



广东外语外贸大学

应用时间序列分析期末课程论文

题目 基于 ARIMA 模型对全球气温变化预测研究

姓名 章静怡

学号 20212005202

所在学院 数学与统计学院

年级专业 经济统计学 2021

指导教师 顾朝晖 职称 讲师

完成时间 2024 年 6 月 5 日

基于 ARIMA 模型对全球气温变化预测研究

章静怡 数学与统计学院

摘要：近年来全球变暖和极端天气频发，这引起了人们对全球气候变化问题的关注。在此背景下，本文针对全球变暖现象进行了量化分析和研究，本文以伯克利地球 1785 至 2015 年的数据作为分析样例，构建 $ARIMA((1,3,4),1,5)$ 模型，结果表明，该模型可以有效地提取时间序列数据变化的潜在特征，实现较高精度的预测。从 2016 年到 2021 年，全球平均气温呈现逐渐上升的趋势。这与全球变暖的趋势吻合。本文的研究展示了 ARIMA 模型在气温预测中的应用，有助于更全面地理解和应对全球气候变化带来的挑战。

关键词：气候预测 ARIMA 全球变暖

Abstract: In recent years, global warming and the frequent occurrence of extreme weather events have drawn attention to the issue of global climate change. Against this backdrop, this paper focuses on the quantitative analysis and study of global warming phenomena. Using Berkeley Earth data from 1785 to 2015 as the analysis sample, we constructed an $ARIMA((1,3,4),1,5)$ model. The results indicate that this model can effectively extract the underlying characteristics of time series data and achieve high prediction accuracy. From 2016 to 2021, global average temperatures show a gradual upward trend, consistent with the trend of global warming. This study not only demonstrates the application of the ARIMA model in temperature prediction but also aids in a more comprehensive understanding and response to the challenges posed by global climate change.

Key words: Climate Prediction, ARIMA, Global Warming

目 录

摘 要:	1
Abstract:	1
一、引言	2
二、文献综述	3
三、模型设计	4
1.数据和样本	4
2.基于 R 的绘图方法	4
3.ARIMA 模型搭建	4
4.模型解释	Error! Bookmark not defined.
四、实证分析	5
1. 数据预处理	5
1.1 软件环境	5
1.2 操作步骤	5
2. 数据可视化	6
2.1 趋势分析	6
2.2 季节性数据分析	7
3. 模型构建	8
3.1 平稳性和随机性	8
3.2 自相关图和偏相关图	10
3.3 模型拟合	11
3.4 残差检验	12
五、结论	13
参考文献	16
附 录：原始及预测后数据	17
附 录：代码	25

一、引言

在全球气候变暖的背景下，地球气温的变化趋势及其未来预测一直是国际科学研究的焦点。许多学者通过构建气候模型探索全球气温变化的规律，尝试预测其未来走势，这不仅有助于科学界更准确地理解气候变化的影响，也为政策制定者应对气候变化提供了科学依据。

研究表明，过去一个世纪以来，全球平均气温显著上升，这种变化与人类活动密切相关。根据 IPCC（国际气候变化专门委员会）¹的报告，过去 50 年全球气温增速是前半个世纪的两倍。如果温室气体排放持续增长，21 世纪末全球平均气温可能比工业化前水平高出 3.7 至 4.8 摄氏度。这种气温变化将对全球水资源(Zhou Z et al., 2023)、食品供应(Richards et al., 2023)、健康(Zhou Z et al., 2023)以及生态系统(Wei W et al., 2024)等产生深远影响。在这种背景下，对气温的预测研究显得尤为重要。

现有学者倾向于使用机器学习和深度学习方法（刘家辉等, 2020），但这些方法往往对大量数据和复杂的特征工程依赖较大。同时，虽然卷积神经网络 (Tripathy and Prusty, 2023)和长短期记忆网络等深度学习模型在气温预测中表现出强大的数据处理和预测能力，但其训练过程复杂且计算资源消耗较大，限制了其在实践中的应用范围。

据此，本文选择了 ARIMA（自回归积分滑动平均）模型进行气温预测。ARIMA 模型的优势在于能够处理时间序列数据中的自相关特性，尤其适用于线性数据的短期预测，并且在处理季节性波动方面表现优越。相比于其他复杂的机器学习方法和深度学习方法，ARIMA 模型具有计算简单、易于实现的特点，同时在短期预测中具有较高的精度和可靠性。

本文借助伯克利地球中 1785-2015 年全球陆地平均气温，构建 ARIMA 模型，具体为 ARIMA((1,3,4),1,5)进行气温预测。该方法能够有效处理时间序列数据中的非线性趋势和季节性波动，预测结果显示 2016 年至 2020 年气温将先上升后略有下降，2021 年有所回升，与全球变暖的共识一致，显示了模型的高精度和可靠性。

¹ [IPCC — Intergovernmental Panel on Climate Change](#)

综上所述，本文的研究不仅展示了 ARIMA 模型在气温预测中的应用和验证，也提出了当前气温预测研究的一些挑战和局限性。未来的工作可以进一步探索更加准确和可靠的气温预测模型，并结合多源数据和先进技术，以更全面地理解和应对全球气候变化带来的挑战。

二、文献综述

全球气温变化是当前气候科学研究中的重要课题。准确的气温预测不仅有助于科学界理解气候变化的动态，还能够为农业(Cruz Maceín et al., 2023)、能源管理(Zhou Z et al., 2023)、灾害预防(Yang et al., 2023)等领域提供有力支持。本文旨在对近年来全球气温预测的研究方法和成果进行综述，分析其优缺点，并阐述本文所采用的 ARIMA 模型及其研究意义。

现有学者对全球气温预测的研究方法可以归纳为以下三类。首先，传统时间序列模型如 ARIMA（自回归移动平均模型）在气温预测中应用广泛。ARIMA 模型通过分析时间序列数据的自相关特性来进行预测(oung, 2017)。其优势在于模型简单，适用于线性数据，且在短期预测中表现良好。然而，ARIMA 模型的局限性在于难以捕捉数据中的非线性关系(Romilly, 2005)。

随着计算技术的发展，机器学习方法在气温预测中的应用逐渐增加。XGBoost（极限梯度提升）作为一种集成学习方法，通过集成多个决策树模型来提高预测精度（汪礼原等, 2022）。这种方法在处理非线性数据方面具有显著优势，但其预测效果高度依赖于大量数据和复杂的特征工程（刘家辉等, 2020）。此外，深度学习方法如卷积神经网络（CNN）、长短期记忆网络（LSTM）和门控循环单元网络（GRU）在气温预测中显示出强大的数据处理和预测能力。尽管这些方法在预测精度上有所提升，但其训练过程复杂且计算资源消耗较大（蒋松谕等, 2023）。

另一个重要的研究方向是分解-集合模型。分解-集合模型通过将复杂时间序列分解为若干子序列，再对每个子序列进行预测并集合结果。CEEMDAN-WOA-TCN 模型利用自适应噪声完备集合经验模态分解（CEEMDAN）和鲸鱼优化算法（WOA）优化的时序卷积网络（TCN）进行气温预测(Tripathy and

Prusty, 2023)。这种方法在捕捉数据的非平稳和非线性特性方面具有优势，但模型复杂性较高，参数选择和模型组合存在较高难度。

尽管机器学习和深度学习等方法具有较高的预测精度，但其黑箱性质使得结果难以解释，给实际应用和决策支持带来挑战。此外，深度学习和集成学习方法通常需要大量的训练数据和高计算资源，这在数据匮乏或资源有限的情况下限制了其应用。

基于以上考虑，本文采用 ARIMA 模型，并利用 1785-2015 年的气温数据进行研究。ARIMA 模型通过分析时间序列的自相关特性，能够有效地对历史数据进行建模和预测。利用长时间跨度的数据，揭示气温变化的长期趋势特征，从而提高气温预测的准确性和可靠性。相比于高端的机器学习和深度学习模型，ARIMA 模型的价值在于其透明性和可解释性，使得预测结果更易于理解和应用。此外，ARIMA 模型对计算资源的要求较低，适用于资源有限的研究环境。

三、模型设计

1. 数据和样本

本文选取了 1785 年 1 月至 2015 年 12 月的伯克利地球数据（Berkeley Earth Data）²，以探究全球气温的季节性变化以及长期趋势。数据筛选与提取工作由 Python 3.10 完成，用于提取日期和全球平均气温值。随后，通过 ARIMA 模型对年度数据进行拟合和预测，从而为全球气温的未来变化提供科学依据。

2. 基于 R 的绘图方法

本研究基于 Windows 系统，采用 R 4.3.1 进行数据可视化。具体绘图过程使用了 R 语言的基础包`ggplot2`和`plotly`函数。`ggplot2`用于静态图形绘制，`plotly`则用于交互式图形分析。通过这两种工具的结合，实现了数据结果的多维展示和互动分析，提升了结果的可视化效果。

3. ARIMA 模型搭建

² [Data Overview - Berkeley Earth](#)

ARIMA 模型，即自回归积分滑动平均模型，是一种结合自回归（Auto-Regressive, AR）、差分（Differencing, I）和滑动平均（Moving Average, MA）过程的时间序列分析方法。其数学表达式为：

$$Y_t = \mu + \sum_{i=1}^p \varphi_i Y_{t-i} + \sum_{j=1}^q \theta_j \varepsilon_{t-j} + \varepsilon_t$$

