全球气温不确定性的ARMA模型开发与分析

Mahmud Hasan[*1]、Gauree Wathodkar[†1]和Mathias Muia[‡1]

[1]密西西比大学数学系，美国密西西比州牛津

2023年3月6日

arXiv:2303.02070v1 [stat.ME] 3 Mar 2023

摘要

可基于统计方法建立陆地和海洋表面的温度不确定性模型。本文开发了一种新颖的时间序列温度不确定性模型，即自回归移动平均（ARMA）(1，1)模型。该模型针对观测年平均温度异常 $X(t)$开发，该异常由某一年度（t）的真实（潜在）全球异常 $Y(t)$与正态变量$w(t)$组合而成。不确定性以 $w(t)$的方差表示，该方差分解为陆地表面温度（LST）不确定性、海表温度（SST）不确定性及相应的不确定性来源。本文对该ARMA模型进行了分析，并将其与自回归（AR）和自回归积分移动平均（ARIMA）模型进行比较；数据取自NASA戈达德空间研究所地表温度分析。对自相关函数（ACF）、偏自相关函数（PACF）、正态分位数-分位数（Normal Q-Q）图、残差密度以及正态变量 $w(t)$方差的统计分析表明，对于 $d = 1$，2，ARMA(1， 1)的拟合效果优于AR(1)和ARIMA(1， d，1)。

关键词： ARMA(1，1)、AR(1)、ARIMA(1，1，1)、自相关函数、偏自相关函数、陆地表面温度（LST）不确定性、海表温度（SST）不确定性。

1 引言

温度不确定性会在多个方面对天文学研究产生重大影响。使用望远镜和其他天文仪器进行的观测往往对温度敏感。温度变化时，仪器的灵敏度和响应也会随之改变，若未能准确监测温度，就会导致测量误差

[*]mhasan4@go.olemiss.edu
[†]gkwathod@go.olemiss.edu
[‡]mnmuia@go.olemiss.edu

并需要进行校正。温度变化也会影响天文数据的质量。例如，用于观测恒星光线的探测器温度波动会给数据引入噪声，而且这种噪声很难与真实信号区分开来。因此，温度控制和准确的温度测量是天文学研究许多领域的重要考量，研究人员也常常竭尽全力，尽量减少温度不确定性对研究结果的影响。

海表温度（SST）的温度不确定性 [17] 和陆地表面温度（LST） [18] 通常采用统计方法建模。用于描述温度不确定性的一种常见时间序列模型是自回归移动平均（ARMA） [13]模型。在该模型中，某一时间点的温度被建模为其过去取值以及前几个时间点的残差（误差）的函数。然而，可以基于海表温度（SST）和陆地表面温度（LST）的不确定性构建ARMA模型。其他时间序列模型还包括自回归、自回归积分移动平均（ARIMA） [13]，季节性自回归积分移动平均（SARIMA） [13] 以及广义自回归条件异方差（ GARCH） [13] ，这些模型均可用于温度不确定性数据。

全球历史气候网络月度数据集（GHCNm）提供陆地表面温度 [20]。海表温度则通过扩展重建海表温度（ERSST） [19]分析确定。ERSST [15]使用最新可用的国际综合海洋-大气数据集（ICOADS） [15]以及ARMA和ARIMA等统计方法，从稀疏数据中实现稳定重建。

James Hansen 于20世纪70年代末定义了 GISS温度分析方案，当时需要一种估算全球温度变化模型的方法，以便与一维全球气候模型进行比较。该分析方法由 [1]完整记录。该分析根据地球站点网络在以下三个时期的观测时段，对 GISS-ER [2] 气候模型的长期运行结果进行子采样。 [14]基于全球和区域平均温度异常时间序列分析，开发了另一种复杂的统计不确定性模型。最近， [12] 采用插值方法，利用假定的基于距离的优化技术生成基于克里金法的场。在本文中，我们正在开发一种受 [6] 启发的统计不确定性模型，其表现优于其他一些模型。然而，如何针对该模型优化数据仍有待探索，这可能成为未来由相应优化技术 [7],[8]开展的研究。

不确定性模型基于现有方法，用于预测温度不确定性。这些模型通过估计前一年数据的概率，改进了 GISTEMP 的不确定性分析。本文构建了一个 ARMA（1，1）温度不确定性模型。我们将时间（年份）t 的年度均值温度异常 $X(t)$ 视为真实（潜在）全球异常 $Y(t)$ 与随机变量 $w(t) \sim N(0, \sigma^2)$ 的线性组合。不确定性定义为 $w(t)$ 的方差，可分解为陆地表面温度（LST）不确定性、海表温度（SST）不确定性及相应的不确定性来源。此外，由 $X(t)$ 和 $Y(t)$ 得出的差分序列在引入系统偏差后形成 ARMA 模型。

我们通过将 ARMA 模型与 AR 和 ARIMA 模型进行比较，并考察自相关函数（ACF）、偏自相关函数（PACF）和残差密度等时间序列性质，验证了 ARMA 模型的合理性。我们增加了一项新的性质——不确定性，用于衡量模型与相应数据的拟合程度。为分析模型，我们采用 NASA 戈达德空间研究所地表温度分析项目的数据；该项目提供了涵盖 1880 年至今、具有月分辨率的综合全球地表温度数据集。我们使用 Python 中的 autoarima 函数为相应数据建立模型。本文分为若干章节。

在第 2 节中，我们针对时刻 t 的观测年平均温度异常 $X(t)$ 建立了 ARMA模型。变量 $X(t)$ 被分解为某一年度 t 的温度真实（潜在）全球异常$Y(t)$ 和正态变量 $w(t)$。随后，正态变量被分解为陆地温度异常和海洋温度异常，其真实异常在 $t-1$ 处等于观测平均温度异常 $X(t)$。我们使用 Python 中的 autoarima 函数对所有模型进行了数据分析。

在第 3 节中，建立模型后，我们发现所提出的 ARMA(1，1) 模型为讨论自相关函数（自相关函数）、偏自相关函数（偏自相关函数）、正态分位数-分位数（正态 Q-Q）图以及残差密度如何影响 ARMA模型提供了研究空间。变量 $w(t)$ 的方差会影响模型，这一点已在数据分析与讨论部分进行了说明。我们通过数学评估和相应参数的作用，分别阐释了每项属性。

在第4节中，我们对结果进行了详细分析和讨论。首先，使用增广迪基-富勒检验（ADF检验）验证了GISTEMP数据的非平稳性。随后，对AR(1)、ARMA(1，1)、ARIMA(1，1，1)和ARIMA(1，2，0)模型进行了详细分析，包括模型诊断和残差分析。我们根据参数估计对模型进行了比较。所有模型均使用Python中的autoarima 函数进行拟合。

最后一节通过理论发现和数值分析总结了我们得到的结果。

1.1 模型的初步定义

我们可以从 [13]中获得模型的相应定义和符号。本节将讨论模型的初步定义及相应的系数参数：AR、ARMA和ARIMA。

自回归（AR）

对于当前时刻 t的值，阶数为 p的自回归$\mathbf{AR(p)}$ 模型

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \phi_3 X_{t-3} + \ldots\ldots\ldots + \phi_p X_{t-p} + w_t \tag{1}$$

X_t 是平稳的，参数 $\phi_1, \phi_2, \phi_3, \ldots\ldots\ldots\phi_p$ 为常数，且 $\phi_p \neq 0$ 和 $w_t \sim WN(0, \sigma_w^2)$。通过后移算子，$\mathbf{AR(p)}$ 可写为

$$\phi(B) X_t = \epsilon_t \tag{2}$$

其中 $\phi(B) = (1 - \phi_1 B - \phi_2 B^2 - ..... - \phi_p B^p)$。 $AR(1)$ 的自相关函数为

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)} = \phi^h, \ h \geq 0 \tag{3}$$

其中，自协方差函数 $\rho(h)$ 满足 $\rho(h) = \phi\psi(h-1) \quad h = 1, \ 2, \ \cdots$.

