

# 基于相关分析和回归分析的北冰洋海冰面积波动研究

## 摘 要

北极是全球气候的驱动器之一，作为全球气候的冷源，北极海冰变化影响到全球尺度的气候系统。本文通过时间序列分解、Pearson 相关性分析和线性回归分析等方法来探究北极海冰变化与全球尺度的气候系统的关系。

针对问题一，本文通过时间序列分解对北极各地区的海冰面积随时间变化进行分解，得到北极各区域的海冰面积变化的波动性、周期性。使用 Mann-Kendall 检验方法分析各地区海冰面积在时间上的变化趋势呈整体下降趋势。

针对问题二，本文分别按每年和每月的北极海冰平均面积与全球平均气温进行 Pearson 相关性分析，皆得到二者具有强相关性且呈负相关，再通过对每年和每月的北极海冰平均面积与全球平均气温进行一元线性回归分析，进一步印证了北极海冰平均面积与全球平均气温呈负相关。最终得到结论：北极海冰面积变化和气候变化的相互关系为负相关。

针对问题三，本文对北极每月海冰面积进行时间序列分解，得到北极海冰面积变化趋势数据，然后对北极海冰面积变化趋势和全球二氧化碳浓度进行 Pearson 相关性分析，得到二者具有极强相关性且呈负相关，再对北极海冰面积变化趋势和全球二氧化碳浓度进行一元线性回归分析，进一步印证了北极海冰面积变化趋势和全球二氧化碳浓度呈负相关。最终得到结论：海冰面积变化趋势与二氧化碳温室效应的关系为负相关。

针对问题四，本文使用 Pearson 相关性分析分别分析温度，海冰面积，二氧化碳浓度和海平面高度的相关性，选取对海平面高度显著水平高的因子：海冰面积和二氧化碳浓度，然后对全球平均海平面高度、北极海冰面积和二氧化碳浓度进行多元线性回归分析，得到回归方程，接着根据布容法则计算得海岸线被侵蚀距离和海平面高度的关系式，呈正相关。最终得到结论：北极海冰面积变化和地球陆地面积变化的关系为负相关。

针对问题五，本文根据文献和数据选取北极海冰面积和全球海平面高度作为气候变暖的危害性指标，使用 Pearson 相关性分析分别分析出与北极海冰面积和海平面高度相关性显著的因子，然后使用逐步回归的逐步法对其进行分析，得到北极海冰面积与全球二氧化碳浓度以及全球海平面高度和全球二氧化碳浓度的回归方程。接着根据北极海冰面积和全球二氧化碳浓度呈负相关、全球海平面高度和全球二氧化碳浓度呈正相关的结论，提出如何减少北极海冰融化和降低海平面高度来降低气候变暖的危害的相关政策。

**关键词：**时间序列分析 MK 检验 Pearson 相关性分析 线性回归分析 逐步回归分析

## 一、 问题介绍

### 1.1 研究背景

全球变暖对某些地区的影响将比其他地区更为显著，而北极正是气候变化的重点地区之一。近年来，北极海冰的面积不断创下新低，并给生活在这里的人类和动植物造成了巨大的影响。极地海冰作为冷源对大气环流和全球气候的影响极其重要，因此如何对北冰洋的海冰面积波动进行定量分析，已经成为研究气候变化的不利因素的危害的一个重要课题。

### 1.2 问题重述

#### 问题一：

基于北极各个区域的海冰面积变化的有关数据，建立数学模型，研究北极各个区域的海冰面积变化的规律，分析其波动性、周期性和变化规律。

#### 问题二：

基于气候变化的相关数据，建立数学模型，研究海冰面积变化和气候变化的相互关系。

#### 问题三：

基于二氧化碳的温室效应的相关数据，建立数学模型，研究海冰面积变化的趋势与二氧化碳的温室效应的相互关系。

#### 问题四：

建立数学模型，研究北极海冰面积变化对地球陆地面积的影响。

#### 问题五：

设置合适的指标，建立数学模型说明人类要改变现有的政策以减小气候变暖的危害。

## 二、 问题分析

#### 问题一分析：

通过时间序列分解对北极各地区的海冰面积随时间变化进行分解，得到北极各区域的海冰面积变化的波动性、周期性。使用 Mann-Kendall 检验方法分析各地区海冰面积在时间上的变化趋势。