其中： Y_t 是当前值， μ 是常数项， φ_i 是自回归系数， ε_t 是随机误差项， θ_j 是滑动平均系数。

4. 模型解释

p （自回归系数）表示序列值滞后 p 阶的自相关程度。 d （差分次数）表示使非平稳时间序列变成平稳序列所需的最少差分次数。 q （滑动平均系数）表示误差项滞后 q 阶的移动平均程度。

对时间序列数据进行自相关图（ACF）和偏自相关图（PACF）的平稳性检验。若序列为非平稳序列，则需要对其进行差分处理，并再次进行平稳性检验，直到满足平稳性条件。满足条件的差分次数即为 d 值。

通过上述方法确定 ARIMA 模型的参数后，采用 SAS 中的 `ARIMA` 库进行模型拟合和验证。

四、实证分析

1. 数据预处理

1.1 软件环境

本文的数据预处理使用 Python 3.10，通过 pandas 库进行数据操作。所有工作在 Windows 系统下使用 PyCharm 集成开发环境（IDE）完成。

1.2 操作步骤

首先，导入必要的软件包并设置环境，确保脚本正确运行，指定原始数据文件的路径以便正确解析。导入全球气温数据后，将日期列（dt）转换为日期

时间格式，以保证时间序列分析的准确性。随后，提取所需的日期和陆地平均气温数据，并保存为新的 CSV 文件“Transformed_GlobalTemperatures.csv”。

具体如表 1 所示

表 1 处理后示例数据

dt	LandAverageTemperature
1785-01-01	2.515
1785-02-01	4.08
1785-03-01	5.724
1785-04-01	13.732
1785-05-01	15.757
1785-06-01	14.747
1785-07-01	16.183
1785-08-01	12.947
1785-09-01	8.844
1785-10-01	9.376

2. 数据可视化

2.1 趋势分析

本文通过全球气温数据集中提取年份信息。并按年份对数据进行分组，计算每年的平均气温及其不确定性。通过计算不确定性范围的上下界，绘制年度平均气温及其不确定性范围的折线图（如图 1 所示）。

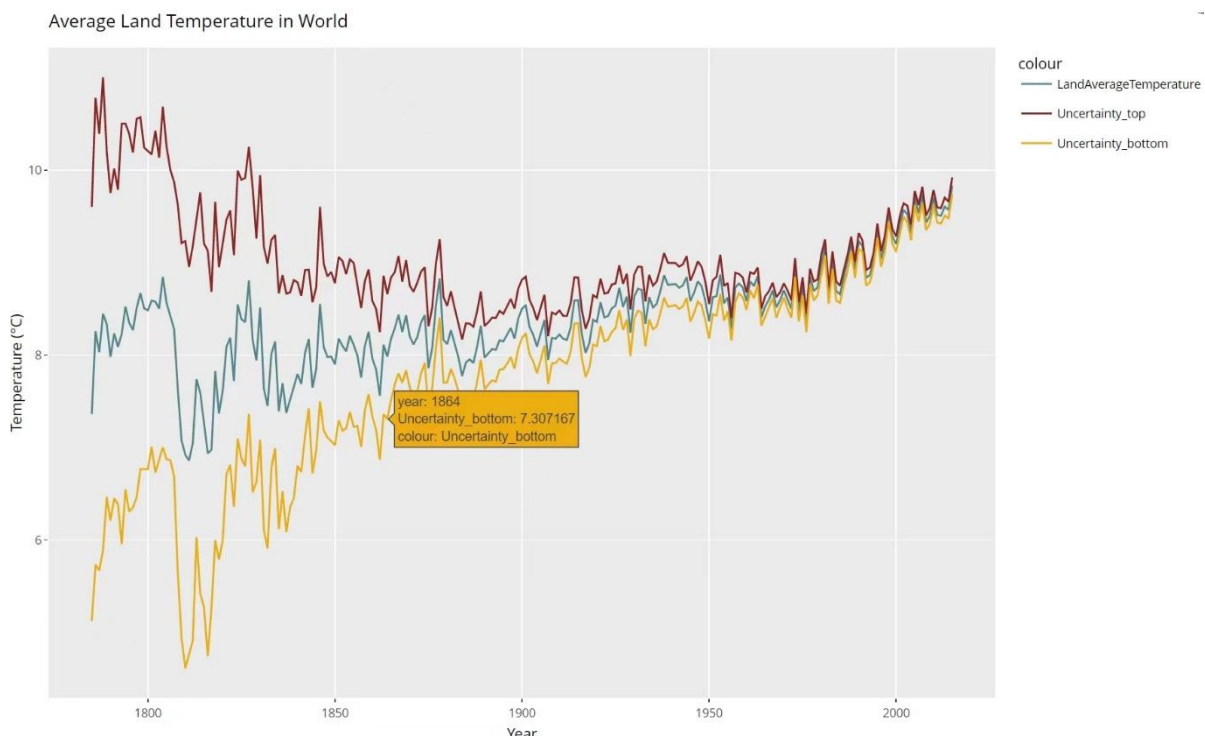
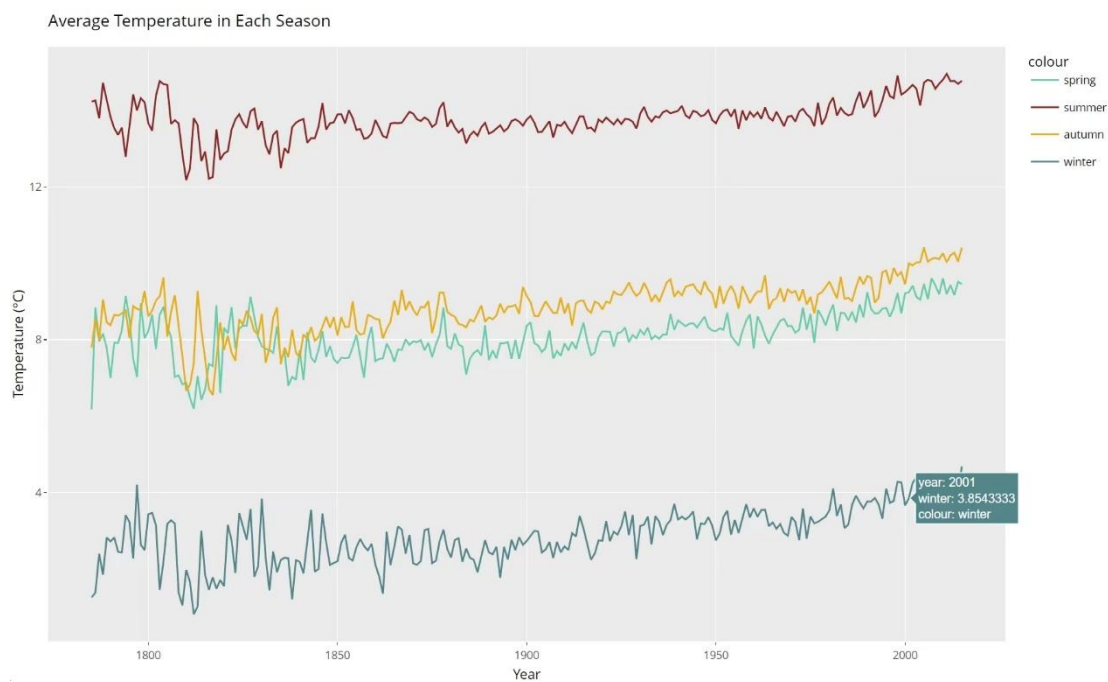


图 1 全球年度平均气温及其不确定性（1785-2015）

图 1 展示了全球气温在过去两个多世纪中的变化趋势及其不确定性范围，表明了气温持续上升的事实，并揭示了气候变化的长期影响。通过对不确定性范围的分析，可以看出，尽管数据存在一定的误差和波动，但总体趋势仍具有较高的可信度。

2.2 季节性数据分析



图二：季节性平均气温变化（1785-2015）

为进一步分析季节性温度变化，从数据集中提取月份信息，并根据月份划分季节（春季：3-5月，夏季：6-8月，秋季：9-11月，冬季：12-2月）。随后，按年份和季节分组，计算每个季节的平均气温。通过绘制季节性平均气温的折线图（具体如图2所示），可以清晰地看到不同季节的气温变化趋势。

据图2可知，各季节气温均呈现上升趋势，特别是在20世纪末，气温上升速度显著加快，其中夏季和冬季气温的上升幅度较大。通过对各季节气温变化的分析，可以看出全球变暖对各季节的影响具有一致性和显著性。

3. 模型构建

3.1 平稳性和随机性

由于月度数据量过于庞大，为简化计算过程和提升模型拟合效率，本研究在使用SAS软件进行ARIMA模型拟合时，将数据转化为年度数据。这一转换不仅减少了数据量，同时也保留了长期趋势和季节性变化的关键信息，确保了模型的准确性和可靠性。

