{\mu} = w_1\mu_1 + w_2\mu_2 + \dots + w_N\mu_N$ 。因此，对于模型平均方法来说，关键在于权重  $w_i$  的选择。

在本次研究中，我们所采用的模型平均方法主要包括四种，基于信息准则的模型选择方法 S-AIC、S-BIC，以及频率模型平均方法 MMA 和 JMA，四种方法的具体概述如下：

#### (1) 基于信息准则的模型选择方法

AIC 和 BIC 方法在模型选择中较为常用，能够衡量模型的复杂度以及拟合程度，二者的表达形式分别如下：

$$\text{AIC} = 2k - 2\ln L \quad (3)$$

$$\text{BIC} = k \ln N - 2\ln L \quad (4)$$

其中， $N$  为样本数量， $L$  为模型的似然函数， $L$  越大，该模型越接近真实模型。 $k$  为模型中参数的个数，参数个数的增加会导致似然函数增大，容易造成过拟合现象，因此公式引入惩罚项来约束参数维度，在似然函数尽可能大的同时，惩罚项应当尽可能小。两种模型选择方法的主要不同在于二者的惩罚项不同。AIC 或 BIC 的值越小，表明该模型所含信息量越大，且保证了模型尽可能简洁，拟合程度越高。

在此基础上，研究者们认为最优模型隐藏在众多候选模型中，这些模型存在不同的参数估计，进一步假设两个模型的惩罚项相同，模型具有相同数量的参数和训练样本，提出了 S-AIC 和 S-BIC 方法，其权重计算的具体表达式如下：

$$w_k = \frac{\exp(-xIC_k/2)}{\sum_k \exp(-xIC_k/2)} \quad (5)$$

式中  $xIC_k$  代表第  $k$  个模型计算所得到的 AIC 或 BIC 值。AIC 是负两倍期望对数似然的近似度量，因此  $\exp(-\phi_k/2)$  近似地为模型  $k$  的相对似然。从信息论角度来看，AIC 又是 K-L 距离的近似无偏估计，因此  $\exp(-\phi_k/2)$  可以看作是模型  $k$  为最小 K-L 距离模型的机率。

#### (2) MMA

考虑提出  $K$  个模型，其中第  $k$  个模型为： $y = X_{(k)}\theta_k + e = \hat{\mu}_{(k)} + e$ ，其中  $\theta_k$  基于第  $k$  个近似模型的最小二乘估计为  $\hat{\theta}_k = (X'_{(k)}X_{(k)})^{-1} X'_{(k)}y$ ，相应的， $\mu$  的最小二乘估计值为  $\hat{\mu}_{(k)} = X_{(k)}\hat{\theta}_k = X_{(k)}(X'_{(k)}X_{(k)})^{-1} X'_{(k)}y$ 。

记权重向量  $w = (w_1, \dots, w_k)'$ ，并限制在如下的权重集合中：

$$H_n = \left\{ w \in [0,1]^K : \sum_{k=1}^K w_k = 1 \right\} \quad (6)$$

此时， $\mu$  的模型平均估计为：

$$\hat{\mu}(w) = \sum_{k=1}^K w_k \hat{\mu}_{(k)} \quad (7)$$

此方法通过极小化 Mallows 准则来进行权重选择，权重选择的 Mallows 准则为：

$$C_n(w) = w' \hat{e}' \hat{e} w + 2\sigma^2 w' \phi \quad (8)$$

其中， $\hat{e} = (\hat{e}_1, \dots, \hat{e}_k)$ ， $\hat{e}_k = y - \hat{\mu}_{(k)}$  为第  $k$  个模型估计的残差向量， $\phi = (\phi_1, \dots, \phi_k)$  是第  $k$  个模型中参数的有效数量。

通过极小化 Mallows 准则得到的权重如下：

$$\hat{w}_k = \arg \min_{w \in H_n} C_n(w) \tag{9}$$

(3) JMA

用 Jackknife 准则求解模型平均问题，首先要计算 Jackknife 残差。第  $m$  个估计值为  $\tilde{\mu}^m = (\tilde{\mu}_1^m, \tilde{\mu}_2^m, \dots, \tilde{\mu}_k^m)'$ ，其中  $\tilde{\mu}_i^m$  表示除去第  $i$  个观测值后所得到的观测值，同样可以写为  $\tilde{\mu}^m = \tilde{P}_m y$ ，Hansen 证实  $\tilde{P}_m$  和  $P_m$  的关系，即  $\tilde{P}_m = D_m (P_m - I) + I$ 。