#### 问题二分析：

分别按每年和每月的北极海冰平均面积与全球平均气温进行 Pearson 相关性分析，得到二者的相互关系，再通过对每年和每月的北极海冰平均面积与全球平均气温进行一元线性回归分析，进一步探究二者关系，最终得到北极海冰面积变化和气候变化的相互

关系。

### 问题三分析：

对北极每月海冰面积进行时间序列分解，得到北极海冰面积变化趋势数据，然后对北极海冰面积变化趋势和全球二氧化碳浓度进行 Pearson 相关性分析，得到二者关系，再对北极海冰面积变化趋势和全球二氧化碳浓度进行一元线性回归分析，进一步探究二者关系。最终得到海冰面积变化趋势与二氧化碳温室效应的关系。

### 问题四分析：

使用 Pearson 相关性分析分别分析温度，海冰面积，二氧化碳浓度和海平面高度的相关性，选取对海平面高度显著水平高的因子，然后对全球平均海平面高度和选取的因子进行多元线性回归分析，得到回归方程，接着根据布容法则计算得海岸线被侵蚀距离和海平面高度的关系式。最终得到北极海冰面积变化和地球陆地面积变化的关系。

### 问题五分析：

根据文献和数据选取适当指标作为气候变暖的危害性指标，使用 Pearson 相关性分析分别分析出与指标相关性显著的因子，然后使用逐步回归的逐步法对其进行分析，得到指标与主要影响指标的回归方程。接着根据得到的回归方程，提出如何降低气候变暖的危害的相关政策。

## 三、模型假设

- (1) 假设提供的实测数据是准确的，能较真实地反映实际情况。
- (2) 仅考虑问题中的核心因素（全球气温、 $CO_2$  的温室效应和北冰洋海冰面积），不考虑次要因素的影响（人口膨胀和对自然资源的过度开发）。
- (3) 假设回归模型中的扰动项服从独立的正态分布。
- (4) 假设回归模型具有同方差性，即误差项  $\varepsilon$  的方差为常数。

## 四、符号说明

符号	文字说明	单位
S	北冰洋海冰面积	$10^6 km^2$
Y	北冰洋海冰面积变化趋势	/
T	全球气温	$^{\circ}C$
C	$CO_2$ 浓度	ppm
H	全球海平面高度	/
R	海岸线被侵蚀的距离	/
r	相关系数	/
p	显著性水平判断因子	/
$R^2$	曲线回归的拟合程度	/

## 五、问题一：基于时间序列分解的海冰面积变化模型

### 5.1 数据来源

北冰洋海冰面积数据来自美国雪冰数据中心的 EASE2-GRID 北半球的周积雪和海冰面积数据集，包含 1978 年至 2022 年的数据。

### 5.2 模型建立

#### 5.2.1 时间序列分解的乘法模型

北极的海冰面积变化具有一定的规律，可以通过时间序列分解<sup>[1]</sup>的乘法模型来进行研究 1978 年至 2022 年各月份海冰面积的总体趋势和其波动性周期性。

时间序列可以表示为 4 个因素的函数，这四个因素分别为：长期趋势  $T_t$ 、季节变动  $S_t$ 、循环波动  $C_t$ 、不规则波动  $I_t$ 。

- 长期趋势：较长时期内的一种趋势。
- 季节波动：季节变化引起的变动。
- 循环波动：以若干年为期限，不具规则的周期性变动。
- 不规则波动：由于众多偶然因素对时间序列造成的影响。