自回归移动平均（ARMA）

如果时间序列为 x_t; $t = 0, \ \pm 1, \ \pm 2,$ ....., 则当其平稳且满足以下条件时，称为 ARMA($p, \ q$):

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \phi_3 X_{t-3} + ... + \phi_p X_{t-p} + \theta_1 w_{t-1} + \theta_2 w_{t-2} + \theta_3 w_{t-3} + ... + \theta_q w_{t-q} \tag{4}$$

其中 $\phi_p, \ \theta_q \neq 0$ ，且 $\sigma_w^2 \geq 0$。参数 p 和 q 分别称为自回归阶数和移动平均阶数。

自回归积分移动平均（ARIMA）

如果过程 X_t 满足以下条件，则称其为 ARIMA($p, \ d, \ q$):

$$\nabla^d X_t = (1 - B)^d X_t \tag{5}$$

是针对季节性参数 d 和后移参数 B 的 ARMA(p, q)。一般来说，我们将模型写作

$$\phi(B)(1 - B)^d X_t = \theta(B) w_t. \tag{6}$$

AR、ARMA 和 ARIMA 模型的系数对于决定模型捕捉时间序列行为的能力起着关键作用，而系数的选择会对模型的预测产生重大影响。在我们的模型中，系数参数被引入为温度方差，这会影响该模型与其他属性的可比性。

2 不确定性 ARMA 模型

令 $Y(t)$ 表示年份 t 的真实（潜在）全球温度异常，我们将计算得到的（观测到的）年平均温度异常视为

$$X(t) = Y(t) + w(t). \tag{7}$$

随机变量 $w(t) \sim N(0, \sigma^2)$。因此，我们将全球平均异常计算中的不确定性定义为

$$W(t) = \sigma^2. \tag{8}$$

我们可以将总不确定性分解为

$$W(t) = \sigma^2 = \sigma_L^2 + \sigma_S^2. \tag{9}$$

其中，不确定性分解为两个部分：由陆地计算中的不确定性导致的全球平均异常不确定性 σ_L^2 ，以及由海面计算中的不确定性导致的全球平均异常不确定性 σ_S^2。这意味着必然存在随机变量，分别表示由陆地中的不确定性导致的异常 $w_L(t) \sim N(0, \sigma_L^2)$，以及由海面中的不确定性导致的异常 $w_S(t) \sim N(0, \sigma_S^2)$，据此我们可以将（1）写为

$$X(t) = Y(t) + w_L(t) + w_S(t). \tag{10}$$

采用修正后的 GISTEMP 方法，针对 14 个十年时间段分别计算缩减覆盖全球年平均值 $X_i(t)$。其中， i 表示所用的十年， t 表示年份时间。十年 i 的差分序列为

$$D_i(t) = Y(t) - X_i = w_{iL}(t) + w_{iS}(t). \tag{11}$$

我们引入潜在的系统加性偏差 α_i 和乘性偏差 β_i。于是，式（1）可表示为

$$X_i(t) = \alpha_i + \beta_i Y(t) + w_{iL}(t) + w_{iS}(t). \tag{12}$$

然后，将陆地温度异常和海洋温度异常分解为真实异常，其中 $t-1$ 等于观测到的平均温度异常 $X(t)$。于是，我们可以写为

$$X_i(t) = \alpha_i + \beta_i X(t-1) + w_{iL}(t) + w_{iS}(t-1). \tag{13}$$

备注

式（13）表示 $p = 1, q = 1$ 的 ARMA（p, q）模型，其中 $X_i(t) = \alpha_i + \beta_i X(t-1)$ 是一阶自回归，即 AR(1)； $X_i(t) = w_{iL}(t) + w_{iS}(t-1)$ 是一阶移动平均，即 MA(1)。MA(1)的系数 w_{iL} 和 w_{iS} 分别表示陆地表面温度和海洋表面温度的不确定性。通过使用 NASA 戈达德空间研究所地表温度分析的数据，我们表明，与 AR(1) 和 ARIMA 相比，该系数能使 ARMA 模型（13）拟合得更好。除了系数之外，我们还发现，对于其他时间序列性质——自相关函数（ACF）、偏自相关函数（PACF）、正态分位数-分位数图（正态 Q-Q 图）以及残差密度，模型（13）的拟合效果也更好。

3 ARMA(1,1) 的统计特性

本节将说明哪些时间序列性质可以帮助我们拟合出更好的模型。在我们的案例中，自相关函数（ACF）、偏自相关函数（PACF）、正态分位数-分位数图（正态 Q-Q 图）和残差密度都会影响 ARMA模型。另一个影响模型的性质是变量 w（t）的方差，我们将在数据分析部分讨论这一点。

3.1 自相关函数（ACF）

如果将 ARMA(1,1) 写成这种因果形式来表示 $X_t = \sum_{j=1}^{\infty} \psi_j w_{t-j}$,

$$X_t = w_{iL}(t) + (1+\beta_i) w_{iS}(t-1), \text{ for } 0 < \beta_i < 1,$$

then

$$X_t = w_{iL}(t) + (1+\beta_i) w_{iS}(t-1).$$

对应的自相关函数（ACF） [13]可表示为

$$\rho(h) = \frac{(1+\beta_i)^2}{2(1-\beta_i)} \beta_i^{h-1} = \frac{1}{2}(1+\beta_i)\beta_i^{h-1}. \tag{14}$$

在方程（14）中， h 表示时间差， β_i 为参数系数。根据参数值，我们需要判断模型的拟合优度。自相关函数（ACF）定义了数据点与前面的数据点之间的平均相关性，这种关系以时间差（即滞后）为单位。

3.2 偏自相关函数（PACF）

然而，AR(1) 和 ARMA(1,1) 过程是完全相关的，其自相关函数会逐渐衰减，但不会变为零，尽管可能非常接近零。在这种情况下，有时仅根据自相关函数可能无法识别该过程。因此，我们将考虑偏自相关函数（PACF），它与自相关函数结合起来有助于识别模型。零均值平稳时间序列 $\{X_t\}_{t=0,1,2,\cdots}$ 的偏自相关函数定义为

$$\begin{aligned} \phi_{11} &= corr(X_1, X_0) = \rho(1) \\ \phi_{\tau\tau} &= corr(X_\tau - f_{(\tau-1)}, X_0 - f_{(\tau-1)}),\ \tau \geq 2, \end{aligned} \tag{15}$$

其中

$$f_{\tau-1} = f(X_\tau, \cdots, X_1)$$

使均方线性预测误差最小

$$E(X_\tau - f_{(\tau-1)})^2.$$

函数 f 的下标表示该函数所依赖的变量数量。$\phi_{\tau\tau}$ 表示变量 X_t 和 $X_{t-\tau}$ 之间的线性相关性。基本上，参数值 ϕ 用于估计 ARMA模型的拟合程度。