此方法通过极小化 Jackknife 准则来进行权重选择，Jackknife 准则为：

$$J_m(w) = w' \hat{e}' \hat{e} w \tag{10}$$

其中， $\hat{e} = (\hat{e}_1, \dots, \hat{e}_m)$ ， $\hat{e}_m = y - \tilde{\mu}^m$ ， $\tilde{\mu}^m = (D_m (P_m - I) + I) y$ ， $D_m$  为对角阵，它的第  $i$  个对角元素为  $(1 - h_{ii}^m)^{-1}$ ，其中  $h_{ii}^m$  是  $P_m$  的第  $i$  个对角元素。因为  $\tilde{\mu}^m$  的第  $i$  行并不依赖于  $y_i$ ，所以认为  $\hat{e}_m$  是弃一的交叉验证的误差向量，并且弃一的交叉验证在统计学中被称为 Jackknife 方法。与 MMA 方法不同的是，这里仅有  $E(e_i | x_i) = 0$ 。

通过极小化 Jackknife 准则得到的权重如下：

$$\hat{w}_k = \arg \min_{w \in H_n} J_n(w) \tag{11}$$

$$H_n = \left\{ w : 0 \leq w_k \leq 1, \sum_{k=1}^K w_k = 1 \right\} \tag{12}$$

为了保证建立模型的科学性以及了解解释变量之间的相关关系，本次研究首先对所有解释变量进行相关性分析，绘制变量的相关系数矩阵，具体结果如图 4 所示：

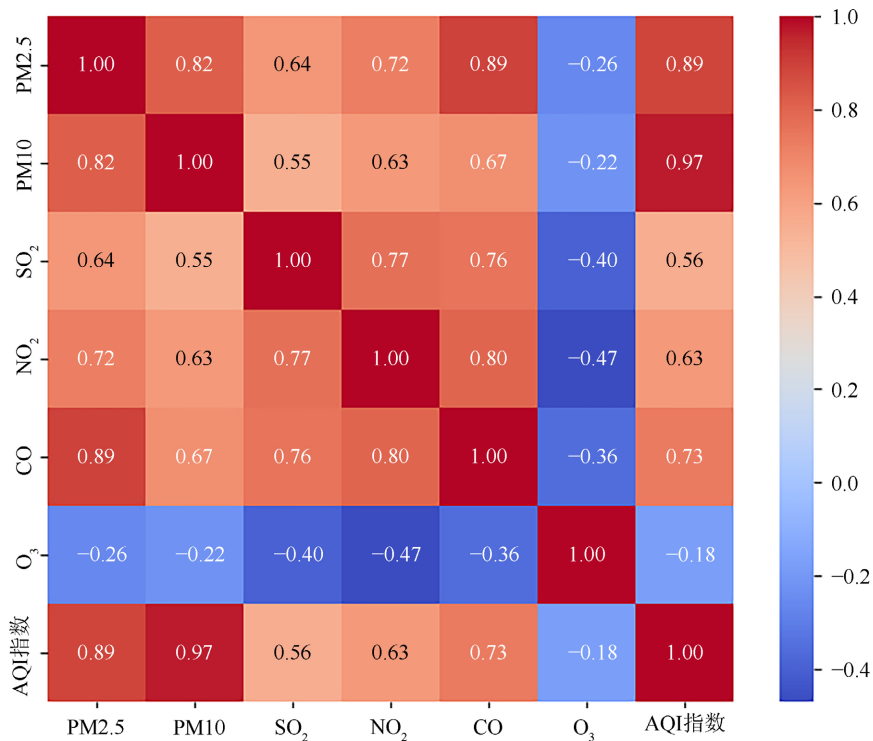


Figure 4. Correlation matrix of explanatory variables  
图 4. 解释变量相关性矩阵

从解释变量相关性矩阵可以看出, AQI 和各个解释变量之间均存在较强的相关性。具体来说, AQI 与 PM2.5、PM10、SO<sub>2</sub>、NO<sub>2</sub> 以及 CO 呈现正相关性, 与 O<sub>3</sub> 呈负相关性; 与此同时, 各个解释变量之间也存在相关性, 比如 PM2.5 和 PM10 之间的相关系数为 0.82, 存在较强正相关性。如果直接对各项指标进行回归分析, 解释变量之间可能会出现多重共线性, 这将会扩大我们模型系数的估计误差。为了进一步确认解释变量之间是否存在多重共线性, 我们通过计算方差膨胀因子来进行定量判断, 具体结果如表 3 所示:

**Table 3.** Variance Inflation Factor (VIF)

**表 3.** 方差膨胀因子

解释变量	PM2.5	PM10	SO <sub>2</sub>	NO <sub>2</sub>	CO	O <sub>3</sub>
VIF	9.79	3.92	2.96	3.93	9.09	1.36

从结果来看, 各个解释变量的方差膨胀因子均在 10 以下, 我们可以认为不存在多重共线性问题。