因此时间序列的函数可表示为：

$$S_t = f(T_t, S_t, C_t, I_t) \quad (1)$$

其函数具有以下模型：

- (1) 乘法模型： $S_t = T_t * S_t * C_t * I_t$ ，其中季节指数和循环指数都是比例数，且  $I_t \sim N(\mu, \sigma^2)$ 。
- (2) 加法模型： $S_t = T_t + S_t + C_t + I_t$ ，其中五个变量都有相同的量纲，且  $I_t \sim N(\mu, \sigma^2)$ 。
- (3) 混合模型： $S_t = T_t * S_t + C_t + I_t$ ，其中除  $S_t$  季节指数为比例数外，其他变量都有相同的量纲，且  $I_t \sim N(\mu, \sigma^2)$ 。

#### 5.2.2 MK(Mann-Kendall) 检验方法

为分析海冰面积在时间上的变化趋势、不同阶段的变化特征、长时间尺度上其变化是否存在突变，以及这些变化的显著性，选择采用 MK(Mann-Kendall) 检验方法<sup>[2]</sup>。

对设定的  $S_1, S_2, \dots, S_n$  时间序列变量， $n$  为时间序列的长度，建立一秩序列  $A_k$ ：

$$A_k = \sum_{i=1}^k r_i \quad (k = 1, 2, \dots, n) \quad (2)$$

其中

$$r_i = \begin{cases} 1, & x_i > x_j \\ 0, & x_i \leq x_j \end{cases} \quad (3)$$

当时间序列独立并且随机的情况下，设定统计量：

$$UF_k = \frac{A_k - E(A_k)}{\sqrt{Var(A_k)}} \quad (k = 1, 2, \dots, n) \quad (4)$$

其中  $UF_1 = 0$ ， $E(A_k)$ 、 $Var(A_k)$  分别为秩序列  $A_k$  的方差和均值，公式如下：

$$\begin{cases} E(A_k) = \frac{k(k+1)}{4} \\ Var(A_k) = \frac{k(k-1)(2k+5)}{72} \end{cases} \quad (5)$$

$UF_k$  标准正态分布，设显著性水平为  $p$ ，若  $|UF_k| > U_p$ ，则说明该序列的变化趋势具有显著性。对  $x$  这一时间序列进行逆向排序，然后再计算逆序列的  $UF_k$ ，则逆向排序序列的  $UB_k$  可以表示为：

$$UB_{n-i+1} = -UF_i \quad (i = 1 - n) \quad (6)$$

对  $UF_k$  和  $UB_k$  序列进行再统计分析能够得到序列  $x$  的变化趋势和突变发生的时间刻。如果计算得到的  $UF_k$  或  $UB_k$  的值大于 0，则表明该序列具有增加的趋势；否则表示具有下降的趋势。

如果两条曲线  $UF_k$  和  $UB_k$  在某时间点上具有交点，且该交点在某置信区间的上下临界线之间，则说明该交点代表的时刻在某一置信区间内是显著的突变发生的时间点。