3.3 正态分布分位数-分位数图（正态 Q-Q 图）

正态分布分位数-分位数图（Q-Q 图）是一种图形方法，用于评估一组样本数据是否近似服从正态分布。它将样本数据的分位数与理论正态分布的分位数进行比较。

在 ARMA（自回归移动平均）模型中，正态Q-Q图可用于评估残差的正态性。残差是观测值与 ARMA模型预测值之间的差值，

如果残差服从正态分布，则表明 ARMA模型已捕捉到数据中的大部分系统性模式，剩余差异则是随机噪声，可用正态分布很好地近似。

换言之，如果ARMA模型的残差能够很好地近似为正态分布，则表明该模型很好地拟合了数据。然而，如果残差明显偏离正态性，则可能表明ARMA模型拟合效果不佳，应考虑采用其他建模技术或对ARMA模型进行修改。

3.4 预测

这里介绍ARMA的预测方法，尽管预测不能作为其性质使用，但我们可以获得参数估计。预测的目标是根据已收集的当前数据，预测时间序列的未来值。对于数据x_1， $x_2.....x_n$ 我们将预测模型写为

$$X_{n+1} = \beta_i X_n + w_{iL}(n) + w_{iS}(n)$$

其中， β_i 为AR参数系数。一步截断预测为

$$\tilde{X}^n_{n+1} = \beta_i X_n + 0 + \tilde{w}^n_n,$$

使用截断预测： $\tilde{w}^n_0 = 0$， $\tilde{w}^n_1 = X_1$。于是

$$\tilde{w}^n_t = X_t - \beta_i X_{t-1} - \tilde{w}^n_{t-1}, t = 2, \cdots, n.$$

近似预测 [13]

$$p^n_{n+m} = \gamma^2_w \left[1 + (1 + \beta^2_i) \sum_{j=1}^{m-1} \beta^2_i (j-1) \right]$$

$$p^n_{n+m} = \gamma^2_w \left[1 + (1 + \beta_i)^2 \frac{(1 - \beta_i^{2(m-1)})}{1 - \beta_i)^2} \right].$$

$1 - \alpha$ 预测 区间，其中 $X^n_{n+m} \pm C_{\alpha/2} \sqrt{p^n_{n+m}}$ 是置信度的程度 .

根据数据计算预测区间时，我们用参数估计值进行替代，从而得到近似预测区间。预测区间给出了系数参数 β_i的估计值。通常，我们需要通过截断预测获得更好的估计值，也可以通过留出部分数据来检验模型稳定性和模型的预测能力。

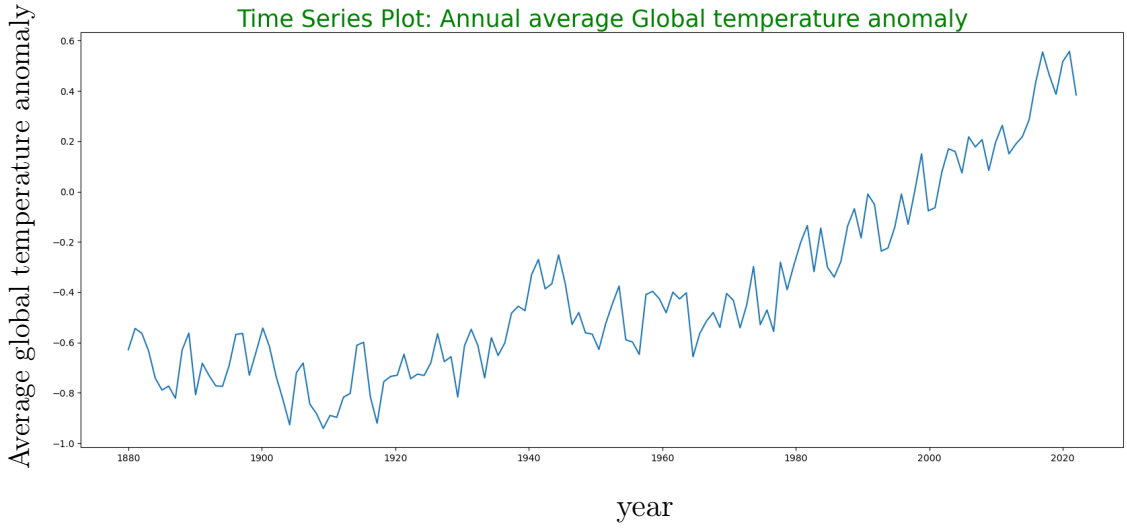
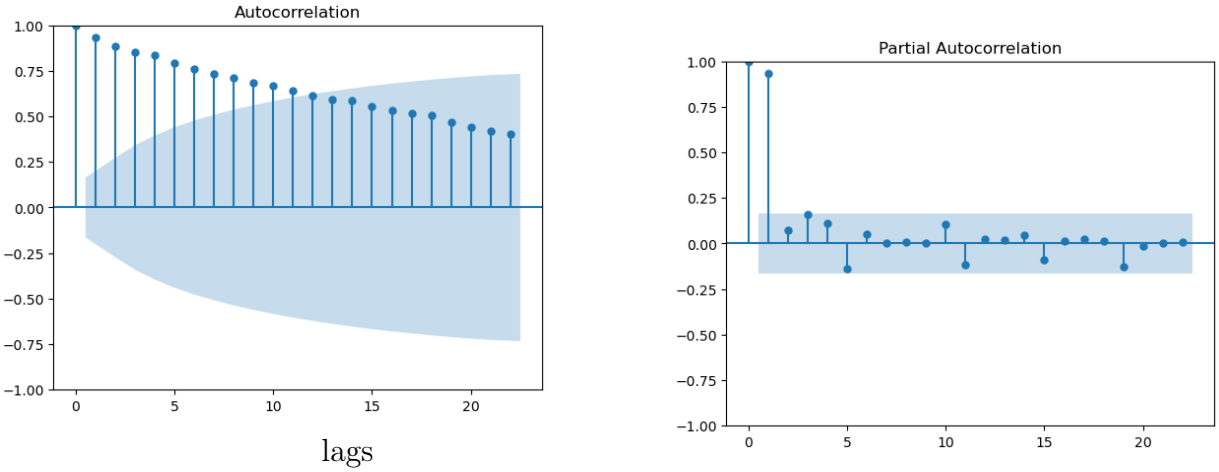


图1：全球温度异常数据时间序列

4 数值数据分析与讨论

我们将考察自1880年以来 NASA 戈达德太空研究所地表温度（GISTEMP）的年度全球温度异常数据。

2019年，Lenssen, N.、G. Schmidt、J. Hansen、M. Menne、A. Persin、R. Ruedy 和 D. Zyss 在 [6]中讨论了 GISTEMP 数据的不同方面。他们讨论了数据年度均值的置信区间，也给出了海洋温度异常的置信区间。此外，他们指出，AR(1)可以作为一种合理的模型，用于比较短时间段内的数据。本文将对 AR(1) 模型进行详细分析，并进一步探讨更复杂的模型，如 ARMA(1,1)和 ARIMA$(1,d,1)$，用于 $d=1$，2。



（a）全球温度数据的自相关　　（b）偏自相关

图2：全球温度数据的自相关函数和偏自相关函数

图1展示了全球年平均气温异常的时间序列图。由于我们考察的是年度平均气温异常，这幅图没有季节性。在图中可以看到，早期全球气温异常呈现出围绕某个均值波动的趋势，一直持续到 1960年。但自1960 年以来，全球气温异常图呈现出明显的上升趋势。其数值显著增加，因此预测全球气温异常的未来数据非常重要。

图2a 显示了全球气温异常时间序列的自相关函数。这是前20 个滞后阶数的自相关值。此处阴影区域表示阈值。在该图中，我们可以看到自相关性缓慢衰减。因此，由于自相关函数图的尖峰超过了阈值区域，该时间序列可能是非平稳的。

