

基于 LMDI 和 ARIMA 的居住建筑碳排放评价与预测

参赛题号： B

参赛队号： 031

参赛队员： 2106050218 徐诚鸿 软件工程 建模、编程、写作

2106050203 余成远 软件工程 建模、编程、写作、作图

2106030211 李懿斐 电子信息工程 资料检索、建模

联系方式： Email: chxu@hhu.edu.cn 联系电话: 15956065027

日期: 2023 年 5 月 1 日

摘要：我国正处于高速发展阶段，对能源资源的需求量较大。从建筑层面来说，我国建筑耗能占全国总能耗的 28%，导致碳排放量急剧上升。因此，在建筑行业实行“双碳”政策，发展低碳经济势在必行。针对与建筑碳减排相关的 5 个问题，本文做了如下工作：

针对问题 1，根据其各组成部分的热导系数计算达到室内外温度平衡时所需的吸热量或放热量，结合每月平均气温、空调制热与制冷系数，计算每月耗能，并进行单位换算，从而得出此建筑年碳排放量为 186.266 千克。

针对问题 2，使用对数平均权重分解法（LMDI），在时间层面上由建筑碳排放总量分解出三个生命周期中易于量化的关键指标，构建评价模型，对南京市建筑全生命周期碳排放量进行评价。

针对问题 3，对问题 2 中 LMDI 影响因子权重相对大小进行排序，选取长期的、相对独立于时效性的三个关键指标。结合 2020-2021 年的相关数据，再次使用 LMDI 方法建立适用于 2021 年的江苏省建筑碳排放评价模型，对 13 个地级市进行评价后，使用 LMDI 残差为 0 这一重要特点，对模型进行有效性检验。

针对问题 4，基于 1999-2019 年江苏省历史居住建筑碳排放情况，构建碳排放时间序列数据，进行平稳性检验后，使用 ARIMA 方法进行建模与调优，使用 AIC 方法得到最优模型为 ARIMA(1, 1, 0)，并预测 2023 年江苏省建筑全过程碳排放量为 190.772 百万吨。

针对问题 5，从建造、使用与拆除的生命周期层面与不同地区建筑水平差异层面给出综合性建议。

关键词：建筑碳排放量、LMDI、残差、AIC、ARIMA

1 问题背景、重述与分析

1.1 问题背景

“双碳”是指碳达峰和碳中和，其政策内容旨在实现通过绿色、环保的生活方式，在 2030 年前实现碳达峰，2060 年前实现碳中和。

我国正处于高速发展阶段，对能源资源的需求量较大。从建筑层面来说，我国建筑耗能占全国总能耗的 28%，导致碳排放量急剧上升。因此，在建筑行业实行“双碳”政策，发展低碳经济势在必行。

低碳建筑是发展绿色建筑行业的关键一环。低碳建筑使用绿色建筑材料，减少对化石能源的依赖。并贯穿建筑生命周期中建造、运行、拆除三个阶段，全方位降低二氧化碳排放量，实现低碳经济。

1.2 问题重述

本题以“双碳”政策和低碳建筑为背景，由小及大、由浅入深地提出了四个问题：

- (1) 结合给定的单层平顶单体建筑实际样例，计算其使用空调调节室温为 18-26 度一年的碳排放量。
- (2) 建立居住建筑的碳排放评价模型，寻找在生命周期中与碳排放相关程度较大且易于量化的指标，对居住建筑整个生命周期的碳排放进行综合评价。
- (3) 使用 (2) 中的模型，分别考虑建筑生命周期三个阶段的碳排放问题。结合相关数据，对 2021 年江苏省 13 个地级市的居住碳排放数据进行综合评价，并分析模型有效性。
- (4) 建立碳排放预测模型，基于江苏省建筑全过程碳排放的历史数据，对 2023 年江苏省建筑全过程的碳排放量进行预测。
- (5) 结合上述分析，为江苏省建筑碳减排提供政策建议。