## 4. 实证分析

### 4.1. 模型建立

本次研究将所选的 6 个变量均视为可选变量, 那么我们可以拟合的模型共有 64 个, 例如第一个模型为只包含常数项的模型  $y = \beta_0 + \varepsilon$ , 第二个模型为  $y = \beta_0 + \beta_1 X_1 + \varepsilon$ , 以此类推, 最后一个模型为全模型, 即包含全部的解释变量, 其形式为:  $y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \varepsilon$ 。

考虑到模型的复杂度以及实际研究时的难易度, 本次研究根据 AIC 信息准则, 通过嵌套的方式选出 6 个最具代表性的模型, 即当含有不同解释变量个数时, 选取此时最优模型, 以此来进行后续模型平均研究。具体模型建立以及变量选择结果如表 4。

**Table 4.** Selected models and included variables

**表 4.** 选定模型及包含的变量情况

模型	选取的解释变量
1	PM2.5 + PM10 + SO <sub>2</sub> + NO <sub>2</sub> + CO + O <sub>3</sub>
2	PM2.5 + PM10 + SO <sub>2</sub> + NO <sub>2</sub> + O <sub>3</sub>
3	PM2.5 + PM10 + NO <sub>2</sub> + O <sub>3</sub>
4	PM2.5 + PM10 + O <sub>3</sub>
5	PM2.5 + PM10
6	PM10

对于上述选择出的最具代表性的模型, 我们通过最小二乘法对其进行拟合, 得到不同的参数估计值以及 6 个不同的模型。对于 S-AIC 和 S-BIC 方法来说, 我们通过计算每个模型的 AIC 和 BIC 值, 进而求得各个模型的最优权重; 对于 MMA 和 JMA 方法, 我们分别最小化 Mallows 准则以及 Jackknife 准则得到最优权重, 四种模型平均方法的具体权重如表 5 所示:

**Table 5.** Model weights from model averaging approach

**表 5.** 模型平均方法各模型权重

模型	W <sub>S-AIC</sub>	W <sub>S-BIC</sub>	W <sub>MMA</sub>	W <sub>JMA</sub>
1	0.122	0.001	0	0.1655

续表

2	0.323	0.036	0	0.1349
3	0.555	0.963	0.05	0.1377
4	0	0	0.614	0.1829
5	0	0	0	0.1829
6	0	0	0.336	0.1962

从权重计算的结果我们可以看出，基于信息准则的模型平均方法所得的权重类似，主要集中在前三个模型，其中模型 3 所占权重最大，而后三个模型权重为 0；而对于 MMA 方法，权重主要集中在模型 3、4 以及模型 6，其中模型 4 所占比重最大，已超过 0.6；JMA 方法对于各个模型的权重相似，集中在 0.15~0.20 之间。