根据原始历史的数据画出的序列图可以看出全球气温呈现上涨趋势，初步判断为不平稳序列，如图3所示。

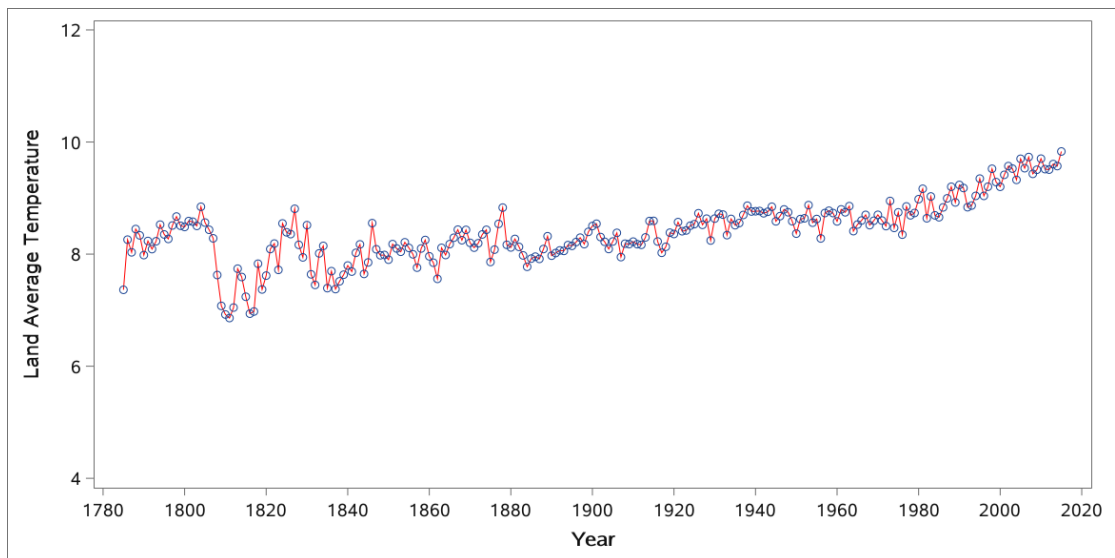


图3 原时序图

故对原序列进行一阶差分，差分后结果如图 4 所示。一阶差分后去除全序列的趋势性，基本可以认定为平稳序列。

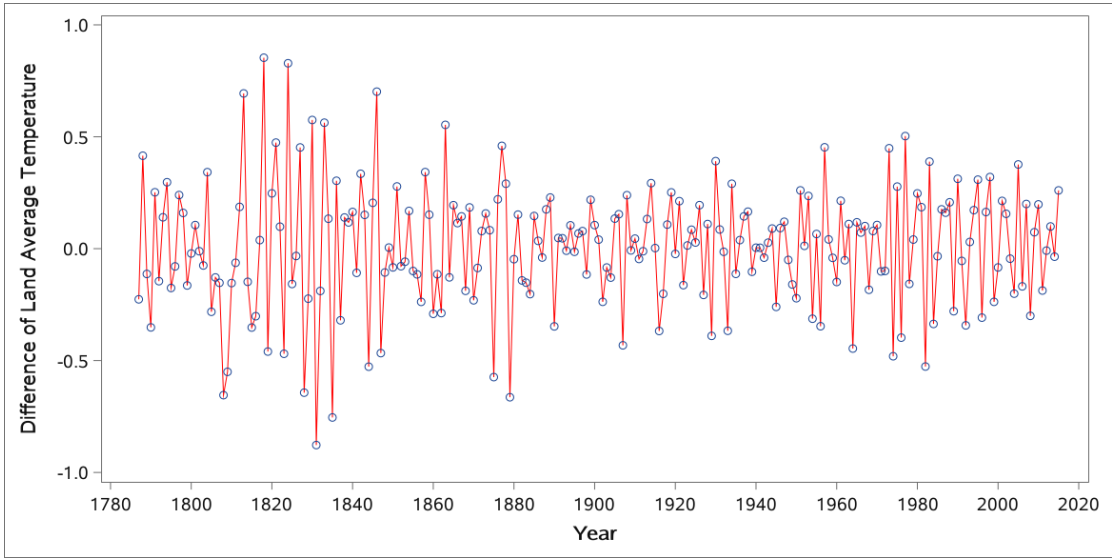


图 4 差分后的时序图

进一步使用 ADF 检验， 检验该序列的平稳性， 结果如下表所示：

表 2 单位根检验结果表

增广 Dickey-Fuller 单位根检验							
类型	滞后	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-294.604	0.0001	-20.75	<.0001		
	1	-483.606	0.0001	-15.64	<.0001		
	2	-718.283	0.0001	-11.94	<.0001		
Single Mean	0	-294.879	0.0001	-20.72	<.0001	214.75	0.0010
	1	-485.979	0.0001	-15.64	<.0001	122.26	0.0010
	2	-729.582	0.0001	-11.94	<.0001	71.27	0.0010
Trend	0	-294.910	0.0001	-20.70	<.0001	214.27	0.0010
	1	-486.951	0.0001	-15.63	<.0001	122.17	0.0010
	2	-739.617	0.0001	-11.97	<.0001	71.71	0.0010

表 3 纯随机性检验结果表

白噪声的自相关检查									
至滞后	卡方	自由度	Pr > 卡方	自相关					
6	32.74	6	<.0001	-0.287	-0.141	0.066	0.046	-0.132	0.117
12	45.49	12	<.0001	-0.066	0.027	0.046	-0.044	-0.156	0.137
18	50.76	18	<.0001	-0.030	-0.086	0.004	0.023	-0.096	0.056
24	56.76	24	0.0002	-0.064	0.107	-0.027	0.038	-0.075	-0.003

结果显示（表 2），p 值均小于 0.05，故可以认为是平稳序列。纯随机检验结果（表 3）显示该序列为非白噪声序列，故可以使用 ARIMA 模型拟合该序列。

3.2 自相关图和偏相关图

下面考察该序列的自相关图和偏相关图的特征，给 ARIMA 模型定阶。

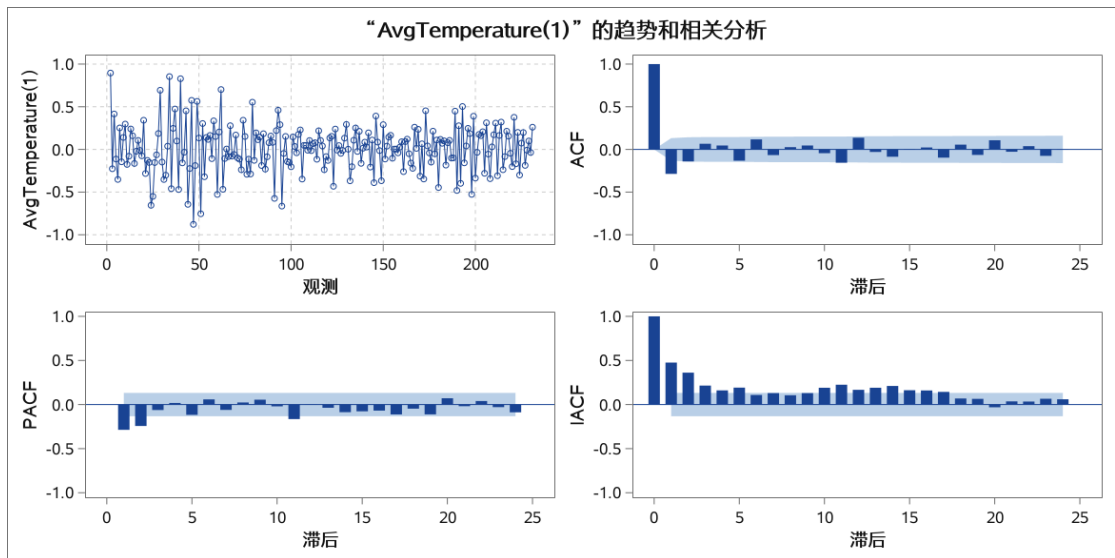


图 5 “AvgTemperature（1）” 的趋势和相关系数图

由图 5 可知，原始序列进行一阶差分后 ACF、PACF 图均为不规则衰减拖尾，故本文采用 ARIMA 模型来进行预测。

3.3 模型拟合

表 4 最小信息准则结果

滞后	最小信息准则										
	MA 0	MA 1	MA 2	MA 3	MA 4	MA 5	MA 6	MA 7	MA 8	MA 9	MA 10
AR 0	-2.65474	-2.76737	-2.77783	-2.76081	-2.78975	-2.85289	-2.902	-2.93742	-2.93319	-2.92311	-2.91019
AR 1	-2.71793	-2.74876	-2.76327	-2.74402	-2.7755	-2.84037	-2.88266	-2.91562	-2.91127	-2.90157	-2.8887
AR 2	-2.7668	-2.7526	-2.73977	-2.72132	-2.75192	-2.8175	-2.85958	-2.89252	-2.88791	-2.87859	-2.8662
AR 3	-2.75608	-2.73638	-2.72091	-2.69907	-2.73674	-2.79916	-2.83945	-2.87181	-2.86962	-2.85866	-2.84533
AR 4	-2.78264	-2.76413	-2.76798	-2.75047	-2.72706	-2.77626	-2.82112	-2.8536	-2.85133	-2.84095	-2.83262
AR 5	-2.83913	-2.82479	-2.82469	-2.80234	-2.78171	-2.75966	-2.80088	-2.83359	-2.83277	-2.8218	-2.82083
AR 6	-2.86908	-2.87217	-2.87207	-2.85044	-2.827	-2.80451	-2.78233	-2.81018	-2.81154	-2.8003	-2.79864
AR 7	-2.88109	-2.90772	-2.90412	-2.88426	-2.86068	-2.84392	-2.8219	-2.80798	-2.81456	-2.80887	-2.80007
AR 8	-2.86205	-2.89954	-2.89812	-2.88258	-2.85907	-2.84224	-2.81909	-2.81175	-2.79143	-2.78526	-2.78291
AR 9	-2.87019	-2.89566	-2.89386	-2.8728	-2.84969	-2.8344	-2.8112	-2.80096	-2.78257	-2.76767	-2.7594
AR 10	-2.85467	-2.88468	-2.88247	-2.85908	-2.83695	-2.8393	-2.8157	-2.80926	-2.78728	-2.76762	-2.74587