1.3 问题分析

根据题目要求，我们对问题做出如下分析：

- (1) 对于给定的单层平顶单体建筑，可根据其各组成部分的热导系数计算达到室内外温度平衡时所需的吸热量或放热量，结合每月平均气温、空调制热与制冷系数计算每月耗能，并进行单位换算，从而得出年碳排放量。
- (2) 居住建筑的建造、运行、拆除生命周期中，影响碳排放的因素众多。可使用对数平均权重分解法 (LMDI)，在时间层面上分解出三个生命周期中易于量化的关键指标，以构建评价模型。
- (3) 基于 (2) 中的模型分析结果，可寻找到适合 2021 年这一具体时间点的、影响因子较大的指标。结合该年的相关数据，使用 LMDI 方法建立适用于该年的碳排放评价模型。并利用 LMDI 相关公式，计算出其理论碳排放数据，与实际排放数据进行比较，从而验证模型的正确性。
- (4) 基于江苏省历史居住建筑碳排放情况，构建碳排放时间序列数据，并使用时间序列分析的相关方法，对 2023 年碳排放量进行预测。
- (5) 给出的建议包括宏观层面与细节层面。在宏观层面，依据建筑生命周期中的关键指标，提供相关政策建议；同时依据江苏省 13 个地级市的评价结果，给出地区性建议；在细节层面，依据上述关键指标，给出具体碳减排的措施，结合 (1) 中的计算结果，给出细节建议。

2 名词解释与符号说明

本文使用的名词与符号说明如表 1所示：

表 1: 名词解释与符号说明

符号	说明
CE, C	碳排放量
K	热导系数
ω	单位温度变化时的热交换量
COP	空调制热性能系数
EER	空调制冷性能系数
HI	居住建筑房地产投资
NRA	新增住宅面积
REC	居民用电
TLA	总居住面积
PR	常住人口
DPN	拆迁人数

3 问题假设

为了方便建立模型，我们对问题作出如下假设：

1. 假设建筑物内温度需要一直保持在 18-26 度，在温度不适宜的时候要通过电来调节温度，消耗一度电相当于 0.28 千克碳排放。
2. 建筑物内空气的传热系数较小，空气的传热量可以忽略不计。
3. 居住建筑碳排放量在涉及到评价时可用居民建筑分布广泛地区的碳排放总量进行替代（视为等比例放大），在预测时，根据问题背景，使用总耗能的 28% 进行计算。

4 问题一：单层平顶单体建筑年碳排放量求解

查阅资料可知，单位时间内热交换量 Q 与热导系数 K 存在如下关系：

$$Q = KA\Delta T \quad (1)$$

其中， A 为材料的导热面积， ΔT 为温度的变化量。同时， COP 是空调的制热能效比， EER 是制冷能效比，计算公式如下：

$$COP = \frac{\text{空调额定制热量}}{\text{额定制热耗电量}} \quad (2)$$

$$EER = \frac{\text{空调额定制冷量}}{\text{额定制冷耗电量}} \quad (3)$$

由此可得出年碳排放量计算式如下：

$$\omega = \sum K_i A_i \quad (4)$$

$$\Delta T = \begin{cases} T_{\text{月平均温度}} - T_{\text{最高室温}} & T_{\text{月平均温度}} > T_{\text{最高室温}} \\ T_{\text{最低室温}} - T_{\text{月平均温度}} & T_{\text{月平均温度}} < T_{\text{最低室温}} \end{cases} \quad (5)$$

$$W_m = \begin{cases} \frac{\omega \cdot \Delta T \cdot t_{\text{month}}}{COP} & T_{\text{月平均温度}} > T_{\text{最高室温}} \\ \frac{\omega \cdot \Delta T \cdot t_{\text{month}}}{EER} & T_{\text{月平均温度}} < T_{\text{最低室温}} \end{cases} \quad (6)$$

$$CE = 0.28 \sum_{i=1}^{12} W_m \quad (7)$$

其中, ω 为单位温度变化时的热交换量, 其值为。

$$\omega = 36.0W/K \quad (8)$$

K_j 为每部分 (墙体、屋顶、门窗、地面) 的热导系数, W 为总热交换量, t_{month} 为每月时间, 使用单位 $kW \cdot h$ 进行表述。

每月的温变 ΔT 列表如表所示:

表 2: 每月温度变化表

月份	1	2	3	4	5	6	7	8	9	10	11	12
$\Delta T/K$	19	16	12	6	0	2	5	6	0	0	3	16

经计算, 该建筑物通过空调调节温度的年碳排放量 CE 的值为 186.266 千克。

5 问题二: 基于 LMDI 的居住建筑碳排放评价模型

5.1 LMDI 评价指标模型建立

在建筑的三个生命周期中, 影响其碳排放的因素众多, 因此需要对这些因素所对应的指标在碳排放量中所占权重进行分析与选择。结合对国内外碳排放研究方法的调研, 模型主要构建方法有主成分分析法 (PCA)、对数平均权重分解法 (LMDI) 等。本部分使用 LMDI 对建筑不同生命周期阶段的碳排放相关指标建立模型, 进行评价。

指数因素分解法是指通过一系列的数学恒等公式变形将某一研究对象表示为相关影响因素的乘积形式, 并根据不同权重的计算每个指标的增量余额。国内外学者对于能源消费和碳排放的研究中, 运用指数因素分解法分析研究对象的影响因素。指数因素分解的基本模型为:

$$V^t = \sum_t X_{1i}^t X_{2i}^t \dots X_{ni}^t \quad (9)$$

其中:

V 代表能源消费量或者碳排放量;

X 表示与 V 相关的影响因素;

i 表示影响因素的区别标志, 如对应的能源种类、产业部门等;

t 表示各个时期。

等式两边取对数并对时间求导, 得到公式的变形:

$$\frac{d \ln V}{dt} = \sum_i \lambda_i \left(\frac{d \ln X_{1i}}{dt} + \frac{d \ln X_{2i}}{dt} + \dots + \frac{d \ln X_{ni}}{dt} \right) \quad (10)$$

对等式两边取定积分：

$$\ln \frac{V^t}{V^0} = \int_0^t \sum_i \lambda_i \left(\frac{d \ln X_{1i}}{dt} + \frac{d \ln X_{2i}}{dt} + \dots + \frac{d \ln X_{ni}}{dt} \right) dt = \sum_i \left(\lambda_i \frac{d \ln X_{1i}}{dt} + \lambda_i \frac{d \ln X_{2i}}{dt} + \dots + \lambda_i \frac{d \ln X_{ni}}{dt} \right) dt \quad (11)$$

上述等式中，对于括号内的定积分无法算出确切数值，通常通过赋予不同的权重函数来计算各个影响因素的贡献值。

LMDI 将时期内两个端点的平均对数函数作为分解权重，即：

$$\begin{cases} L(x, y) = (y - x) / \ln(y - x) \\ L(x, x) = x \end{cases} \quad (12)$$

则权重函数可以表示为：

$$\omega_i = \lambda_i = \frac{(V_t^t - V_t^0) \ln(V_t^t - V_t^0)}{(V^t - V^0) \ln(V^t - V^0)} \quad (13)$$

基于加法的 LMDI 分解公式为：

$$\Delta V = V^t - V^0 = \Delta V_{X1} + \Delta V_{X2} + \dots + \Delta V_{Xk} + \Delta V_{rsd} \quad (14)$$

其中， $\Delta V_{Xk} = \frac{V_t^t - V_t^0}{\ln(V^t/V^0)} \ln D_{Xk}$ 。

而 ΔV_{rsd} 在 LMDI 中，值为 0，即无残差。这是 LMDI 分解法的一个重要特点 [1, 2, 3]。

本部分针对南京市建筑的三个生命周期，分别选取的指标如表 3 所示：

表 3: 建筑生命周期不同阶段指标		
生命周期	指标	符号
建造	居住建筑房地产投资（亿元）	HI
	新增住宅面积（千万平方米）	NRA
运行	居民用电（亿千瓦时）	REC
	总居住面积（万平方米）	TLA
	常住人口（万人）	PR
拆除	拆迁人数（万户）	DPN

构建碳排放量恒等式：

$$CE = \frac{CE}{HI} \times \frac{HI}{NRA} \times \frac{NRA}{REC} \times \frac{REC}{TLA} \times \frac{TLA}{PR} \times \frac{PR}{DPN} \times DPN \quad (15)$$

$$CE = c \cdot h \cdot n \cdot r \cdot t \cdot p \cdot DPN \quad (16)$$

上式中， $c = \frac{CE}{HI}$ 表示碳排放-居住投资转化比， $h = \frac{HI}{NRA}$ 表示居住投资价值， $n = \frac{NRA}{REC}$ 表示新增住宅面积用电转化比， $r = \frac{REC}{TLA}$ 表示住宅面积用电转化比， $t = \frac{TLA}{PR}$ 表示人口居住密度， $p = \frac{PR}{DPN}$ 表示拆迁规模。

则依据式 (13)~(15) 可得出如下分解模型：

$$\Delta CE = CE_t - CE_0 = \Delta C_c + \Delta C_h + \Delta C_n + \Delta C_r + \Delta C_t + \Delta C_p + \Delta C_{DPN} \quad (17)$$

其中， $\Delta C_c = \sum_i F(C_{it}, C_{i0}) \ln(\frac{c_t}{c_0})$,

$\Delta C_h = \sum_i F(C_{it}, C_{i0}) \ln(\frac{h_t}{h_0})$,

$\Delta C_n = \sum_i F(C_{it}, C_{i0}) \ln(\frac{n_t}{n_0})$,

$$\begin{aligned}\Delta C_r &= \sum_i F(C_{it}, C_{i0}) \ln\left(\frac{r_t}{r_0}\right), \\ \Delta C_t &= \sum_i F(C_{it}, C_{i0}) \ln\left(\frac{t_t}{t_0}\right), \\ \Delta C_p &= \sum_i F(C_{it}, C_{i0}) \ln\left(\frac{p_t}{p_0}\right), \\ \Delta C_{DPN} &= \sum_i F(C_{it}, C_{i0}) \ln\left(\frac{DPN_t}{DPN_0}\right), \\ F(x,y) &\text{ 定义为:}\end{aligned}$$

$$F(x,y) = \begin{cases} \frac{a-b}{\ln a - \ln b} (a \neq b) \\ a (a = b) \end{cases} \quad (18)$$