基于表 5 的结果，进一步分析不同模型平均方法的权重分配机制，我们可以发现，各方法对于不同模型的权重差异主要源于模型复杂度和拟合优度的权衡方式不同。基于信息准则的 S-AIC 和 S-BIC 方法的权重计算依赖于 AIC 和 BIC 的取值，所以当某一模型在信息准则上明显优于其他模型时，其所被赋予的权重会更大。说明该类基于信息准则的模型平均方法更倾向于选择更优的少数模型。相比之下，基于留一交叉验证的 JMA 方法在预测过程中的目标是 minimized 预测误差，因此在权重分配上可能会更加均衡。这种均衡的权重分配方法更加注意预测结果的稳健性，通过综合不同预测模型的信息来降低传统的单一预测模型存在的误差，从而提升预测模型的泛化能力。而基于最小化 Mallows 准则的 MMA 方法在一定程度上兼顾了预测的精度和复杂度，因此其权重不像 S-AIC 和 S-BIC 方法过于集中，也不像 JMA 方法过于均衡，不同的模型平均方法呈现出不同的权重侧重。

## 4.2. 预测结果比较

利用上述研究建立的模型以及所求得的权重，对未来青岛市空气质量情况进行预测。通过已有模型，预测青岛市 2023 年 1 月 1 日至 2023 年 1 月 10 日污染物浓度数据，四种模型平均方法的预测值以及真实值的结果如下(表 6)：

**Table 6.** 10-day forecast values based on model averaging

**表 6.** 基于模型平均的 10 天预测值

期数	S-AIC	S-BIC	MMA	JMA	真实值
2023-01-01	139.47054	139.58177	131.68798	135.02377	143
2023-01-02	86.08309	86.13120	85.10730	86.15320	87
2023-01-03	86.40351	86.60269	85.90462	86.63505	85
2023-01-04	136.58732	136.65731	129.29718	132.60657	142
2023-01-05	161.98488	161.87981	152.47799	156.73028	170
2023-01-06	162.38777	162.41969	155.48093	158.32361	153
2023-01-07	106.22274	106.37835	110.37137	109.14403	101
2023-01-08	124.05622	123.99813	121.27023	122.51599	120
2023-01-09	90.61069	90.61788	88.51418	89.76253	87
2023-01-10	37.23318	37.25998	39.45732	39.16200	39

在预测青岛市未来十期空气质量指数时，从结果来看，当我们进行短期预测时，我们应该选择基于信息准则的模型平均方法，S-AIC 和 S-BIC 方法的精确度相似；在进行长期预测时，可以选用 MMA 或者 JMA 方法来进行预测。在实际研究的过程中，研究者应根据自己的研究过程以及最终的研究目的来选择更合适的方法。

**Table 7.** Absolute errors of predictions from model averaging  
**表 7.** 模型平均方法预测的绝对误差

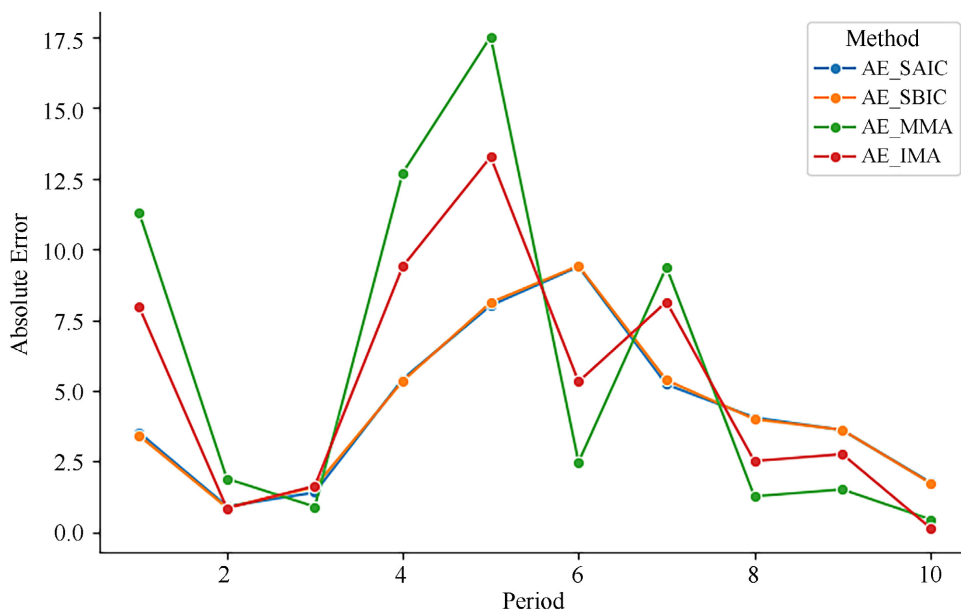
期数	S-AIC	S-BIC	MMA	JMA	最优方法
2023-01-01	3.52946	3.41823	11.31202	7.97623	S-BIC
2023-01-02	0.91691	0.8688	1.8927	0.8468	JMA
2023-01-03	1.40351	1.60269	0.90462	1.63505	MMA
2023-01-04	5.41268	5.34269	12.70282	9.39343	S-BIC
2023-01-05	8.01512	8.12019	17.52201	13.26972	S-AIC
2023-01-06	9.38777	9.41969	2.48093	5.32361	MMA
2023-01-07	5.22274	5.37835	9.37137	8.14403	S-AIC
2023-01-08	4.05622	3.99813	1.27023	2.51599	MMA
2023-01-09	3.61069	3.61788	1.51418	2.76253	MMA
2023-01-10	1.76682	1.74002	0.45732	0.162	JMA
平均值	4.332192	4.350667	5.94282	5.202939	
最优率	0.2	0.2	0.4	0.2	