误差序列模型: AR(19)

最小表值: BIC(0,7) = -2.93742

表 5 ARIMA((1,3,4),1,5)模型结果

条件最小二乘估计					
参数	估计	标准 误差	t 值	近似 Pr > t	滞后
MU	0.0084136	0.0016834	5.00	<.0001	0
MA1,1	-0.62431	0.17053	-3.66	0.0003	1
MA1,2	0.52026	0.10305	5.05	<.0001	2
MA1,3	0.91674	0.19951	4.60	<.0001	3
MA1,4	0.41073	0.10374	3.96	0.0001	4
MA1,5	-0.22342	0.11193	-2.00	0.0472	5
AR1,1	-1.01509	0.15200	-6.68	<.0001	1
AR1,2	0.80841	0.16394	4.93	<.0001	3
AR1,3	0.73222	0.10847	6.75	<.0001	4

根据最小信息准则结果表 4，得知，ARIMA(0,1,7)为最适宜模型，但在模型实际拟合过程中出现残差白噪音检验不通过的情况，最终确定最优模型为带均值项的疏系数模型 ARIMA((1,3,4),1,5)，参数检验结果如表 5 所示模型拟合结果为：

$$\nabla x_t = 0.0084136 + \frac{1 - 0.62431B + 0.52026B^2 + 0.91674B^3 + 0.41073B^4 - 0.22342B^5}{1 + 1.01509B - 0.80841B^3 - 0.73222B^4} \varepsilon_t$$

3.4 残差检验

表 6 残差的自相关系数检验

残差的自相关检查									
至滞后	卡方	自由度	Pr > 卡方	自相关					
6	.	0	.	-0.002	-0.051	0.002	0.017	0.025	-0.002
12	5.96	4	0.2024	0.015	0.074	0.001	-0.082	-0.093	0.015
18	9.79	10	0.4593	-0.001	-0.096	-0.048	-0.027	-0.054	-0.018
24	14.98	16	0.5264	-0.013	0.122	-0.010	0.030	-0.055	-0.036
30	21.97	22	0.4616	0.107	0.021	0.108	0.017	-0.009	-0.053
36	27.26	28	0.5040	-0.070	0.078	0.046	-0.064	-0.032	0.037
42	28.37	34	0.7398	-0.046	-0.005	0.027	-0.033	0.003	-0.007

表 6 展示了残差的自相关性检验结果。该图表中列出了不同滞后期下的卡方统计量、自由度、p 值以及相应的自相关系数。图表中卡方统计量随着滞后期的增加而增加，但都未达到显著水平。在所有滞后期下，p 值均显著大于 0.1。这表明在 95%的置信水平下，不能拒绝残差是白噪声的假设，即残差没有显著的自相关性。

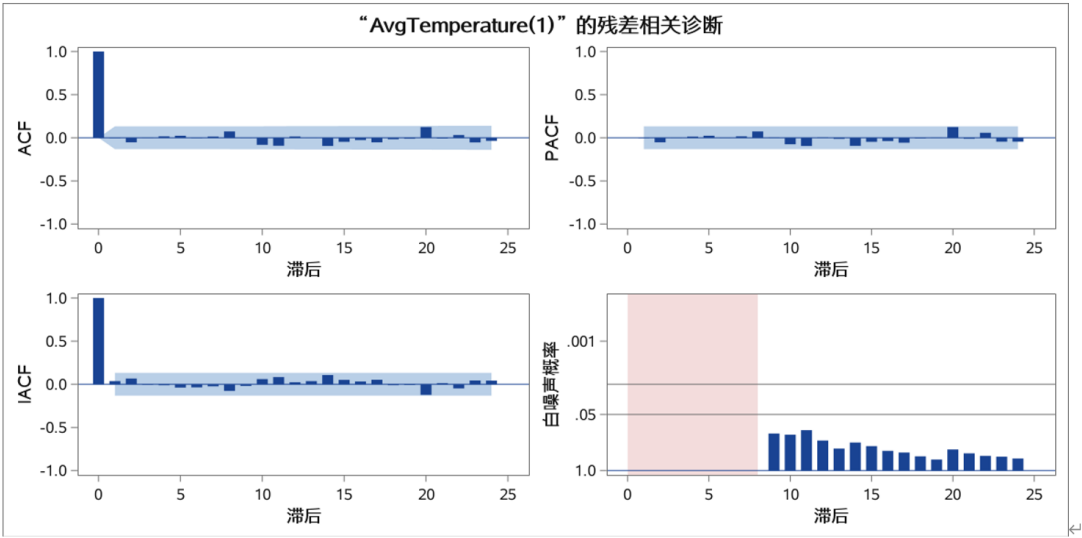


图 6 “AvgTemperature（1）” 的残差相关诊断图

从图 6 中可以看到，ACF 在大部分滞后期内都接近于零，并且大部分点落在置信区间内。进一步表明残差没有显著的自相关性。PACF 在大部分滞后期内也接近于零，并且大部分点落在置信区间内。图中显示标准化残差在零附近波动，没有明显的趋势或模式。

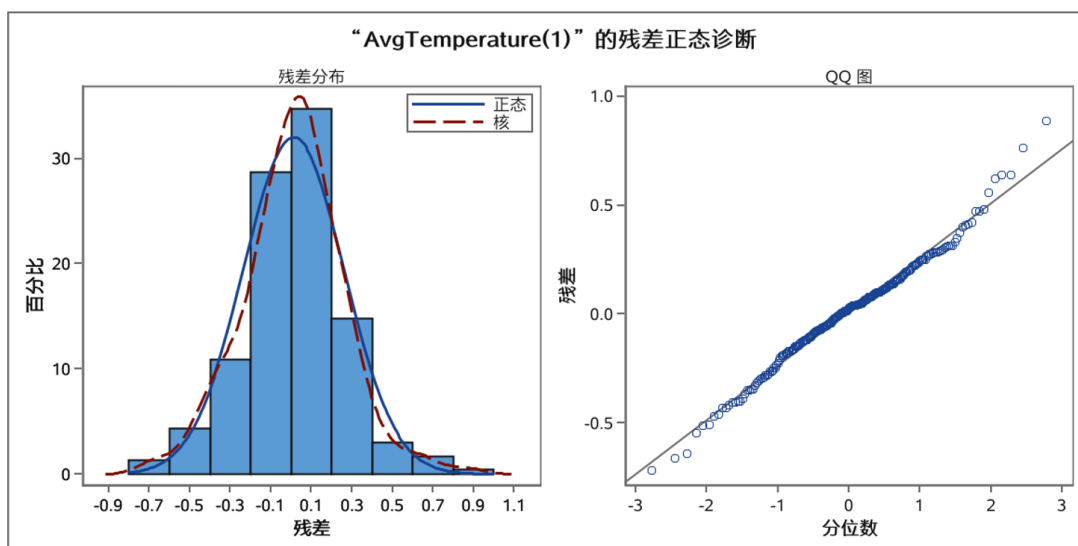


图 7 “AvgTemperature (1)” 的残差正态诊断图

由图 7 中左侧的残差分布直方图显示了残差的频率分布情况。图中蓝色实线代表正态分布，红色虚线表示核密度估计。可以看出，残差大致呈现正态分布，中心位置较为集中，两侧对称性良好，基本符合正态分布的特点。这表明残差没有明显的偏态或峰态，符合模型残差应为随机分布的假设。图 7 中右侧的残差 QQ 图将实际残差与正态分布的理论分位数进行比较。点分布较为接近对角线，说明残差的实际分布与正态分布高度吻合。若残差点均匀分布在对角线附近，则表明残差近似于正态分布。图中，绝大部分点都在对角线附近，进一步验证了残差的正态性假设。

综合以上分析，残差状态基本符合正态性假设，且没有显著的自相关性，表明模型拟合情况良好。残差的正态分布和随机性表明模型对数据的解释力较强，没有显著的系统性偏差，为模型的预测能力提供了良好的基础，保证了预测结果的可靠性和有效性。