定义影响因子权重相对大小为:

$$W(\Delta C_x) = \left| \frac{\Delta C_x}{\Delta CE} \right| \times 100\%, x \in \{c, h, n, r, t, p, DPN\} \quad (19)$$

5.2 基于 LMDI 对居住建筑生命周期碳排放进行综合评价

通过查找资料,得到南京市 2018~2022 年的二氧化碳排放量、居住建筑房地产投资、居民用电等指标因素,具体情况如下表所示。

表 4: 2018-2022 年南京市相关指标因素

年份	二氧化碳 排放量 (百万吨)	居住建筑 房地产 投资 (亿元)	新增 住宅面积 (千万平方米)	居民用电 (亿千瓦时)	总居住面积 (万平方米)	常住人口 (万人)	拆迁人数 (万户)
2018	164.9052799	1626.04	1.27	248.6	1200.6	843.75	0.62
2019	170.3911893	1754.67	1.29	257.9	1253.5	866.67	0.83
2020	178.7823898	1861.05	1.9	275.3	1316.6	931.47	0.98
2021	190.831939	1932.13	1.2	298.6	1387.4	938.74	1.13
2022	207.8234123	2052.38	1.08	331.2	1473.2	949.11	1.3

基于表格中数据,运用 LDMI 评价指标模型,得到 2019~2022 年各指标因素的影响因子权重相对大小如下:

表 5: 影响因子权重相对大小

年份	ΔCE	$W(\Delta C_c)$	$W(\Delta C_h)$	$W(\Delta C_n)$	$W(\Delta C_r)$	$W(\Delta C_t)$	$W(\Delta C_p)$	$W(\Delta C_{DPN})$
2019	5.485909	132.64%	184.89%	64.48%	19.53%	49.86%	809.47%	891.37%
2020	8.391200	22.44%	683.04%	669.66%	33.65%	47.83%	195.58%	345.58%
2021	12.049549	42.53%	762.01%	829.11%	44.25%	68.39%	206.44%	218.36%
2022	16.991473	29.21%	194.31%	245.01%	51.13%	57.47%	151.43%	164.31%

需注意,此处的百分数无实际意义,仅表示其影响因子权重的相对大小。

由此,可作出如下分析:

(1) 碳排放-居住投资转化比 (ΔC_c)

平均影响权重为 56.71%,处于碳排放影响因素指标的第 5 位。投资量与碳排放的关系对碳排放的总量影响不明显。

(2) 居住投资价值 (ΔC_h)

平均影响权重为 456.06%，处于碳排放影响因素指标的第 1 位。在居住建筑建造过程中，投资额度与新增住宅面积大体呈现正相关态势，住宅面积的新增势必涉及建材运输、工地施工等一系列加大碳排放的举动，因此影响权重较大。

(3) 新增住宅面积用电转化比 (ΔC_n)

平均影响权重为 452.61%，处于碳排放影响因素指标的第 2 位。在居住建造与使用过程中，新增面积与电力使用情况的组合对碳排放的影响权重较大。

(4) 住宅面积用电转化比 (ΔC_r)

平均影响权重为 37.14%，处于碳排放影响因素指标的第 7 位。近年来，住宅面积区域稳定，其用电量随全球变暖而逐年上升，因此影响因子为总体上升趋势，但其权重较低，趋于稳定。

(5) 人口居住密度 (ΔC_t)

平均影响权重为 55.89%，处于碳排放影响因素指标的第 6 位。直至 2020 年，人口居住密度对碳排放呈正影响。2021 年，由于国家推进生育政策，人口居住量上升，导致人口密度增大。人口居住密度影响因子基本保持在 50% 至 70%，其影响小于其他因素对碳排放的影响。

(6) 拆迁规模 (ΔC_p)

平均影响权重为 340.73%，处于碳排放影响因素指标的第 4 位。2019 年，由于国家更新拆迁补偿政策，带来了大规模拆迁活动，导致 2019 年拆迁规模对碳排放影响因子达到 809.47%。2019 年以后，拆迁规模大幅度降低，影响因子保持在 150% 至 200% 的水平，结构相对稳定。