表 7 中数据表示四种不同的模型平均方法对空气质量指标 AQI 的预测结果与真实值之间的绝对误差，纵向分析所得结果可以得到每一种方法十次预测的绝对误差的平均值，横向比较每一期的预测结果，可以判断在每期预测中预测精度最高的方法，即最优方法，将各种方法最优次数除以总次数则得到各个方法的最优率。

由结果可以看出，从各方法的绝对误差平均值来看，S-AIC 和 S-BIC 方法表现最好，平均绝对误差仅有 4.33 与 4.35，相反，MMA 方法表现最差，平均绝对误差高达 5.94。而从最优率的角度来看，MMA 方法最优率最高，达到 40%，在 10 期预测中取得最优的次数为 4 次；其次是 S-AIC、S-BIC、JMA 方法，最优率均为 20%。结合平均绝对误差和最优率来看，S-AIC 和 S-BIC 方法表现最优，拟合效果较好，预测精度较高，由于这两种方法均是基于信息量准则加权，所以表现相差不大。由于 S-AIC 和 S-BIC 方法原理简单，操作方便，在对预测精度要求不高时可以作为较好的常用方法。其次是 MMA 方法和 JMA 方法，二者在短期预测中并不存在优势，特别是 MMA 方法，但我们从数据中可以看出，MMA 方法和 JMA 方法在长期预测中是有巨大优势的。

为了更直观地观察四种模型平均方法的预测结果趋势以及具体的波动情况，将上述 10 期预测值的绝对误差绘制成折线图，结果如下所示。

如图 5 所示，四种模型平均方法的绝对误差趋势相似，对于短期预测来说，两种基于信息准则的模型平均方法绝对误差较小，但在长期预测中，MMA 和 JMA 方法的绝对误差低于 S-AIC 和 S-BIC 方法。



**Figure 5.** Line chart of absolute errors from model averaging method  
**图 5.** 模型平均方法预测的绝对误差折线图

为了进一步探究四种方法在实际应用过程中的效果，我们计算了模型预测结果的均方误差。因为在实际应用中，我们并不能直接根据绝对误差的大小来判断方法的优劣，预测的绝对误差大不代表这种方法不好，也可能是数据点存在异常值所导致的，所以我们需要结合均方误差来进一步比较。一般来说，均方误差越小，我们就认为预测的波动越小，预测结果越好，预测精度越高。均方误差的计算结果如下表 8 所示：