### 5.3 模型求解及结果分析

选择使用乘法模型来对各个地区的海冰面积随时间变化进行分解。例如，对于 Baffin 扩展地区，有分解：

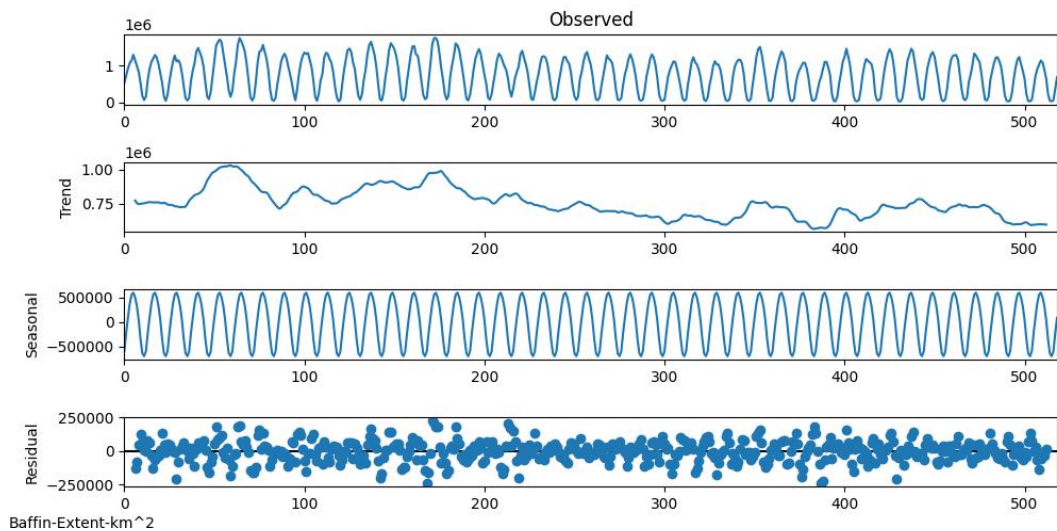


图 1 Baffin 扩展地区海冰面积的时间序列分解图

其中第一幅子图为原观测数据，可以看出来，具有一定的波动性，但是又包含波形的局部上下偏移和不光滑凸起。进行分解后，可以依次得到趋势（Trend），季节性（Seasonal）和残差（Pesidual），分别对应子图二至四。由子图二可见，海冰面积有明显的下降趋势，而对于子图三，则可以看出原数据的波动性以及周期性。

例如对于地区 Baffin，假设：

- H0：对于给定数据，其中没有趋势。
- H1：对于给定数据，其中有趋势。

进行 MK 检验后得出结果  $p = 0.0001$ ，由此可拒绝 H0，接受 H1。并且得出  $z = -3.8039$ ，可以判断趋势向下。

北极其他地区海冰面积变化的波动性、周期性和变化趋势分析见附录。

## 六、问题二：北冰洋海冰面积与全球气温的一元线性回归模型

### 6.1 数据来源

气温变化是气候变化的重要指标，全球气温数据来自英国的 HadCrut、美国的 NASA(航空航天局) 和 NOAA(海洋和大气管理局)，包含 1750 年至 2020 年的数据。

### 6.2 模型建立

#### 6.2.1 相关性分析

对北冰洋海冰面积  $S$  与全球气温  $T$  之间进行相关分析<sup>[3]</sup>，可以表明其内在相互联系的程度大小，通过对皮尔逊相关系数的计算与检验来实现，具体分析步骤如下：

- (1) 计算它们之间不存在相关性的概率  $p_{ST}$  值，判断它们之间相关关系的显著性水平。  
若  $p_{ST} < 0.05$ ，则说明两变量之间存在显著的线性关联，其中  $p_{ST}$  值越小，显著性水平越高。
- (2) 计算它们的相关系数：

$$r_{ST} = \frac{\sum_{i=1}^n (S_i - \bar{S})(T_i - \bar{T})}{\sqrt{\sum_{i=1}^n (S_i - \bar{S})^2 (T_i - \bar{T})^2}} \quad (7)$$

其中  $\bar{S}_i$ 、 $\bar{T}_i$  分别为北冰洋海冰面积  $S$  和全球气温  $T$  的样本平均值， $r_{ST} \in [-1, 1]$ 。

- (3) 根据相关系数分析相关的正负向和相关性程度。当  $r_{ST}$  大于 0 时，说明  $S$  与  $T$  呈正相关关系；若其小于 0，则说明  $S$  与  $T$  呈负相关关系。 $r_{ST}$  向下越靠近 -1 或向上越靠近 1，那么说明  $S$  与  $T$  之间的关系越紧密；若  $r_{ST}$  的值越趋向于 0，则说明  $S$  与  $T$  之间的关系越小。

#### 6.2.2 一元线性回归

当变量之间存在着显著的相关关系时，可以利用一定的数学模型对其进行回归分析。因此建立北冰洋海冰面积  $S$  与全球气温  $T$  的一元线性回归模型<sup>[4]</sup>。该模型假定因变