(7) 拆迁人数 (ΔC_{DPN})

平均影响权重为 404.905%，处于碳排放影响因素指标的第 3 位。同上，拆迁人数对碳排放影响因素在因政策原因达到高峰后呈现相对稳定的变化趋势，且逐年下降。

综上所述，在居住建筑的整个生命周期中，建造需要消耗大量资源，其对碳排放影响较大；而在其使用过程中，电力的使用是影响其碳排放量的主要因素；在拆除阶段，碳排放量受当季政策影响较为明显。

6 问题三：基于改进 LMDI 模型的 2021 年江苏省地级市居住建筑碳排放综合评价

6.1 改进的 LMDI 分解模型建立

根据不同指标因素综合考虑建筑的各个生命周期，由上表可知：近年来，在多项指标中，影响因子最大的为 ΔC_h ，其次为 ΔC_n 。

而对 ΔC_{DPN} 和 ΔC_p 而言，由于其受极值影响较大（对应限时政策原因），因此其影响权重实际上低于 ΔC_h 和 ΔC_n 。

欲构建 2021 年江苏省 13 个地级市居住建筑碳排放的综合评价模型，本部分对 (2) 中 LMDI 分解模型进行改进，得到如下模型：

$$CE = \frac{CE}{NRA} \times \frac{NRA}{REC} \times REC \quad (20)$$

重新记为：

$$CE = m \cdot n \cdot REC \quad (21)$$

其中, m 为新增住宅面积平均碳排放量。

因此有:

$$\Delta CE = CE_t - CE_0 = \Delta C_m + \Delta C_n + \Delta C_{REC} \quad (22)$$

其中,

$$\Delta C_m = \sum_i F(C_{it}, C_{i0}) \ln\left(\frac{m_t}{m_0}\right),$$

$$\Delta C_n = \sum_i F(C_{it}, C_{i0}) \ln\left(\frac{n_t}{n_0}\right),$$

$$\Delta C_{REC} = \sum_i F(C_{it}, C_{i0}) \ln\left(\frac{REC_t}{REC_0}\right)。$$

定义影响因子权重相对大小为:

$$W_x = \left| \frac{\Delta C_x}{\Delta CE} \right| \times 100\%, x \in \{m, n, REC\} \quad (23)$$

6.2 江苏省地级市居住建筑碳排放综合评价

基于 2021 年江苏省 13 个地级市的碳排放数据, 应用上述 LMDI 模型, 得到江苏省 13 地级市建筑碳排放相关指标影响因子权重如下:

表 6: 江苏省地级市相关指标影响因子权重相对大小

	南京	无锡	徐州	常州	苏州	南通	连云港	淮安	盐城	扬州	镇江	泰州	宿迁
$W(\Delta C_m)$	87%	1329%	425%	152%	697%	3064%	10661%	1501%	692%	549%	11270%	1232%	584%
$W(\Delta C_n)$	36%	1088%	457%	352%	652%	2595%	9113%	1378%	680%	435%	10578%	1159%	608%
$W(\Delta C_{REC})$	151%	341%	68%	300%	145%	369%	1448%	223%	111%	214%	792%	173%	76%

需注意, 此处的百分数无实际意义, 仅表示其影响权重的相对大小。

对十三个省市的数据分析如下:

- (1) 南京地区的新增住宅较少, 可能与其可供开发的用地面积不足有关, 因此, 南京市的居住建筑建造碳排放量对其居住建筑总体碳排放量影响不大, 主要影响因素为电力消耗。
- (2) 某城市的 ΔC_m 与 ΔC_n 影响权重与 ΔC_{REC} 相对差别越大, 则说明此城市在 2021 年居住建筑建造碳排放量在总排放量中影响较大。符合此特点的典型城市为连云港与镇江。
- (3) 对其余城市而言, 建造碳排放量相关指标均处于相对高位值, 说明其居住建筑碳排放量主要体现在建造这一生命周期内, 但也不乏例外, 如南京市与常州市。

6.3 模型有效性验证

如前文所述, LMDI 模型具有残差为 0 的特点, 推导如下:

$$\begin{aligned}
\frac{V^t}{V^0} &= \exp\left(\sum_i \omega_i \ln \frac{X_{1i}^t}{X_{1i}^0}\right) \exp\left(\sum_i \omega_i \ln \frac{X_{2i}^t}{X_{2i}^0}\right) \dots \exp\left(\sum_i \omega_i \ln \frac{X_{ni}^t}{X_{ni}^0}\right) \\
&= \exp\left(\sum_i \frac{(V_t^t - V_t^0) \ln(V_t^t/V_t^0)}{(V^t - V^0) \ln(V^t/V^0)} \ln \frac{V^t}{V^0}\right) \\
&= \exp\left(\frac{\ln(V^t - V^0)}{V^t - V^0} \sum_i (V^t - V^0)\right) \\
&= \frac{V^t}{V^0}
\end{aligned} \quad (24)$$