五、 结论

采用 ARIMA((1,3,4),1,5)模型对 2016 年至 2021 年全球气温变化进行短期预测，预测结果如表 7 所示。

表 7 陆地平均气温预测表（2016-2021 年）

变量“AvgTemperature”的预测				
观测	预测	标准误差	95% 置信限	
232	9.7213	0.2538	9.2239	10.2187
233	9.6812	0.2972	9.0988	10.2637
234	9.6337	0.3217	9.0032	10.2643
235	9.6767	0.3461	8.9984	10.3550
236	9.5898	0.3678	8.8689	10.3108
237	9.6143	0.3778	8.8737	10.3548

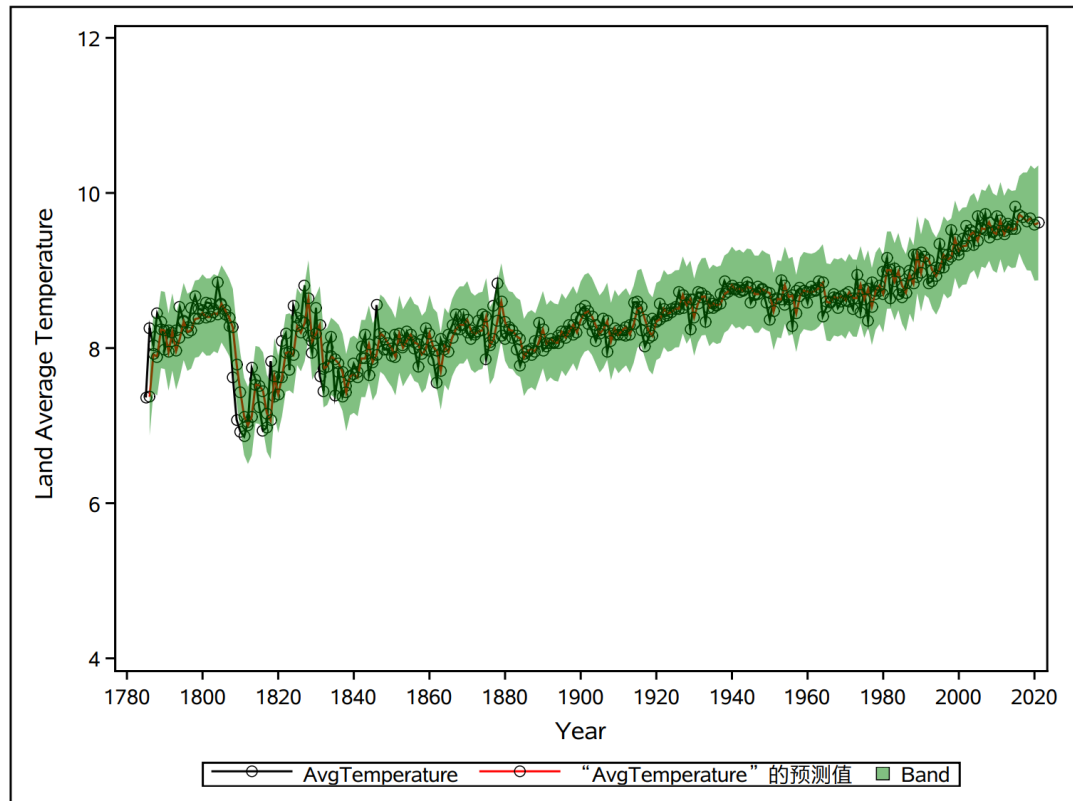


图 8 陆地平均气温预测图（1785-2021 年）

据表 7 数据显示，从 2016 年到 2021 年，全球平均气温呈现逐渐上升的趋势。尽管在 2016 年至 2020 年间预测值有小幅下降的波动，整体趋势在 2021 年开始再次上升。在这六年中，气温变化范围虽然较小，但总体上呈现逐步回升的迹象。图 8 进一步展示了从 1785 年到 2021 年的全球平均气温预测图，从而进一步验证了这一趋势。

鉴于全球气温持续上升，建议各国政府加强气候变化应对措施，制定并实施更为严格的碳排放控制政策，减少温室气体排放。各国应加强气候变化适应

能力建设，特别是对于易受气候变化影响的地区，应提前制定应对极端天气的措施，提升基础设施的抗灾能力。此外，应大力推动绿色技术的发展，鼓励使用清洁能源，如风能、太阳能等替代传统的化石燃料，以减缓气温上升的趋势。同时，加强公众的气候变化教育，提高公众的环保意识和参与度，鼓励个人和社区采取低碳生活方式，共同应对气候变化。考虑到气候变化是全球性问题，需要国际社会的共同努力和合作，各国应在气候变化问题上加强合作，共同推动全球气候治理机制的完善和实施。

该模型在短期内能够较准确地预测气温变化，预测值与实际观测值接近，残差较小。残差分析表明，残差分布近似正态，且残差的自相关性检验显示无显著自相关性，说明模型的拟合效果较好。通过 Ljung-Box Q 检验和 ACF 图的分析，进一步确认残差没有显著的自相关性，标准化残差图显示残差在零附近波动，无明显趋势或模式，进一步验证了模型的合理性。

为了提高模型的预测精度和适应性，在未来的研究中可以引入更多外部影响因素，如二氧化碳浓度、海洋温度等，以提高对长期气温趋势的预测能力。同时，可采用多种预测模型进行对比分析，以验证结果的一致性和可靠性。在实际分析过程中，发现气候系统具有复杂性，在后续研究中应整合更多的外部环境因素，如太阳活动、火山喷发等，以提高模型对复杂气温变化的适应能力。

参考文献

- [1] 蒋松谕,何贞铭,周再文.基于分解-集合的 TCN 日平均气温预测[J].计算机时代,2023,(10):45-49.
- [2] 刘家辉,梅平,刘长征,等.基于 GRU-CNN 模型的云南地区短期气温预测[J].计算机仿真,2023,40(09):472-476.
- [3] 李哲远,张成鹏.基于 ARIMA 模型对山东省生猪出场价格变动的预测[J].当代农村财经,2024,(05):34-39.
- [4] 汪礼原,李全,龚莹洁,等.基于机器学习和时间因子的气温变化模型研究[J].智能计算机与应用,2024,14(03):187-191.
- [5] 王源昊.基于 ARIMA 模型和 LSTM 神经网络的全球气温预测分析[J].科学技术创新,2021,(35):166-170.
- [6] 严迅,铁承城,鄢薇,等.基于 ARIMA 模型和 CNN-LSTM 组合模型的全球气温预测分析[J].科技与创新,2024,(02):19-22.
- [7] 张明玉,周钰博,吴建军,等.ARIMA 模型与 GM(1,1)模型对丙型病毒性肝炎发病的预测效果比较研究[J/OL].预防医学情报杂志,1-6[2024-06-10].
- [8] Cruz Maceín, J. L., Gonzalez-Fernandez, I., Barrutietta, A., Bermejo-Bermejo, V., & Zamorano Rodríguez, J. P. (2023). Adaptation strategies for dealing with global atmospheric change in Mediterranean agriculture: a triple helix approach to the Spanish case study. *Regional Environmental Change*, 23(4), 142.
- [9] Richards C E, Gauch H L, Allwood J M. International risk of food insecurity and mass mortality in a runaway global warming scenario[J]. *Futures*, 2023, 150: 103173.
- [10] Romilly P. Time series modelling of global mean temperature for managerial decision-making[J]. *Journal of environmental management*, 2005, 76(1): 61-70.
- [11] Tripathy D S, Prusty B R. Quantile regression averaging - based probabilistic forecasting of daily ambient temperature[J]. *International Journal of Numerical Modelling: Electronic Networks, Devices and Fields*, 2021, 34(3): e2846.
- [12] Wei W, Zou S, Duan W, et al. Increased population exposures to extreme precipitation in Central Asia under 1.5°C and 2°C global warming scenarios[J]. *Geography and Sustainability*, 2024.
- [13] Yang, J., Huo, Z., Wang, P., Wu, D., Ma, Y., Yao, S., & Dong, H. (2021). Process-based indicators for timely identification of apricot frost disaster on the warm temperate zone, China. *Theoretical and Applied Climatology*, 146(3), 1143-1155.
- [14] Young P C. Data-based mechanistic modelling and forecasting globally averaged surface temperature[J]. *International Journal of Forecasting*, 2018, 34(2): 314-335.
- [15] Zhou Z, Nguyen-Xuan T, Liao H, et al. Characterization of temperature and humidity effects on extreme heat stress under global warming and urban growth in the Pearl and Yangtze River Deltas of China[J]. *Weather and Climate Extremes*, 2024, 44: 100659.