图2b 展示了全球气温异常时间序列的偏自相关函数（PACF）。这里，前两个滞后阶数位于阈值之外，随后PACF突然下降，使得其余所有滞后阶数的PACF均位于阈值之内。

使用增广迪基-富勒检验（ADF检验） [16] ，我们将检验该时间序列是否平稳。为此使用了Python软件包statsmodels.tsa.stattools。ADF检验的结果生成如下：

```
ADF Statistic: 1.0582478050392963
n_lags: 0.9948440491482736
p-value: 0.9948440491482736
Critial Values:
   1%, -3.4793722137854926
Critial Values:
   5%, -2.8830370378332995
Critial Values:
   10%, -2.578233635380623
```

当我们对数据进行ADF检验时，所得p值大于0.05的显著性水平，且ADF统计量高于任何一个临界值。这意味着没有理由拒绝零假设。因此，该时间序列实际上是非平稳的。

因此，我们尝试寻找平稳时间序列。为此，我们对给定数据进行一阶差分。图3展示了给定时间序列的一阶差分。它显然没有呈上升趋势。整个数据中都呈现出均值回复行为。

此外，我们绘制了一阶差分时间序列的自相关图（图4a）。在第一个滞后之后，我们注意到自相关函数突然急剧下降。第二个和第三个滞后之间的差异很小。此后几乎所有滞后阶数的自相关函数值都在阈值范围内，但第二和第三个滞后的值超出了阈值。接着，我们绘制了一阶差分时间序列的偏自相关图（图4b）。我们注意到

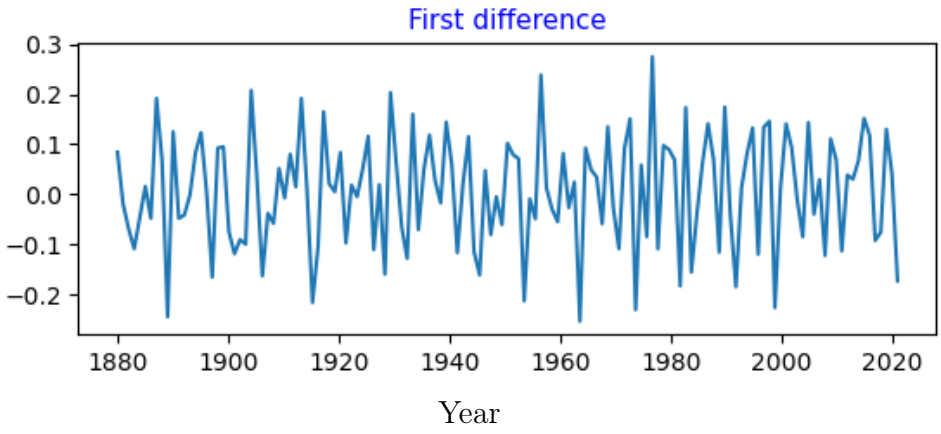
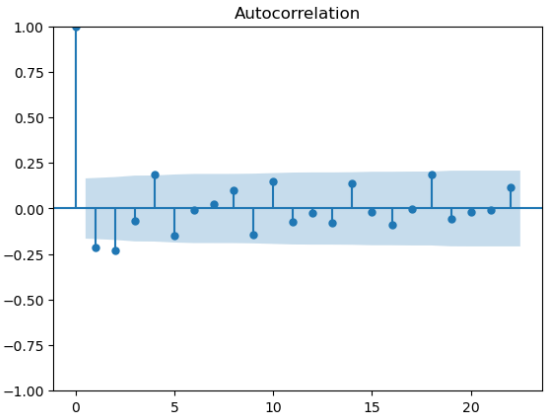
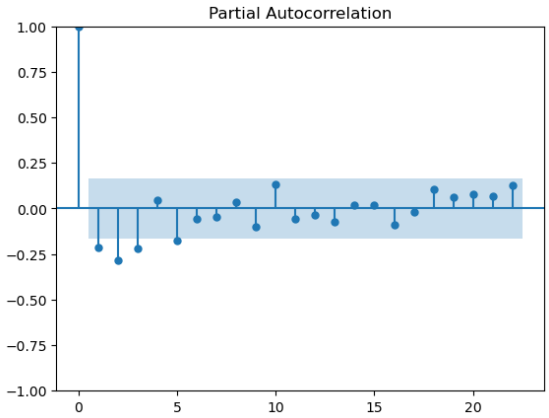


图 3：全球温度异常数据的一阶差分时间序列



(a) 一阶差分的自相关函数



(b) 一阶差分的偏自相关函数

图 4：一阶差分的自相关函数和偏自相关函数

第一个滞后之后，偏自相关函数突然下降，但此后缓慢衰减。前四个滞后的值都超出了阈值。这表明它并不是适用于给定时间序列的良好模型。

图 5 显示了全球温度数据的二阶差分，也体现了均值回归。接下来，我们绘制二阶差分的自相关函数和偏自相关函数。图 6a 显示了二阶差分时间序列的自相关函数。前两个滞后值的幅度远超阈值，而之后的滞后值均在阈值范围内。根据前两个滞后值的表现，我们可以认为该时间序列可能存在过度差分。图 6b 显示了二阶差分时间序列的偏自相关函数图。前四个滞后值均超出阈值，且其幅度远高于阈值。这证实该时间序列存在过度差分，因此，这并非针对给定全球温度数据拟合的最佳模型。

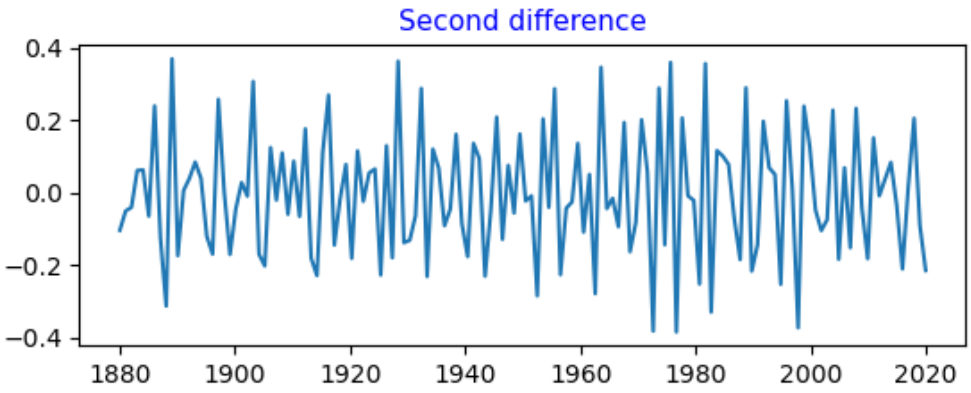
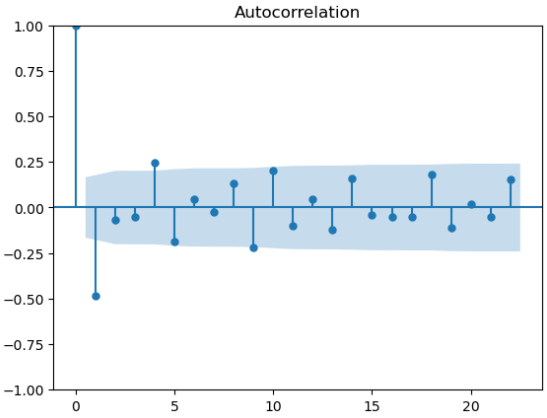
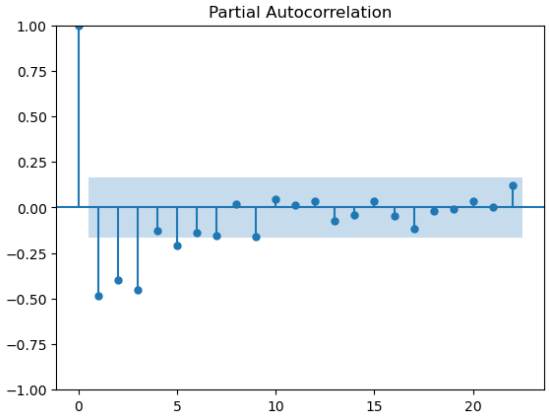