江苏 13 个地级市的 ΔCE , ΔC_m , ΔC_n 与 ΔC_{REC} 列表如下:

经检验, 数据满足 $\Delta CE = \Delta C_m + \Delta C_n + \Delta C_{REC}$, 无残差, 说明模型有效。

表 7: 江苏地级市 ΔCE , ΔC_m , ΔC_n 与 ΔC_{REC}

	南京	无锡	徐州	常州	苏州	南通	连云港	淮安	盐城	扬州	镇江	泰州	宿迁
ΔC_m	-10.50	-40.07	-68.03	3.04	-123.89	-37.22	-18.23	-24.52	-24.46	-6.23	-55.30	-28.26	-8.21
ΔC_n	4.33	32.80	73.20	-7.05	115.96	31.52	15.59	22.51	24.07	4.94	51.90	26.59	8.54
ΔC_{REC}	18.21	10.28	10.82	6.00	25.71	4.48	2.47	3.64	3.93	2.43	3.88	3.96	1.07
ΔCE	12.04	3.01	16.00	2.00	17.77	-1.21	-0.17	1.63	3.53	1.13	0.49	2.29	1.40

7 问题四：基于 ARIMA 模型的碳排放量预测估值

7.1 ARIMA 模型构建总述

本部分以江苏省建筑全过程碳排放量为研究对象，根据 1999-2019 的江苏省建筑全过程碳排放量，预测 2023 年江苏省建筑全过程碳排放量。在对时间序列进行平稳化的基础上，按照 AIC 方法确定最优的差分整合移动平均自回归模型（ARIMA），对时间序列的规律性进行建模，并得到预测估值。

1999-2019 的江苏省建筑全过程碳排放量如下表所示：

表 8: 1999-2019 年江苏省建筑全过程碳排放量

年份	江苏省建筑全过程碳排放量（百万吨）	年份	江苏省建筑全过程碳排放量（百万吨）
1999	207.85	2010	546.29
2001	217.31	2011	613.11
2002	232.36	2012	619.99
2003	268.13	2013	638.44
2004	314.2	2014	621.07
2005	386.71	2015	634.16
2006	419.5	2016	653.12
2007	436.35	2017	645.05
2008	462.05	2018	644.97
2009	485.97	2019	636.58

7.2 基于历史数据和 ARIMA 模型的碳排放量预测

7.2.1 序列的平稳性检验与最佳参数选取

ARIMA 模型作为一种时间序列模型，要求序列具有平稳性。即时间序列的统计性质，如均值、方差、自相关系数等相对稳定，从而保证预测模型的适用性和准确性。本部分通过单位根检验方法（ADF）判断序列是否平稳 [4]。

ADF 检验参数如表所示：

ADF 检验流程如下：

- 选取差分阶数 0, 1, 2，分别进行假设显著性检验。
- 作出原假设 H_0 ：时间序列不平稳。
- 对数据进行 t 检验，通过以下公式得到碳排放样本的单总体 t 检验统计量 T 的值，该值 T 值

ADF 检验表							
变量	差分阶数	t	P	AIC	临界值		
					1%	5%	10%
碳排放量	0	-3.084	0.028**	105.013	-4.012	-3.104	-2.691
	1	0.134	0.968	106.788	-4.069	-3.127	-2.702
	2	-1.288	0.635	105.175	-4.069	-3.127	-2.702

注：***、**、* 分别代表 1%、5%、10% 的显著性水平

用于比较样本平均数与总体平均数差异是否显著。

$$T = \frac{\bar{X} - \mu_0}{\frac{s}{\sqrt{n}}} \quad (25)$$

$$\bar{x} = \frac{\sum_{i=0}^n x_i}{n} \quad (26)$$

$$s = \sqrt{\frac{\sum_{i=0}^n (x_i - \bar{x})^2}{n}} \quad (27)$$

- 取 (1%, 5%, 10%) 作为置信区间的三个边界值, 将 ADF 检验得到的 T 值与之比较。如果 T 值同时小于 (1%, 5%, 10%), 则说明可以很好地拒绝假设 H_0 , 即该序列为平稳的时间序列。