**Table 8.** Mean Squared Error (MSE) of model averaging methods  
**表 8.** 模型平均方法的均方误差

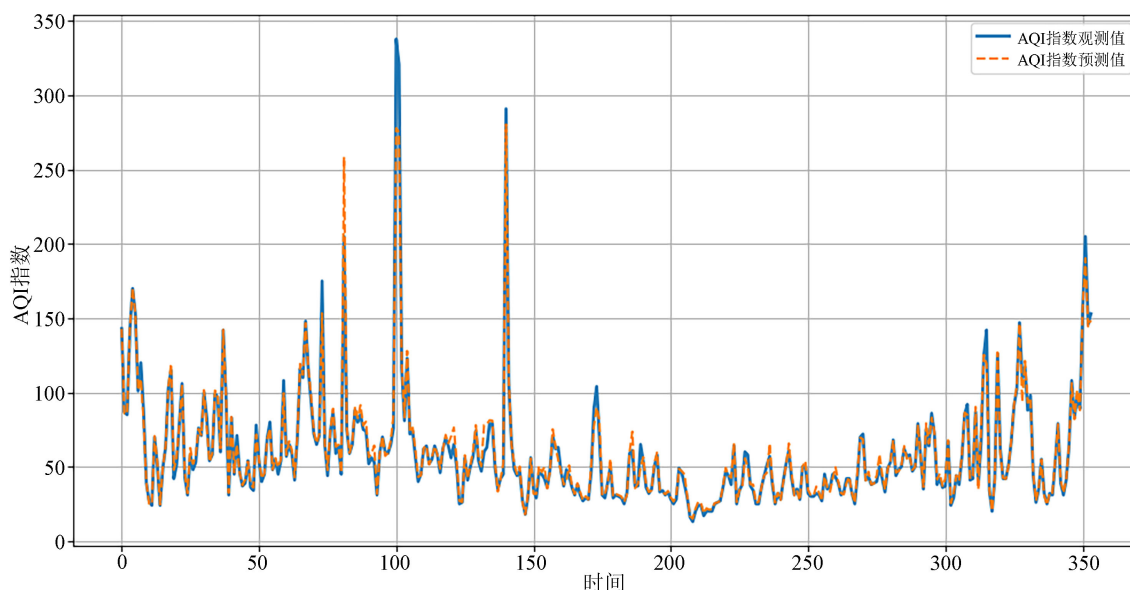
	S-AIC	S-BIC	MMA	JMA
MSE	47.30368	47.4457	40.25261	41.6656

通过计算结果可知，四种模型平均方法的均方误差差别不大，相比之下，MMA 方法的均方误差最小，为 40.25，其次为 JMA 方法，其 MSE 为 41.67。相较于 MMA 和 JMA 方法，S-AIC 和 S-BIC 方法预测结果的均方误差较大，均大于 47.3，预测结果波动性强，最不稳定。

为了探究本次研究所提出模型的有效性，我们选择与传统机器学习方法随机森林模型进行结果对比。随机森林方法作为一种典型的集成学习方法，能够有效捕捉各个解释变量之间的非线性关系，在空气质量预测领域具有广泛应用，因此，在本次研究中，选择随机森林模型作为基准模型，有助于验证本文所提出的模型平均方法的预测精准度和有效性。

在随机森林模型构建的过程中，为了确保对比的公平性和一致性，采用与模型平均方法相同的输入变量，即选择六种大气污染物作为解释变量，并采取相同的数据集划分方式，同时，验证相同的模型评价指标。

通过拟合随机森林模型，为了直观地展示模型预测结果，绘制观测值与预测值的对比曲线图，具体如下(图 6)：



**Figure 6.** Prediction results of random forest model  
**图 6.** 随机森林模型预测结果

具体模型评价指标结果如下：

**Table 9.** Prediction metrics of random forest model  
**表 9.** 随机森林模型预测指标

模型评价指标	数值
MSE	41.68
MAE	3.01

根据表 9 的评价指标结果，从整体结果来看，本次研究中所提出的多种模型平均方法和随机森林模型相比在均方误差方面表现相似，其中 MMA 方法优于随机森林方法，说明模型平均方法在预测空气质量的过程中存在较好的稳定性。

## 5. 结论和展望

### 5.1. 研究结论

在基于已有的对于青岛市空气质量分析研究的基础上，本研究基于 Python 爬虫方法，获取了青岛市 2018 年至 2023 年空气质量日数据，并将这些数据作为初始数据进行后续的模型平均分析。

基于本研究所使用的四种模型平均方法以及所得的预测结果来看，我们可以得出以下结论：比较四种模型平均方法，结合平均绝对误差和最优率来看，S-AIC 和 S-BIC 方法表现最优，拟合效果较好，预测精度较高，由于这两种方法均是基于信息量准则加权，所以表现相差不大。其次是 MMA 方法和 JMA 方法，二者在短期预测中并不存在优势，特别是 MMA 方法，但从数据中可以看出，MMA 方法和 JMA 方法在长期预测中是有巨大优势的。