量 S 主要受自变量 T 的影响，它们之间存在着近似的线性函数关系，即：

$$S = \alpha_0 + \alpha_1 T + \varepsilon_1 \tag{8}$$

其中  $\alpha_0$  和  $\alpha_1$  是回归系数， $\varepsilon_1$  是随机误差。 $R_{ST}^2$  代表曲线回归的拟合程度，越接近 1 效果越好，公式如下：

$$R_{ST}^2 = 1 - \frac{\sum_i (\hat{S}_i - S_i)^2}{\sum_i (S_i - \bar{S})^2} \tag{9}$$

其中  $S_i$  为真实值， $\hat{S}_i$  为预测值， $\bar{S}$  为样本平均值。

6.3 模型求解及结果分析

6.3.1 北冰洋月平均海冰面积与全球月平均气温的相互关系

将北冰洋月平均海冰面积数据与全球月平均气温数据按照日期绘制折线图进行比较。

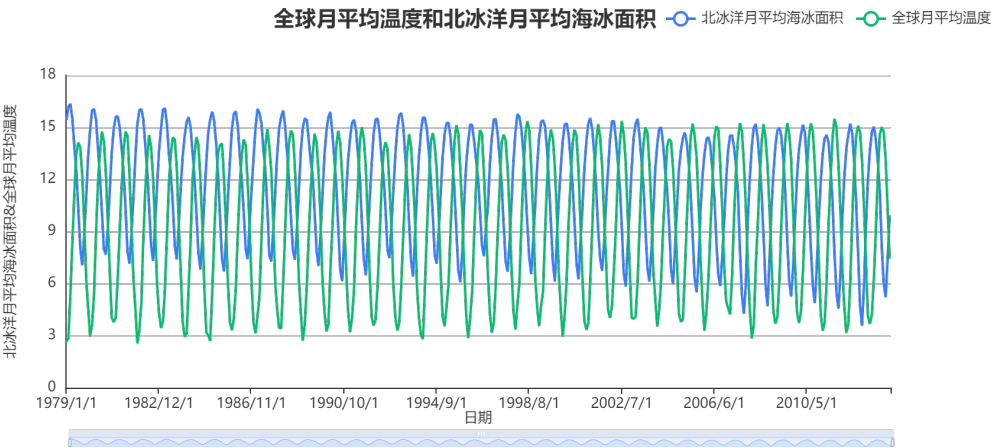


图 2 全球月平均温度与北冰洋月平均海冰面积比较

由图大致判断北冰洋月平均海冰面积与全球月平均气温具有负相关性，计算二者相关系数为  $r_{月ST} = -0.668$ ， $p_{月ST} = 0.0003$ ，为强相关且显著性水平高，初步推断北冰洋月平均海冰面积与全球月平均气温呈负相关关系。

采用最小二乘法对全球月平均气温数据和北冰洋月平均海冰面积进行线性回归拟合，结果如下：

线性回归分析结果 n=419									
	非标准化系数		标准化系数	t	p	VIF	R²	调整R²	F
	B	标准误	Beta						
常数	16.346	0.284	-	57.597	0.000***	-	0.445	0.444	F=334.863 P=0.000***
全球月平均温度	-0.513	0.028	-0.667	-18.299	0.000***	1.000			
因变量：北冰洋月平均海冰面积									