在差分为 0 阶时, 显著性 P 值为 0.028**, 水平上呈现显著性, 拒绝原假设, 该序列为平稳的时间序列。

赤池信息准则 (AIC) 可以衡量所估计模型的复杂度和此模型拟合数据的优良性。通过以下公式计算 AIC 的值可以寻找能够最好地解释数据但包含最少自由参数的模型。AIC 的值越小, 数据拟合的优良性更佳。

$$AIC = 2k - 2\ln(L) \quad (28)$$

其中, k 是参数的数量, L 是似然函数。

在 ARIMA 模型中, 时间序列主要分为季节性 (p)、趋势 (d) 和噪声 (q) 主要方面。当 p、d、q 分别取不同参数值时, ARIMA 模型对数据的拟合性能不同。通过 AIC 方法可以寻找到 p、d、q 参数的最优解, 使预测模型效果更好。经 AIC 处理, 求得 (1,1,0) 是 (p,d,q) 的最佳参数。

7.2.2 ARIMA 模型预测碳排放量估值

基于 AIC 方法确认最优参数, 构建 ARIMA(1,1,0) (基于 0 差分数据)。
模型拟合结果如下图所示:

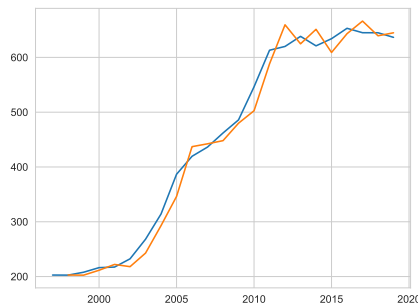


图 1: ARIMA 拟合结果

误差均值 $ERROR_{mean} = 17.372$, $R^2 = 0.984$ 代表时间序列拟合程度（越接近 1 效果越好）。残差均值分析与正态性检验如图 2、图 3 所示。

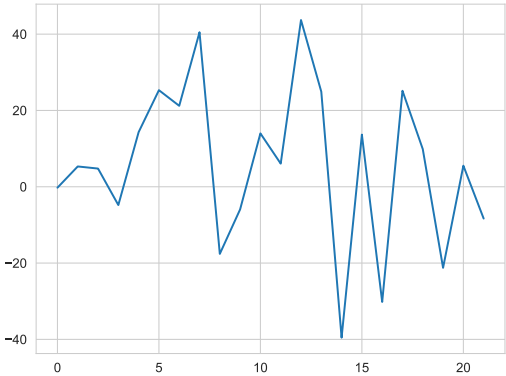


图 2: 残差均值

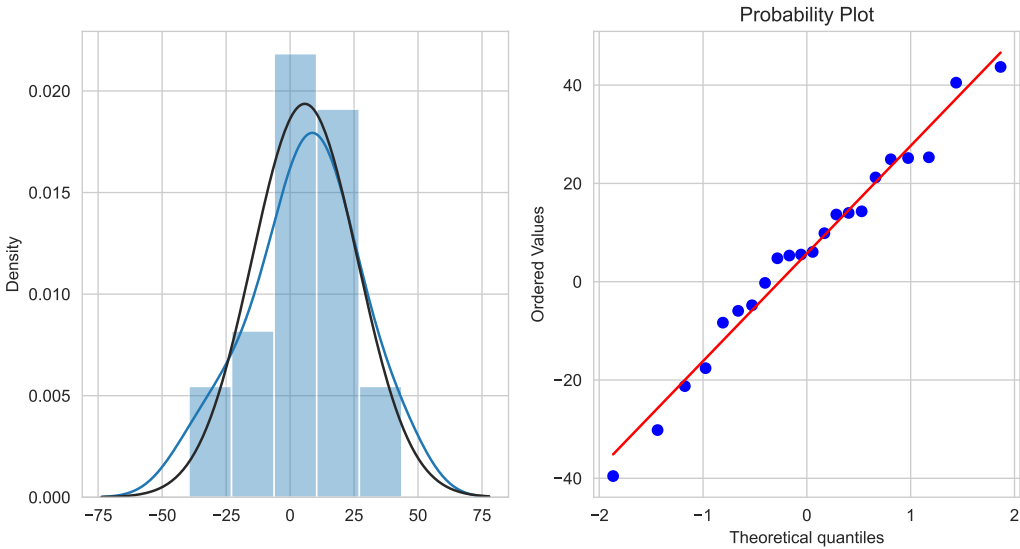


图 3: 残差正态性检验

由此可见，模型在 1999-2019 年时间序列上表现优秀。对最近 5 期数据应用模型进行预测，得到结果如下：

预测值	
阶数（年份）	预测结果（百万吨）
1（2020）	640.79
2（2021）	651.50
3（2022）	665.55
4（2023）	681.329