图 5：全球温度数据的二阶差分时间序列



（a）二阶差分的自相关函数



（b）二阶差分的偏自相关函数

图 6：二阶差分的自相关函数和偏自相关函数

4.1 拟合 AR(1) 模型

我们尝试使用 Python 中的 autoarima 函数对给定数据拟合 AR(1)模型。以下是结果摘要：

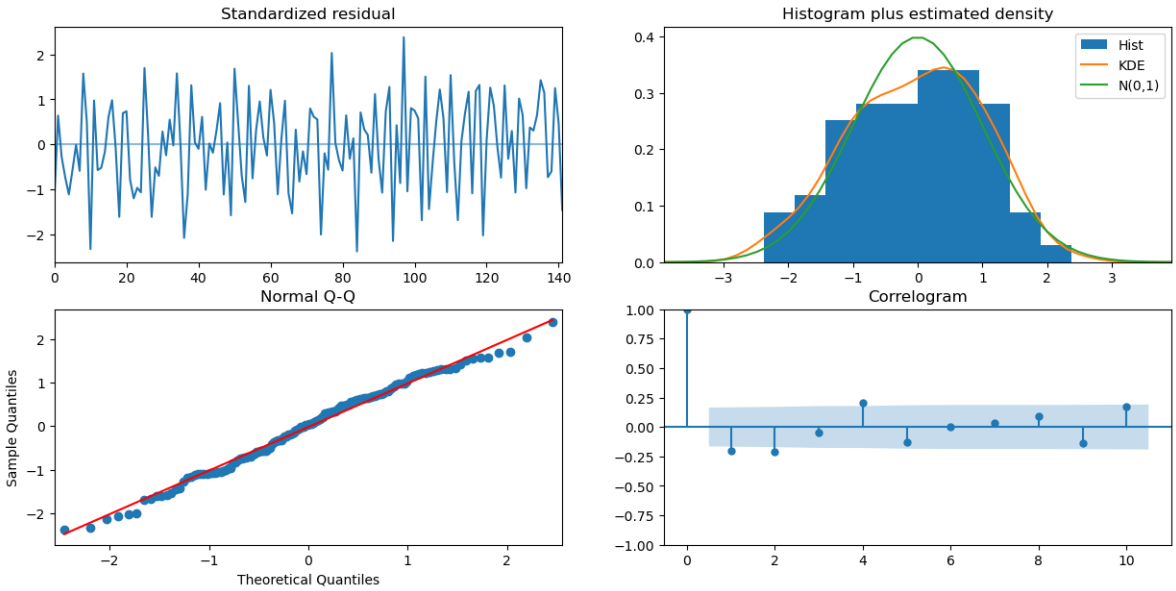


图7：AR(1)模型的诊断

SARIMAX Results

Dep. Variable:	y	No. Observations:	142
Model:	SARIMAX(1, 0, 0)	Log Likelihood	109.574
Date:	Thu, 03 Nov 2022	AIC	-215.148
Time:	11:58:08	BIC	-209.236
Sample:	0	HQIC	-212.746
	- 142		
Covariance Type:	opg		

	coef	std err	z	P>|z|	[0.025	0.975]
ar.L1	0.9786	0.018	55.357	0.000	0.944	1.013
sigma2	0.0122	0.002	7.132	0.000	0.009	0.016

Ljung-Box (L1) (Q):	5.82	Jarque-Bera (JB):	2.70
Prob(Q):	0.02	Prob(JB):	0.26
Heteroskedasticity (H):	1.26	Skew:	-0.17
Prob(H) (two-sided):	0.44	Kurtosis:	2.42

对于 AR(1)模型，我们得到系数 $\phi = 0.9786$ 和 $\sigma = 0.0122$。该模型的偏度为 -0.17，峰度为 2。42。

图 7展示了 AR(1) 模型的诊断结果。第一张图展示了标准化残差。第二张图展示了数据的直方图、标准正态分布 (0，1) 曲线（绿色）以及核密度估计（KDE）图（橙色），该图对给定数据进行了平滑处理。第三张图是正态分布分位数-分位数（正态 Q-Q）图，我们可以清楚地看到，大多数样本分位数和理论分位数在均值附近都符合正态分布。但在两个标准差之外，数据点偏离了参考线。此外，从第一个标准差范围向第二个标准差范围移动时，数据点开始逐渐偏离参考线。最后是相关图。在这里，我们可以看到，该自相关函数图与原始时间序列一阶差分的自相关函数图非常相似，因此这不是拟合效果最佳的模型，我们需要为这组数据

找到更好的模型。

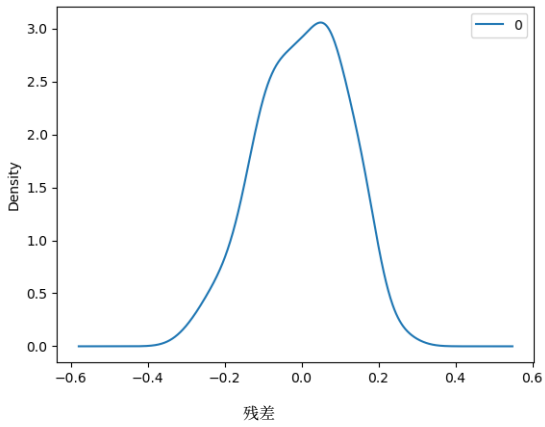


图8：AR(1) 的残差密度

图8 显示了残差密度。大多数残差接近0。此外，从残差分析中可以看出，AR(1)模型的残差最大值为0。266279。

4.2 拟合 ARMA(1，1) 模型

为了获得更好的拟合效果，我们使用 Python 中的 autoarima 函数对给定数据拟合 ARMA(1，1) 模型。以下是结果摘要：

SARIMAX Results

Dep. Variable:	y	No. Observations:	142
Model:	SARIMAX(1, 0, 1)	Log Likelihood	116.156
Date:	Thu, 03 Nov 2022	AIC	-226.313
Time:	11:42:56	BIC	-217.445
Sample:	0	HQIC	-222.709
	- 142		
Covariance Type:	opg		

	coef	std err	z	P>|z|	[0.025	0.975]
ar.L1	0.9938	0.009	107.567	0.000	0.976	1.012
ma.L1	-0.4365	0.075	-5.847	0.000	-0.583	-0.290
sigma2	0.0111	0.002	6.324	0.000	0.008	0.015

Ljung-Box (L1) (Q):	0.98	Jarque-Bera (JB):	4.51
Prob(Q):	0.32	Prob(JB):	0.10
Heteroskedasticity (H):	1.11	Skew:	-0.13
Prob(H) (two-sided):	0.73	Kurtosis:	2.17

对于 ARMA(1，1) 模型，我们得到的系数为 $\phi = 0.9938$、$\theta = -0.4365$ 和 σ =0。0111。该模型的偏度为 -0.13，峰度为 2。17。