附录：原始及预测后数据³

观测	Year	AvgTemperature	FORECAST	STD	L95	U95	RESIDUAL
1	1785	7.36300
2	1786	8.25817	7.37141	0.25378	6.87401	7.8688	0.88675
3	1787	8.03183	7.92005	0.25378	7.42264	8.4175	0.11178
4	1788	8.44717	7.88698	0.25378	7.38958	8.3844	0.56018
5	1789	8.33417	8.23803	0.25378	7.74062	8.7354	0.09614
6	1790	7.98233	8.22724	0.25378	7.72983	8.7246	-0.24491
7	1791	8.23425	7.94925	0.25378	7.45185	8.4467	0.28500
8	1792	8.08900	8.20738	0.25378	7.70997	8.7048	-0.11838
9	1793	8.22917	7.96128	0.25378	7.46387	8.4587	0.26789
10	1794	8.52600	8.12655	0.25378	7.62914	8.6240	0.39945
11	1795	8.35033	8.34247	0.25378	7.84507	8.8399	0.00786
12	1796	8.27100	8.20340	0.25378	7.70599	8.7008	0.06760
13	1797	8.51058	8.23356	0.25378	7.73615	8.7310	0.27703
14	1798	8.67025	8.37307	0.25378	7.87567	8.8705	0.29718
15	1799	8.50575	8.38486	0.25378	7.88745	8.8823	0.12089
16	1800	8.48425	8.45321	0.25378	7.95580	8.9506	0.03104
17	1801	8.58967	8.39994	0.25378	7.90253	8.8973	0.18973
18	1802	8.57817	8.40189	0.25378	7.90448	8.8993	0.17628
19	1803	8.50300	8.45563	0.25378	7.95823	8.9530	0.04737
20	1804	8.84483	8.43096	0.25378	7.93355	8.9284	0.41387
21	1805	8.56275	8.57087	0.25378	8.07346	9.0683	-0.00812
22	1806	8.43408	8.49007	0.25378	7.99266	8.9875	-0.05599
23	1807	8.28142	8.39977	0.25378	7.90237	8.8972	-0.11836
24	1808	7.62683	8.26591	0.25378	7.76850	8.7633	-0.63908
25	1809	7.07642	7.79445	0.25378	7.29704	8.2919	-0.71804
26	1810	6.92267	7.43540	0.25378	6.93799	7.9328	-0.51273
27	1811	6.85950	7.11721	0.25378	6.61980	7.6146	-0.25771
28	1812	7.04625	7.00351	0.25378	6.50611	7.5009	0.04274

³ 由于月度数据共 2760 条，为节省篇幅，为此处为经 python 预处理后，后经 SAS 代码自动计算后的年度数据，共计 230 条。表中第 2，3 列为原始数据，剩余为预测后的数据，分别为预测值、标准差、上限、下限和残差。

观测	Year	AvgTemperature	FORECAST	STD	L95	U95	RESIDUAL
29	1813	7.73992	7.11629	0.25378	6.61888	7.6137	0.62363
30	1814	7.59175	7.52965	0.25378	7.03224	8.0271	0.06210
31	1815	7.23933	7.51730	0.25378	7.01989	8.0147	-0.27796
32	1816	6.93750	7.44589	0.25378	6.94848	7.9433	-0.50839
33	1817	6.97567	7.15971	0.25378	6.66231	7.6571	-0.18405
34	1818	7.82892	7.06577	0.25378	6.56836	7.5632	0.76315
35	1819	7.36917	7.63102	0.25378	7.13362	8.1284	-0.26186
36	1820	7.61625	7.40461	0.25378	6.90720	7.9020	0.21164
37	1821	8.09008	7.61791	0.25378	7.12050	8.1153	0.47217
38	1822	8.18833	7.93635	0.25378	7.43894	8.4338	0.25198
39	1823	7.71867	7.95140	0.25378	7.45400	8.4488	-0.23274
40	1824	8.54717	7.90870	0.25378	7.41129	8.4061	0.63847
41	1825	8.38933	8.27856	0.25378	7.78115	8.7760	0.11077
42	1826	8.35658	8.19814	0.25378	7.70073	8.6956	0.15844
43	1827	8.80875	8.32755	0.25378	7.83015	8.8250	0.48120
44	1828	8.16533	8.63500	0.25378	8.13759	9.1324	-0.46966
45	1829	7.94183	8.08875	0.25378	7.59134	8.5862	-0.14692
46	1830	8.51675	8.18542	0.25378	7.68801	8.6828	0.33133
47	1831	7.63917	8.29969	0.25378	7.80228	8.7971	-0.66053
48	1832	7.45017	7.73254	0.25378	7.23513	8.2299	-0.28237
49	1833	8.01300	7.76615	0.25378	7.26874	8.2636	0.24685
50	1834	8.14675	7.89482	0.25378	7.39741	8.3922	0.25193
51	1835	7.39250	7.85265	0.25378	7.35524	8.3501	-0.46015
52	1836	7.69583	7.80250	0.25378	7.30509	8.2999	-0.10666
53	1837	7.37567	7.68952	0.25378	7.19212	8.1869	-0.31386
54	1838	7.51492	7.42591	0.25378	6.92850	7.9233	0.08901
55	1839	7.63225	7.63242	0.25378	7.13501	8.1298	-0.00017
56	1840	7.79625	7.66273	0.25378	7.16533	8.1601	0.13352
57	1841	7.68817	7.61883	0.25378	7.12142	8.1162	0.06933
58	1842	8.02275	7.86599	0.25378	7.36858	8.3634	0.15676
59	1843	8.17367	7.86496	0.25378	7.36755	8.3624	0.30871
60	1844	7.64633	8.04991	0.25378	7.55250	8.5473	-0.40357
61	1845	7.85117	7.82204	0.25378	7.32463	8.3194	0.02913

观测	Year	AvgTemperature	FORECAST	STD	L95	U95	RESIDUAL
62	1846	8.55275	7.91047	0.25378	7.41306	8.4079	0.64228
63	1847	8.08600	8.19280	0.25378	7.69539	8.6902	-0.10680
64	1848	7.97950	8.15045	0.25378	7.65305	8.6479	-0.17095
65	1849	7.98442	8.06664	0.25378	7.56924	8.5641	-0.08223
66	1850	7.90067	7.99802	0.25378	7.50061	8.4954	-0.09735
67	1851	8.17858	7.88790	0.25378	7.39049	8.3853	0.29068
68	1852	8.10017	8.18032	0.25378	7.68291	8.6777	-0.08015
69	1853	8.04183	8.00321	0.25378	7.50580	8.5006	0.03862
70	1854	8.21050	8.08933	0.25378	7.59192	8.5867	0.12117
71	1855	8.11075	8.17127	0.25378	7.67386	8.6687	-0.06052
72	1856	7.99592	8.07306	0.25378	7.57565	8.5705	-0.07714
73	1857	7.75817	8.04859	0.25378	7.55118	8.5460	-0.29042
74	1858	8.10025	7.91952	0.25378	7.42211	8.4169	0.18073
75	1859	8.25200	7.97770	0.25378	7.48029	8.4751	0.27430
76	1860	7.96083	8.18730	0.25378	7.68989	8.6847	-0.22646
77	1861	7.84650	8.01512	0.25378	7.51772	8.5125	-0.16862
78	1862	7.55858	7.96167	0.25378	7.46426	8.4591	-0.40309
79	1863	8.11183	7.70197	0.25378	7.20456	8.1994	0.40986
80	1864	7.98358	8.02308	0.25378	7.52567	8.5205	-0.03949
81	1865	8.17783	7.95159	0.25378	7.45418	8.4490	0.22625
82	1866	8.29217	8.13502	0.25378	7.63761	8.6324	0.15715
83	1867	8.43633	8.23972	0.25378	7.74232	8.7371	0.19661
84	1868	8.24792	8.29848	0.25378	7.80107	8.7959	-0.05057
85	1869	8.43208	8.29816	0.25378	7.80075	8.7956	0.13392
86	1870	8.20133	8.36507	0.25378	7.86766	8.8625	-0.16374
87	1871	8.11508	8.22162	0.25378	7.72421	8.7190	-0.10653
88	1872	8.19383	8.17815	0.25378	7.68074	8.6756	0.01569
89	1873	8.35108	8.21522	0.25378	7.71781	8.7126	0.13587
90	1874	8.43350	8.22827	0.25378	7.73086	8.7257	0.20523
91	1875	7.85958	8.40458	0.25378	7.90717	8.9020	-0.54499
92	1876	8.08008	8.02912	0.25378	7.53171	8.5265	0.05096
93	1877	8.53958	8.11692	0.25378	7.61951	8.6143	0.42266
94	1878	8.82975	8.35656	0.25378	7.85915	8.8540	0.47319