即 2023 年江苏省建筑全过程碳排放量预测为 $681.329 \times 28\% = 190.772$ （百万吨）。

8 问题五：江苏省建筑碳减排的政策建议

基于上述四个问题的解决过程及结果，提出如下碳减排建议：

- (1) 针对可用建筑面积较少、居住人口较多的地区（如南京、常州等），改善居住用电结构，如推行分时计价、分段计价等。
- (2) 在建筑使用过程中，政府应当推行科学用电，尽量少用空调等大功率电器，以降低电量消耗，从而实现减排。
- (3) 针对建筑施工项目较多的地区，推行环保化、低碳化施工措施，同时推进新能源运输建设，在建造全流程层面减少碳排放量。
- (4) 限时政策（如拆迁、生育等）可能会影响居住建筑的碳排放量，因此研究建筑减排时应当考虑到其他领域一系列相关政策可能带来的时间性连锁反应。

9 模型优缺点分析

9.1 模型优点

9.1.1 LMDI

第一，易于公式化。一旦指定了 IDA 标识，就可以很容易地推导出 LMDI 公式。且其结果不包含无法解释的残留项，简化了结果的解释。

第二，乘式 LMDI 给出的结果具有可加性，且乘法分解与加性分解之间存在简单的关系，使得使用乘法和加法方案的单独分解变得不必要。

第三，LMDI 在聚合上是一致的。可以将子组水平上的影响估计汇总起来，以在组水平上给出相应的影响，这一特性在多级聚合研究中很有用。

9.1.2 ARIMA

模型简单，只需要内生变量而不需要借助其他外生变量，适宜处理小规模、短期数据。

9.2 模型缺点

9.2.1 LMDI

LMDI 的缺点:LMDI 公式包含的对数项和变量不能有负值。

9.2.2 ARIMA

第一，ARIMA 要求时序数据是稳定的，或者是通过差分化后是稳定的。

第二，ARIMA 本质上只能捕捉线性关系，而不能捕捉非线性关系。采用 ARIMA 模型预测时序数据，对其稳定性要求较高。

第三，ARIMA 使用的数据量有限，在构建碳排放预测模型时仅仅使用 1999-2019 的全过程建筑碳排放量作为指标进行预测。如需进一步提高预测准确率，应当适量减小时间序列的步长，增大数据量。

10 模型推广

1. 对于建筑、能源领域，LMDI 方法是较为通用的分解方法，其以易于公式化、无残差等特点而广受青睐。本文应用该模型的过程可与机器学习、深度学习相关方法结合，进行对比实验，从而应用于更多的因素分解与评价类问题。
2. ARIMA 模型对短期、小规模时序数据的拟合与预测效果较好，但不适用于长期、大规模的数据处理。欲推广这一模型，应当引入相关性较大的生命周期中的其他指标作为辅助预测量，利用 LSTM 等方法进行预测。

参考文献

- [1] Xiwang Xiang,Xin Ma,Zhili Ma,Minda Ma and Weiguang Cai,Python-LMDI:A Tool for Index Decomposition Analysis of Building Carbon Emissions,Buildings:2022,12,83
- [2] 袁蒂, 保定市交通运输业碳排放计量分析与减排路径研究, 华北电力大学, 2012, 27-28
- [3] B.W.Ang,The LMDI approach to decomposition analysis:a practical guide,Energy Policy 33:2005 867-871
- [4] 吴孟操, 基于 ARIMA 模型的上海市碳排放权价格分析 [J]. 经济研究导刊,2022(26):29-31.

附录

LMDI 部分实现代码:

```
1 import numpy as np
2 import pandas as pd
3 import math
4
5 def F(a, b):
6     if a == b:
7         return a
8     else:
9         return (a - b) / (math.log(a) - math.log(b))
10
11 data = pd.read_excel(r"./data/problem02_data.xlsx")
12 data = data.values
13 data = data[:, 1:8]
14 emission = data.copy()
15 for i in range(0, 5):
16     for j in range(0, 6):
17         emission[i, j] = emission[i, j] / emission[i, j + 1]
18 res_LMDI = emission.copy()
19 for i in range(1, 5):
20     for j in range(0, 7):
21         res_LMDI[i, j] = F(data[i, 0], data[i - 1, 0]) *
22             math.log(emission[i, j] / emission[i - 1, j])
23 res_LMDI = res_LMDI[1:, :]
24 delta_C = []
25 for i in range(1, 5):
26     delta_C.append(data[i, 0] - data[i - 1, 0])
27 delta_C = np.asarray(delta_C)
28 weight_LMDI = res_LMDI.copy()
29 for i in range(0, 4):
30     for j in range(0, 7):
31         weight_LMDI[i, j] /= delta_C[i]
32 print(weight_LMDI)
33 print(delta_C)
34
```