在图9 中，我们展示了 ARMA(1,1) 模型的诊断结果。第一幅图显示了标准化残差。将 ARMA(1,1) 与 AR(1)的直方图进行比较，可以看出核密度估计图拟合得更好，也更接近标准正态图。在第三幅图中可以看到，大多数样本分位数和理论分位数在正态分布

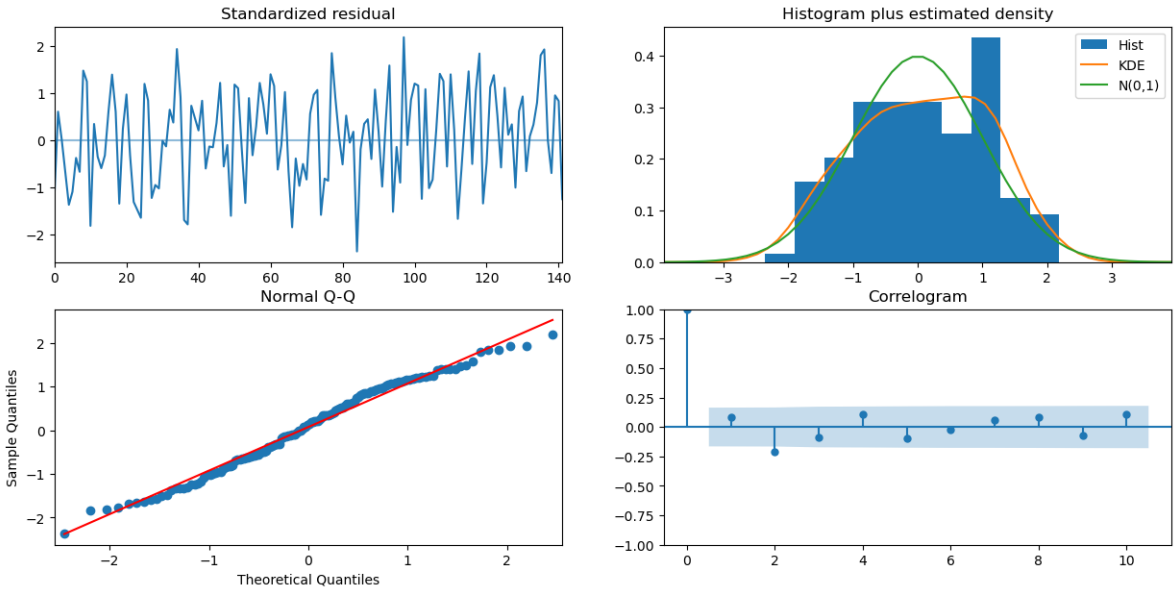


图9：ARMA(1,1) 模型的诊断

均值附近拟合良好。在这里，第一和第二个标准差之间，偏离参考线的点很少。与 AR(1) 图相比，该图清楚地表明，ARMA(1,1) 对给定数据的拟合效果更好。这里的相关图不同于原始时间序列的自相关函数图，第二个滞后位于阈值以内。

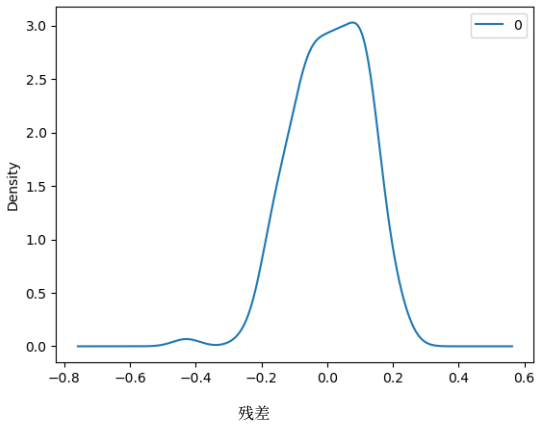


图10：ARMA(1,1) 的残差密度

图 10 显示了残差的密度。大多数残差都接近 0。从 ARMA 残差的分析(1,1) 中可以看出，残差的最大值为 0.232144，这小于 AR(1) 模型的最大残差值。因此，ARMA(1,1) 模型明显优于 AR(1)。

4.3 拟合 ARIMA(1,1,1) 和 ARIMA (1,2,0) 模型

一般来说，我们知道 ARIMA (p, d, q) 模型比 ARMA 模型拟合效果更好。因此，我们尝试对 GISTEMP 数据进行一阶和二阶差分后拟合 ARIMA 模型。对于

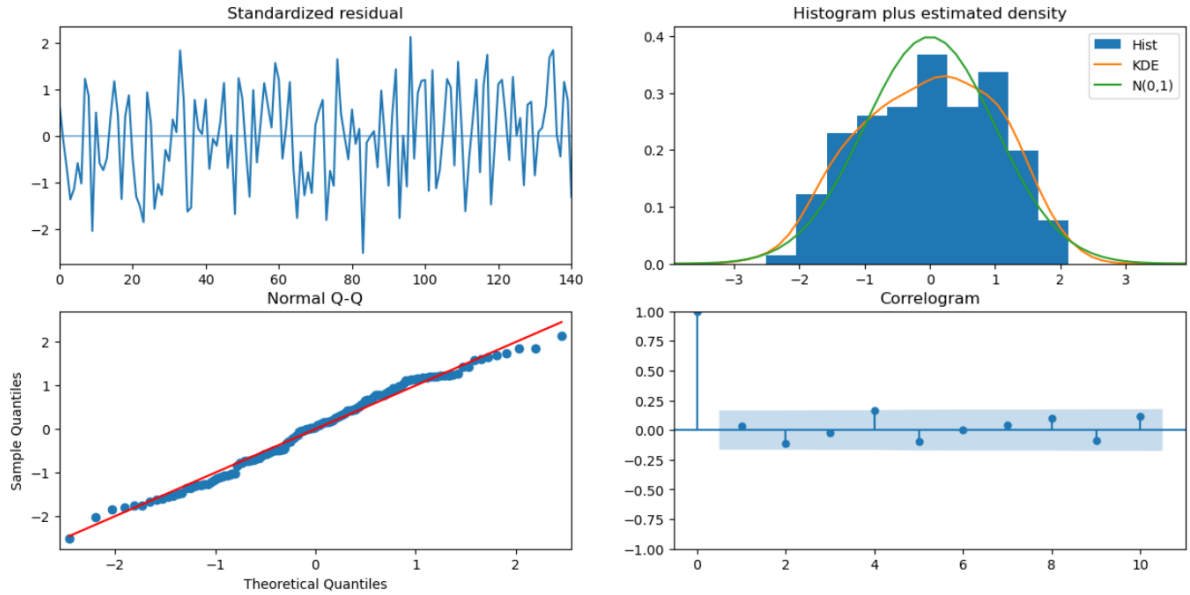


图 11：ARMA(1，1，1)模型的诊断

$d = 1$，模拟结果表明 ARIMA(1, 1,1) 是最佳模型；对于 $d = 2$，拟合效果最佳的模型是 ARIMA(1,2,0)。

与之前的拟合类似，我们尝试使用 Python 中的 autoarima 函数拟合 ARIMA(1，1，1)模型。以下是结果摘要：

SARIMAX Results

Dep. Variable:	y	No. Observations:	142
Model:	SARIMAX(1, 1, 1)	Log Likelihood	121.537
Date:	Sat, 21 Jan 2023	AIC	-235.073
Time:	17:54:13	BIC	-223.278
Sample:	0	HQIC	-230.280
	- 142		
Covariance Type:	opg		

	coef	std err	z	P>|z|	[0.025	0.975]
intercept	0.0050	0.002	2.023	0.043	0.000	0.010
ar.L1	0.3652	0.152	2.398	0.016	0.067	0.664
ma.L1	-0.7617	0.111	-6.879	0.000	-0.979	-0.545
sigma2	0.0104	0.002	6.436	0.000	0.007	0.014