综上所述，从绝对误差、最优率和均方误差以及长短期预测结果来看，MMA 方法在本次研究中具有较大的优势，绝对误差较小，最优率最大，均方误差最小，且在长期预测中具有较高稳定性，是表现最好的方法。

## 5.2. 未来展望

本次研究通过对青岛市空气质量分析实例比较四种模型平均方法, 通过比较预测的绝对误差均值以及均方误差来判断方法的准确程度和在预测中的稳定性, 进一步分析四种模型平均方法的优势和劣势, 但在本次研究中仍存在许多局限性, 具体如下:

首先, 在收集原始数据的过程中, 由于天气后报网记录缺失, 数据存在许多缺失值, 本次研究所采用的处理方法是将缺失值删除, 但是直接删除缺失值可能会导致信息丢失、样本偏差、统计功效降低等问题, 因此在后续的研究中, 为了保证研究结果的准确性以及精确度, 应使用适当的替代方法来处理缺失值, 比如均值填补以及插值法, 以最大限度利用数据。

第二, 在本次研究的过程中, 考虑到模型平均的预测效率以及相关结果, 仅基于 AIC 准则选取了 6 个最优模型, 并未将所有模型考虑在内, 存在一定偏差。在后续的研究过程中, 应该包含所有可能的模型, 使分析结果更有一般性, 更具说服力。

最后, 在之前的研究中, 模型平均方法被证明是有效的, 其克服了模型选择的不确定性和不稳健性。但是在提高精度的同时, 当模型中不确定因素增加时, 模型平均的计算量将会呈指数倍增长, 导致模型平均在高维数据下的优势并不明显。并且现在的模型平均方法大都基于多元线性回归模型, 会存在模型不能很好拟合的情况, 导致模型平均方法的表现并不突出。后续研究中还可引入非参数模型以及时间序列模型对空气质量进行分析, 并将各种模型的模型平均结果进行对比, 得出进一步的结论。

## 参考文献

- [1] 宋艳玲, 郑水红, 柳艳菊, 张永山. 2000-2002 年北京市城市大气污染特征分析[J]. 应用气象学报, 2005(S1): 116-122.
- [2] 刘丽, 唐晓婵, 马东卓, 等. 2013-2019 年青岛市空气质量变化特征分析[J]. 中国环境监测, 2020, 36(6): 66-71.
- [3] 武苗苗. 山东省城市区域空气污染的时空变化特征研究[D]: [硕士学位论文]. 上海: 华东师范大学, 2016.
- [4] 徐红霞, 刘汝海, 王艳, 张燕燕, 汪明宇, 张玉卿. 2014-2016 年山东省空气质量变化及典型城市污染天气污染物来源分析[J]. 中国人口·资源与环境, 2017, 27(S2): 132-135.
- [5] Min, C. and Zellner, A. (1993) Bayesian and Non-Bayesian Methods for Combining Models and Forecasts with Applications to Forecasting International Growth Rates. *Journal of Econometrics*, **56**, 89-118. [https://doi.org/10.1016/0304-4076\(93\)90102-b](https://doi.org/10.1016/0304-4076(93)90102-b)
- [6] Buckland, S.T., Burnham, K.P. and Augustin, N.H. (1997) Model Selection: An Integral Part of Inference. *Biometrics*, **53**, 603-618. <https://doi.org/10.2307/2533961>
- [7] Hansen, B.E. (2007) Least Squares Model Averaging. *Econometrica*, **75**, 1175-1189. <https://doi.org/10.1111/j.1468-0262.2007.00785.x>
- [8] Zhang, X.Y. and Zou, G.H. (2011) Model Averaging Method and Its Application in Forecast. *Statistical Research*, **28**, 98-103. (In Chinese)
- [9] Hansen, B.E. and Racine, J.S. (2012) Jackknife Model Averaging. *Journal of Econometrics*, **167**, 38-46. <https://doi.org/10.1016/j.jeconom.2011.06.019>