观测	Year	AvgTemperature	FORECAST	STD	L95	U95	RESIDUAL
95	1879	8.16583	8.59572	0.25378	8.09831	9.0931	-0.42989
96	1880	8.11875	8.33195	0.25378	7.83454	8.8294	-0.21320
97	1881	8.27092	8.23611	0.25378	7.73870	8.7335	0.03481
98	1882	8.12892	8.22302	0.25378	7.72561	8.7204	-0.09410
99	1883	7.97692	8.15374	0.25378	7.65633	8.6511	-0.17682
100	1884	7.77367	8.12192	0.25378	7.62451	8.6193	-0.34825
101	1885	7.91967	7.87951	0.25378	7.38211	8.3769	0.04015
102	1886	7.95450	7.96338	0.25378	7.46597	8.4608	-0.00888
103	1887	7.91492	7.99195	0.25378	7.49454	8.4894	-0.07704
104	1888	8.09050	7.95154	0.25378	7.45413	8.4489	0.13896
105	1889	8.31900	8.09199	0.25378	7.59458	8.5894	0.22701
106	1890	7.97167	8.23722	0.25378	7.73981	8.7346	-0.26555
107	1891	8.01933	8.05958	0.25378	7.56217	8.5570	-0.04024
108	1892	8.06583	8.11886	0.25378	7.62146	8.6163	-0.05303
109	1893	8.05675	8.07822	0.25378	7.58082	8.5756	-0.02147
110	1894	8.16100	8.06504	0.25378	7.56763	8.5624	0.09596
111	1895	8.14700	8.20856	0.25378	7.71115	8.7060	-0.06156
112	1896	8.21492	8.13603	0.25378	7.63862	8.6334	0.07889
113	1897	8.29333	8.21787	0.25378	7.72046	8.7153	0.07546
114	1898	8.17850	8.30103	0.25378	7.80362	8.7984	-0.12303
115	1899	8.39633	8.20205	0.25378	7.70464	8.6995	0.19428
116	1900	8.50158	8.36178	0.25378	7.86437	8.8592	0.13980
117	1901	8.54192	8.44854	0.25378	7.95113	8.9459	0.09338
118	1902	8.30442	8.47187	0.25378	7.97446	8.9693	-0.16745
119	1903	8.22017	8.40588	0.25378	7.90847	8.9033	-0.18571
120	1904	8.09092	8.29092	0.25378	7.79351	8.7883	-0.20000
121	1905	8.22517	8.18179	0.25378	7.68439	8.6792	0.04337
122	1906	8.37908	8.24189	0.25378	7.74448	8.7393	0.13720
123	1907	7.94717	8.34596	0.25378	7.84855	8.8434	-0.39879
124	1908	8.18608	8.08403	0.25378	7.58662	8.5814	0.10205
125	1909	8.17825	8.25319	0.25378	7.75579	8.7506	-0.07494
126	1910	8.22300	8.17278	0.25378	7.67537	8.6702	0.05022
127	1911	8.17700	8.22969	0.25378	7.73228	8.7271	-0.05269

观测	Year	AvgTemperature	FORECAST	STD	L95	U95	RESIDUAL
128	1912	8.16542	8.27496	0.25378	7.77755	8.7724	-0.10954
129	1913	8.29775	8.17817	0.25378	7.68077	8.6756	0.11958
130	1914	8.59050	8.30556	0.25378	7.80816	8.8030	0.28494
131	1915	8.59317	8.50324	0.25378	8.00583	9.0006	0.08993
132	1916	8.22525	8.52445	0.25378	8.02704	9.0219	-0.29920
133	1917	8.02333	8.36789	0.25378	7.87048	8.8653	-0.34456
134	1918	8.13083	8.21660	0.25378	7.71919	8.7140	-0.08576
135	1919	8.38225	8.15696	0.25378	7.65955	8.6544	0.22529
136	1920	8.35875	8.34253	0.25378	7.84512	8.8399	0.01622
137	1921	8.57100	8.37187	0.25378	7.87446	8.8693	0.19913
138	1922	8.40800	8.50909	0.25378	8.01168	9.0065	-0.10109
139	1923	8.42217	8.44927	0.25378	7.95186	8.9467	-0.02710
140	1924	8.50683	8.46295	0.25378	7.96554	8.9604	0.04388
141	1925	8.53442	8.50453	0.25378	8.00712	9.0019	0.02989
142	1926	8.72883	8.50919	0.25378	8.01179	9.0066	0.21964
143	1927	8.52208	8.68418	0.25378	8.18677	9.1816	-0.16210
144	1928	8.63217	8.55329	0.25378	8.05588	9.0507	0.07887
145	1929	8.24242	8.63153	0.25378	8.13412	9.1289	-0.38911
146	1930	8.63358	8.39836	0.25378	7.90095	8.8958	0.23522
147	1931	8.71925	8.57074	0.25378	8.07334	9.0682	0.14851
148	1932	8.70525	8.66025	0.25378	8.16284	9.1577	0.04500
149	1933	8.33825	8.66693	0.25378	8.16952	9.1643	-0.32868
150	1934	8.62833	8.52215	0.25378	8.02475	9.0196	0.10618
151	1935	8.51575	8.57687	0.25378	8.07946	9.0743	-0.06112
152	1936	8.55400	8.54970	0.25378	8.05230	9.0471	0.00430
153	1937	8.69833	8.56714	0.25378	8.06973	9.0645	0.13119
154	1938	8.86367	8.69586	0.25378	8.19845	9.1933	0.16781
155	1939	8.76033	8.72971	0.25378	8.23231	9.2271	0.03062
156	1940	8.76467	8.81003	0.25378	8.31262	9.3074	-0.04536
157	1941	8.76867	8.75259	0.25378	8.25518	9.2500	0.01608
158	1942	8.72842	8.77208	0.25378	8.27467	9.2695	-0.04366
159	1943	8.75508	8.73198	0.25378	8.23458	9.2294	0.02310
160	1944	8.84500	8.78628	0.25378	8.28887	9.2837	0.05872

观测	Year	AvgTemperature	FORECAST	STD	L95	U95	RESIDUAL
161	1945	8.58483	8.77604	0.25378	8.27863	9.2734	-0.19120
162	1946	8.67692	8.69543	0.25378	8.19803	9.1928	-0.01852
163	1947	8.79708	8.69449	0.25378	8.19708	9.1919	0.10259
164	1948	8.74675	8.76462	0.25378	8.26721	9.2620	-0.01787
165	1949	8.58675	8.72987	0.25378	8.23247	9.2273	-0.14312
166	1950	8.36525	8.70851	0.25378	8.21110	9.2059	-0.34326
167	1951	8.62567	8.47166	0.25378	7.97425	8.9691	0.15401
168	1952	8.63825	8.63531	0.25378	8.13791	9.1327	0.00294
169	1953	8.87375	8.62443	0.25378	8.12702	9.1218	0.24932
170	1954	8.56067	8.80897	0.25378	8.31156	9.3064	-0.24830
171	1955	8.62625	8.65595	0.25378	8.15854	9.1534	-0.02970
172	1956	8.27958	8.67854	0.25378	8.18114	9.1760	-0.39896
173	1957	8.73250	8.44707	0.25378	7.94967	8.9445	0.28543
174	1958	8.77375	8.67119	0.25378	8.17378	9.1686	0.10256
175	1959	8.73267	8.74164	0.25378	8.24424	9.2391	-0.00898
176	1960	8.58375	8.72727	0.25378	8.22986	9.2247	-0.14352
177	1961	8.79767	8.71856	0.25378	8.22115	9.2160	0.07911
178	1962	8.74592	8.73543	0.25378	8.23803	9.2328	0.01048
179	1963	8.85600	8.77554	0.25378	8.27813	9.2729	0.08046
180	1964	8.40967	8.84134	0.25378	8.34394	9.3388	-0.43168
181	1965	8.52742	8.59600	0.25378	8.09859	9.0934	-0.06858
182	1966	8.59933	8.58435	0.25378	8.08694	9.0818	0.01498
183	1967	8.70008	8.66017	0.25378	8.16276	9.1576	0.03991
184	1968	8.51633	8.64546	0.25378	8.14805	9.1429	-0.12913
185	1969	8.59508	8.66782	0.25378	8.17041	9.1652	-0.07273
186	1970	8.70092	8.61695	0.25378	8.11954	9.1144	0.08397
187	1971	8.59925	8.71829	0.25378	8.22088	9.2157	-0.11904
188	1972	8.49958	8.64619	0.25378	8.14878	9.1436	-0.14660
189	1973	8.94825	8.64242	0.25378	8.14501	9.1398	0.30583
190	1974	8.46750	8.81770	0.25378	8.32030	9.3151	-0.35020
191	1975	8.74483	8.62879	0.25378	8.13138	9.1262	0.11604
192	1976	8.34725	8.76493	0.25378	8.26752	9.2623	-0.41768
193	1977	8.85025	8.53625	0.25378	8.03884	9.0337	0.31400

观测	Year	AvgTemperature	FORECAST	STD	L95	U95	RESIDUAL
194	1978	8.69275	8.73495	0.25378	8.23755	9.2324	-0.04220
195	1979	8.73342	8.80557	0.25378	8.30816	9.3030	-0.07215
196	1980	8.98033	8.69818	0.25378	8.20077	9.1956	0.28215
197	1981	9.16583	9.00475	0.25378	8.50734	9.5022	0.16108
198	1982	8.63917	9.00648	0.25378	8.50907	9.5039	-0.36731
199	1983	9.02817	8.85558	0.25378	8.35818	9.3530	0.17258
200	1984	8.69183	8.98721	0.25378	8.48980	9.4846	-0.29537
201	1985	8.65800	8.80672	0.25378	8.30931	9.3041	-0.14872
202	1986	8.83358	8.71464	0.25378	8.21723	9.2120	0.11894
203	1987	8.99442	8.94174	0.25378	8.44433	9.4391	0.05268
204	1988	9.20158	8.82875	0.25378	8.33134	9.3262	0.37284
205	1989	8.92200	9.20386	0.25378	8.70645	9.7013	-0.28186
206	1990	9.23417	8.96807	0.25378	8.47066	9.4655	0.26610
207	1991	9.17942	9.18243	0.25378	8.68502	9.6798	-0.00301
208	1992	8.83658	9.14136	0.25378	8.64395	9.6388	-0.30478
209	1993	8.86658	9.00264	0.25378	8.50524	9.5001	-0.13606
210	1994	9.03875	8.92855	0.25378	8.43114	9.4260	0.11020
211	1995	9.34708	9.03042	0.25378	8.53301	9.5278	0.31667
212	1996	9.03892	9.20092	0.25378	8.70351	9.6983	-0.16200
213	1997	9.20258	9.13775	0.25378	8.64035	9.6352	0.06483
214	1998	9.52267	9.17456	0.25378	8.67715	9.6720	0.34811
215	1999	9.28508	9.40506	0.25378	8.90765	9.9025	-0.11997
216	2000	9.20117	9.25876	0.25378	8.76135	9.7562	-0.05759
217	2001	9.41458	9.31345	0.25378	8.81604	9.8109	0.10113
218	2002	9.57042	9.31883	0.25378	8.82142	9.8162	0.25158
219	2003	9.52558	9.45872	0.25378	8.96131	9.9561	0.06686
220	2004	9.32458	9.50116	0.25378	9.00375	9.9986	-0.17658
221	2005	9.70092	9.38479	0.25378	8.88738	9.8822	0.31613
222	2006	9.53250	9.54795	0.25378	9.05054	10.0454	-0.01545
223	2007	9.73217	9.52864	0.25378	9.03123	10.0260	0.20353
224	2008	9.43175	9.62329	0.25378	9.12588	10.1207	-0.19154
225	2009	9.50525	9.49950	0.25378	9.00210	9.9969	0.00575
226	2010	9.70308	9.46636	0.25378	8.96895	9.9638	0.23672