Ljung-Box (L1) (Q):	0.17	Jarque-Bera (JB):	4.19
Prob(Q):	0.68	Prob(JB):	0.12
Heteroskedasticity (H):	1.13	Skew:	-0.13
Prob(H) (two-sided):	0.68	Kurtosis:	2.20

对于 ARIMA(1，1，1)模型，系数为 $\phi = 0.3652, \theta = -0.7617$ 和 $\sigma = 0.0104$。该模型的偏度为 -0.13，峰度为 2。20 。

在图 11 中展示了 ARIMA(1,1,1) 模型的诊断。第一张图显示了标准化残差。比较 ARMA(1,1) 与 ARIMA(1, 1,1)的直方图，可以看到，核密度估计图对 ARMA(1,1) 的拟合效果更好，也更接近标准正态分布

图。ARIMA(1, 1,1) 的正态 Q-Q 图和相关图与 ARMA(1,1) 的相比没有太大差异。

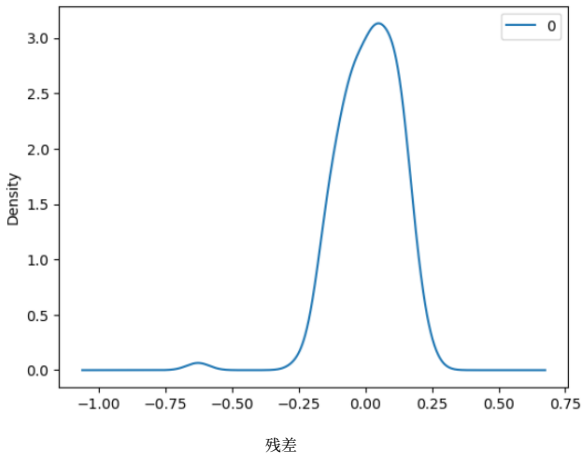


图 12：ARIMA(1，1，1) 的残差密度

图12 显示了残差密度。大多数残差都接近0。1。通过分析 ARIMA(1，1，1) 的残差，我们可以看到残差的最大值为0。240199，高于 ARMA(1，1)模型的最大残差值。因此，ARMA(1，1) 比 ARIMA(1，1，1)更优。

此外，我们尝试拟合一个差分阶数为 $d = 2$ 的 ARIMA 模型。Python 中的 autoarima 函数显示，对于 $p \leq 1$和 $q \leq 1$，ARIMA(1，2，0) 是给定数据的最佳拟合模型。拟合 ARIMA(1，2，0) 后，我们得到以下结果：

```
Best model:  ARIMA(1,2,0)(0,0,0)[0]
Total fit time: 0.518 seconds
                               SARIMAX Results
==============================================================================
Dep. Variable:                      y   No. Observations:                  142
Model:               SARIMAX(1, 2, 0)   Log Likelihood                66.333
Date:                Sun, 22 Jan 2023   AIC                        -128.667
Time:                        14:29:58   BIC                        -122.783
Sample:                             0   HQIC                       -126.276
                                  - 142
Covariance Type:                  opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ar.L1         -0.4902      0.086     -5.680      0.000      -0.659      -0.321
sigma2         0.0227      0.003      7.581      0.000       0.017       0.029
===================================================================================
Ljung-Box (L1) (Q):                   5.87   Jarque-Bera (JB):                 2.05
Prob(Q):                              0.02   Prob(JB):                         0.36
Heteroskedasticity (H):               1.24   Skew:                            -0.23
Prob(H) (two-sided):                  0.46   Kurtosis:                         2.64
===================================================================================
```

对于 ARIMA(1，2，0) 模型，系数为 $\phi = -0.4902$、 $\theta = 0$ 和 $\sigma = 0.0227$。该模型的偏度为 -0.23 ，峰度为 2。64。

在图 13 中，我们给出了 ARIMA$(1, 2, 0)$ 模型的诊断结果。第一幅图显示了标准化残差。比较 ARMA$(1,1)$ 和 ARIMA$(1, 2, 0)$的直方图，

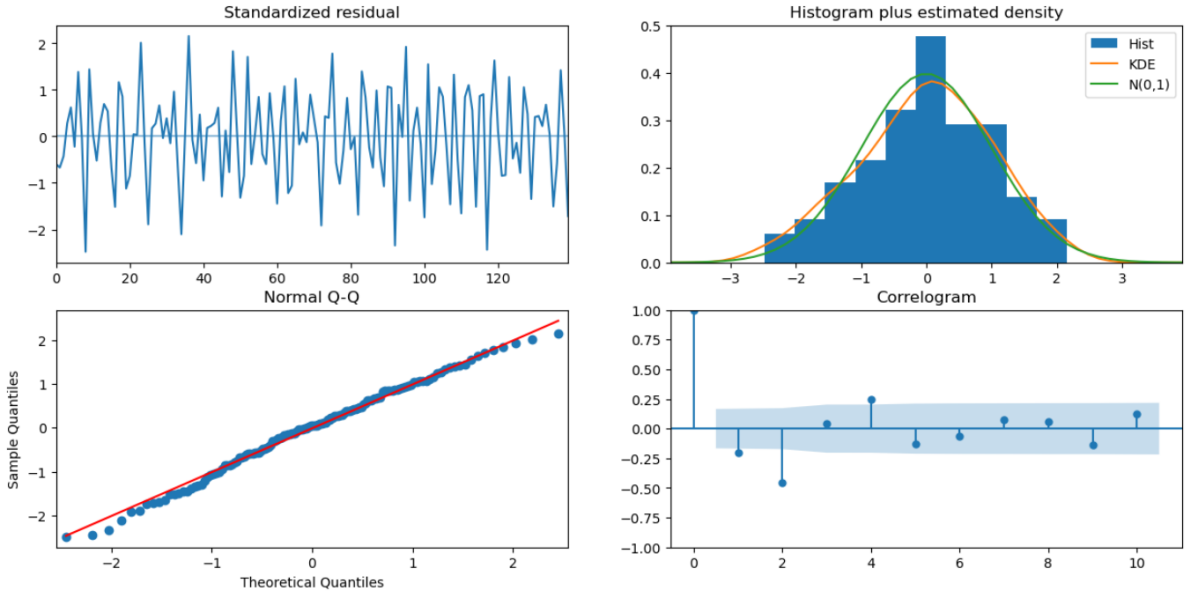


图 13：ARMA(1，2，0)模型的诊断

我们可以看到，这里的核密度估计图拟合得并不好。ARIMA(1，2，0) 的正态 Q-Q 图与 ARMA(1，1) 的看起来没有太大区别。但在相关图中，我们注意到自相关函数的行为有所不同。这里，自相关函数在第一个滞后后骤然下降，而第二和第三个滞后值位于阈值之外。此后，大多数滞后阶数都在阈值范围内，但显然没有趋近于零。这清楚地表明，该模型拟合得并不好。

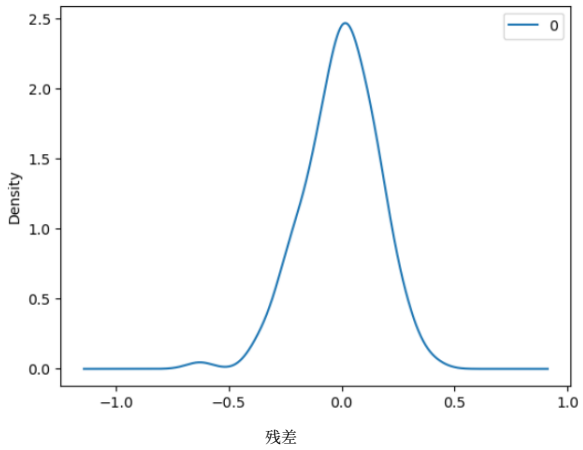


图 14：ARIMA(1，2，0)的残差密度

此外，图14 中的残差密度进一步证实了该结果。ARIMA$(1，2，0)$ 模型的大多数残差都接近0。通过对残差的分析可以看出，残差的最大值为0.398333，这高于ARMA$(1，1)$ 模型的最大残差值。

因此，通过分析差分 $d = 1$和 $d = 2$的最佳拟合模型，即 ARIMA(1，1，1) 和 ARIMA(1，2，0) 模型，可以看出，ARMA(1，1) 模型比这些模型拟合得更好。

5 结论：

我们基于陆地和海洋表面不确定性所导致的全球平均异常不确定性，对新建立的 ARMA (1，1) 模型进行了分析，并将其与 AR(1)、ARIMA(1，1，1) 和 ARIMA(1，2，0) 进行了比较。一般而言，ARIMA (p，d，q) 模型被认为比 ARMA(p，q) 模型拟合得更好。但对于戈达德空间研究所从 1880年 至今的地表温度 (GISTEMP) 数据，我们发现了不同的结果。这里，ARMA(1，1) 模型比 ARIMA(1，1，1) 和 ARIMA(1，2，0) 模型拟合得更好。此外，模拟结果表明，ARIMA(1，2，0) 比 ARIMA(1，2，1) 拟合得更好，因此可以得出结论：对于 $d = 1$ 和 $d = 2$，ARMA(1，1) 是比 ARIMA(1，d，1) 更适合给定数据的模型。

ARMA(1,1) 的预测是无偏的，且随着提前期增加，预测误差方差会无限增大。对于非平稳序列，当我们对遥远的未来进行预测时，预测结果存在很大的不确定性。此外，根据正态Q-Q图、相关图和核密度估计图，我们可以认为 ARMA(1,1) 模型拟合得更好。此外，残差分析也进一步验证了这一结果。

这些结果仅适用于戈达德太空研究所地表温度（GISTEMP）数据。作者希望说明，本文中的分析和结果未必适用于其他数据。

数据可用性声明

数据的原始贡献取自 https://data.giss.nasa.gov/gistemp/graphs/，模拟使用 Python 完成。如有进一步疑问，请联系通讯作者。

作者贡献

所有列出的作者均为本研究付出了大量的智力努力。

致谢

作者感谢密西西比大学数学系副教授 Sang 博士开设时间序列分析课程，我们从中学习并获得了有关时间序列模型的统计学理念。

利益冲突 作者声明，本研究在不存在任何可能被视为潜在利益冲突的商业或财务关系的情况下开展。

参考文献

[1] Hansen, J.E. 和 S. Lebedeff，1987：实测地表气温的全球趋势。《地球物理研究杂志》，92，13345-13372，doi：10.1029/JD092iD11p13345。[2] Hansen, J., Sato, M., Ruedy, R., Kharecha, P., Lacis, A., Miller, R. 等（2007）。采用*GISS*模型对*1880*年—*2003*年的气候模拟。《气候动力学》，29（7-8），661-696。[3] Hansen, J., Ruedy, R., Sato, M., Lo, K.（2010）。全球地表温度变化。《地球物理学评论》，48，RG4004，https://doi.org/10.1029/2010RG000345

[4] Hawkins, E., Ortega, P., Suckling, E., Schurer, A., Hegerl, G., Jones, P. 等（2017）。估计工业化前时期以来的全球气温变化。《美国气象学会公报》，98（9），1841-1856，https://doi.org/10.1175/BAMS-D- 16-0007.1[5]Huang, B., Banzon, V. F., Freeman, E., Lawrimore, J., Liu, W., Peterson, T. C. 等（2015）。扩展重建海表温度第*4*版（*ERSSTv4*）。第一部分：升级与相互比较。《气候杂志》，28（3），911-930。https://doi.org/10.1175/JCLI-D-14-00006.1[6] Lenssen, N., G. Schmidt, J. Hansen, M. Menne, A. Persin, R. Ruedy 和 D. Zyss，2019：*GISTEMP*不确定性模型的改进。《地球物理研究杂志：大气》，124，第12期，6307-6326，doi：10.1029/2018JD029522。[7] Mahmud Hasan，（2014），设施布局规划问题中布局设计的新算法方法，硕士论文，数学系，贾汉吉尔纳加尔大学[8] Mahmud Hasan, AR Khan, N Ghosh, M Uddin - 2015，设施布局规划问题中布局设计算法的开发，《物理科学杂志》，维迪亚萨加尔大学，第20卷，35-42[9] Mahmud Hasan, S Hossain, MK Ahamed, MS Uddin，（2017），使用模糊*TOPSIS*法选择有效电子设备的可持续方式，American Academic Scientific Research Journal for，35（1），342-351[10] Menne, M. J., Williams, C. N., Gleason, B. E., Rennie, J. J., Lawrimore, J. H.（2018）。全球历史气候网络月度温度数据集，第*4*版。《气候杂志》，31，9835-9854。https://doi.org/10.1175/jcli-d-18-0094.1[11]Menne, M. J., Williams, C. N., Palecki, M. A.（2010）。关于美国地表温度记录的可靠性。《地球物理研究杂志》，115，D11108。https://doi.org/10.1029/2009jd013094[12] Morice, C. P., Kennedy, J. J., Rayner, N. A., Jones, P. D.（2012）。使用观测估计值集合量化全球和区域气温变化中的不确定性：*HadCRUT4*数据集。《地球物理研究杂志》，117，D08101[13] Robert H. Shumway, Davids Stoffer，《时间序列分析及其应用：*R*示例》，第*2*版。施普林格统计学丛书。[14] Robert A. Rohde1, Zeke Hausfather 等（2021），伯克利地球陆地/海洋温度记录[15] Freeman, E., Woodruff, S. D., Worley, S. J., Lubker, S. J., Kent, E. C., Angel, W. E. 等（2016）。*ICOADS*第*3.0*版发布：历史海洋气候记录的重大更新。《国际气候学杂志》，*37*（*5*）。[16] Mushtaq, Rizwan，增广迪基-富勒检验（2011年8月17日）。可从SSRN获取：https://ssrn.com/abstract=1911068 或 http://dx.doi.org/10.2139/ssrn.1911068

[17] Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell, E. C. Kent 和 A. Kaplan，2003年：19世纪末以来海表温度、海冰和夜间海洋气温的全球分析。《地球物理研究杂志》，108，4407，https://doi.org/10.1029/2002JD002670。[18] Quattrochi, D.A.；Luvall, J.C.《地表处理中的热红外遥感》；CRC出版社：美国佛罗里达州博卡拉顿，2004年。[19] Huang, B.、Thorne, P. W.、Smith, T. M.、Liu, W.、Lawrimore, J.、Banzon, V. F. 等（2016年）。进一步探讨和量化扩展重建海表温度（ERSST）第4版（v4）的不确定性。《气候杂志》，29（9），3119Ã¢3142。[20]Menne, M. J.、Williams, C. N.、Gleason, B. E.、Rennie, J. J.、Lawrimore, J. H.（2018年）。全球历史气候网络月度气温数据集，第4版。《气候杂志》，31，9835Ã¢9854。