观测	Year	AvgTemperature	FORECAST	STD	L95	U95	RESIDUAL
227	2011	9.51600	9.64294	0.25378	9.14553	10.1403	-0.12694
228	2012	9.50733	9.46582	0.25378	8.96841	9.9632	0.04152
229	2013	9.60650	9.56367	0.25378	9.06626	10.0611	0.04283
230	2014	9.57067	9.52901	0.25378	9.03161	10.0264	0.04165
231	2015	9.83100	9.53773	0.25378	9.04032	10.0351	0.29327
232	2016	.	9.72129	0.25378	9.22388	10.2187	.
233	2017	.	9.68121	0.29717	9.09877	10.2637	.
234	2018	.	9.63371	0.32171	9.00316	10.2643	.
235	2019	.	9.67670	0.34609	8.99837	10.3550	.
236	2020	.	9.58985	0.36784	8.86890	10.3108	.
237	2021	.	9.61426	0.37784	8.87371	10.3548	.

附录：代码

1. Python 数据预处理

```
import pandas as pd
import os
import warnings
warnings.filterwarnings('ignore')
print(f'Running script: {os.path.abspath(__file__)}')
print("Hi, PyCharm")
print("Script is running")
file_path = r"C:\Users\立早\Downloads\气温\GlobalTemperatures.csv"
print(f"Using file path: {file_path}")
global_temp = pd.read_csv(file_path)
global_temp['dt'] = pd.to_datetime(global_temp['dt'], errors='coerce')
output_file_path = r"C:\Users\立早\Downloads\气温\Transformed_GlobalTemperatures.csv"
global_temp[['dt', 'LandAverageTemperature']].to_csv(output_file_path, index=False)
print(f'Transformed data has been saved to {output_file_path}')
```

2. R 语言可视化

```
library(ggplot2)
library(dplyr)
library(plotly)
library(tidyr)
file_path <- "C:/Users/立早/Downloads/气温/GlobalTemperatures.csv"
global_temp <- read.csv(file_path)
global_temp$year <- as.integer(substr(global_temp$dt, 1, 4))
data <- global_temp %>%
  group_by(year) %>%
  summarise(LandAverageTemperature = mean(LandAverageTemperature, na.rm =
TRUE),
            LandAverageTemperatureUncertainty =
mean(LandAverageTemperatureUncertainty, na.rm = TRUE))
data <- data %>%
  mutate(Uncertainty_top = LandAverageTemperature +
LandAverageTemperatureUncertainty,
         Uncertainty_bottom = LandAverageTemperature -
LandAverageTemperatureUncertainty)
p <- ggplot(data, aes(x = year)) +
  geom_line(aes(y = LandAverageTemperature, color =
"LandAverageTemperature")) +
  geom_line(aes(y = Uncertainty_top, color = "Uncertainty_top")) +
```

```

    geom_line(aes(y = Uncertainty_bottom, color = "Uncertainty_bottom")) +
    labs(title = "Average Land Temperature in World", x = "Year", y = "Temperature
(°C)") +
    scale_color_manual(values = c("LandAverageTemperature" = "cadetblue4",
"Uncertainty_top" = "brown4", "Uncertainty_bottom" = "darkgoldenrod2"))
ggplotly(p)
global_temp$month <- as.integer(substr(global_temp$dt, 6, 7))
global_temp$season <- ifelse(global_temp$month %in% c(3, 4, 5), "spring",
                             ifelse(global_temp$month %in% c(6, 7, 8),
"summer",
                                     ifelse(global_temp$month %in% c(9, 10,
11), "autumn", "winter"))))
season_data <- global_temp %>%
  group_by(year, season) %>%
  summarise(mean_temp = mean(LandAverageTemperature, na.rm = TRUE)) %>%
  pivot_wider(names_from = season, values_from = mean_temp)
p_season <- ggplot(season_data, aes(x = year)) +
  geom_line(aes(y = spring, color = "spring")) +
  geom_line(aes(y = summer, color = "summer")) +
  geom_line(aes(y = autumn, color = "autumn")) +
  geom_line(aes(y = winter, color = "winter")) +
  labs(title = "Average Temperature in Each Season", x = "Year", y = "Temperature
(°C)") +
  scale_color_manual(values = c("spring" = "mediumaquamarine", "summer"
="brown4", "autumn" = "darkgoldenrod2", "winter" = "cadetblue4"))
ggplotly(p_season)

```

3. SAS 模拟代码

```

/* 删除现有的 WORK.IMPORT 表 */
proc datasets lib=work nolist;
    delete IMPORT;
quit;
/* 定义文件路径 */
FILENAME REFFILE
'/home/u63877733/sasuser.v94/mylib/Transformed_GlobalTemperatures.csv';
/* 手动导入数据 */
data WORK.IMPORT;
    infile REFFILE delimiter=' ' MISSOVER DSD firstobs=2;
    informat dt yymmdd10.;
    informat LandAverageTemperature best32.;
    format dt yymmdd10.;
    input dt LandAverageTemperature;

```

```

        if not missing(dt) and not missing(LandAverageTemperature); /* 去除缺失值
*/
run;
/* 新建一个逻辑库 mylib, 并保存数据 */
libname mylib '/home/u63877733/sasuser.v94/mylib';
data mylib.Temperatures;
    set WORK.IMPORT;
run;
/* 创建一个新的数据集来计算年度平均温度 */
proc sql;
    create table mylib.AvgTemperatures as
    select year(dt) as Year,
           mean(LandAverageTemperature) as AvgTemperature
    from mylib.Temperatures
    group by calculated Year;
quit;
/* 计算一阶差分 */
data mylib.AvgTemperatures_diff;
    set mylib.AvgTemperatures;
    if _N_ > 1 then Diff_AvgTemperature = dif(AvgTemperature);
run;
/* ODS PDF 输出设置 */
ods pdf file="/home/u63877733/sasuser.v94/mylib/temperature_analysis.pdf"
style=journal;
ods word file="/home/u63877733/sasuser.v94/mylib/temperature_analysis.docx";
/* 绘制清洗后的数据时序图 */
proc sgplot data=mylib.AvgTemperatures;
    series x=Year y=AvgTemperature / markers lineattrs=(color=red);
    xaxis label="Year" values=(1785 to 2020 by 10);
    yaxis label="Land Average Temperature" values=(4 to 12 by 2);
run;
/* 绘制差分后的时序图 */
proc sgplot data=mylib.AvgTemperatures_diff;
    where not missing(Diff_AvgTemperature);
    series x=Year y=Diff_AvgTemperature / markers lineattrs=(color=red);
    xaxis label="Year" values=(1785 to 2020 by 10);
    yaxis label="Difference of Land Average Temperature";
run;
/* 重新检查原始数据的平稳性和纯随机性 */
proc arima data=mylib.AvgTemperatures;
    identify var=AvgTemperature(1) nlag=24 stationarity=(adf) minic p=(0:10)
q=(0:10);
run;
/* 预测未来 6 年的值 */

```

```

proc arima data=mylib.AvgTemperatures;
    identify var=AvgTemperature(1) nlag=24;
    estimate p=(1,3,4) q=5;
    forecast lead=6 id=Year out=results;
run;
/* 检查预测结果 */
proc print data=results; run;
/* 绘制预测结果 */
proc sgplot data=results;
    series x=Year y=AvgTemperature / markers lineattrs=(color=black);
    series x=Year y=Forecast / markers lineattrs=(color=red);
    band x=Year lower=L95 upper=U95 / transparency=0.5 fillattrs=(color=green);
    xaxis label="Year" values=(1785 to 2025 by 10);
    yaxis label="Land Average Temperature" values=(4 to 12 by 2);
run;
/* 结束 ODS PDF 和 Word 输出 */
ods pdf close;
ods word close;

```