图 3 北冰洋月平均海冰面积对全球月平均气温数据线性回归分析结果表

$$S_{\text{月}} = 16.323 - 0.515T_{\text{月}} \quad (10)$$

拟合效果如图：

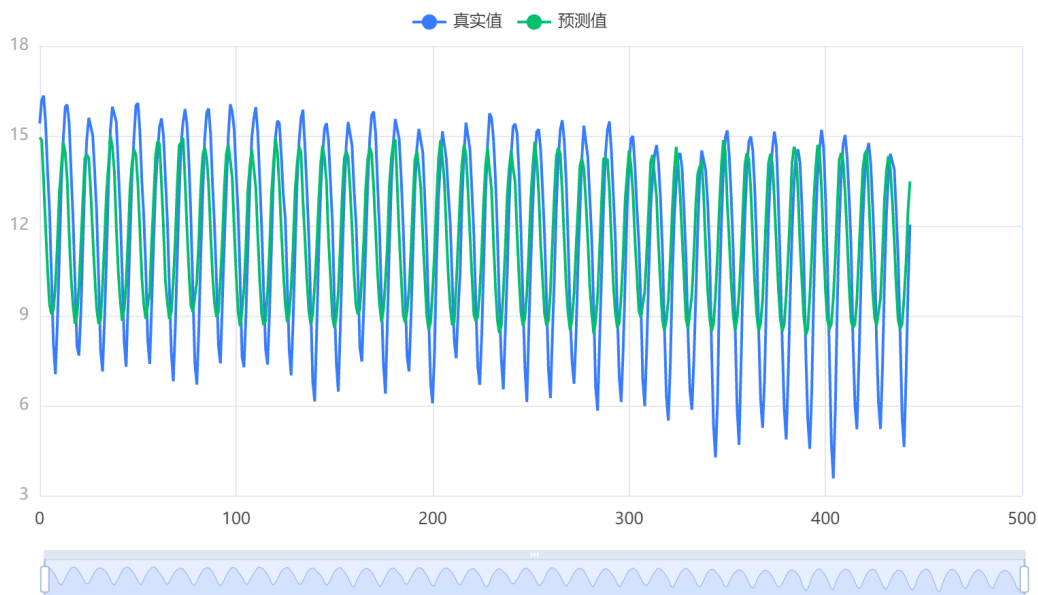


图4 北冰洋月平均海冰面积对全球月平均温度的拟合曲线

得北冰洋月平均海冰面积与全球月平均气温呈负相关关系。

### 6.3.2 北冰洋年平均海冰面积与全球年平均气温的相互关系

将北冰洋年平均海冰面积数据与全球年平均气温数据按照日期绘制折线图进行比较。

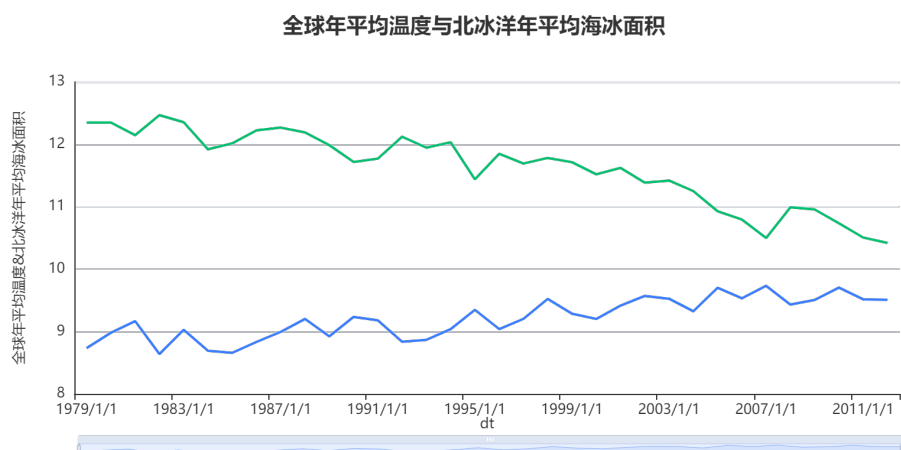


图5 全球年平均温度与北冰洋年平均海冰面积比较

由图大致判断北冰洋年平均海冰面积与全球年平均气温具有负相关性，计算二者相关系数  $r_{\text{年}ST} = -0.830$ ， $p_{\text{年}ST} = 0.0002$ ，为极强相关且显著性水平高，初步推断北冰洋年平均海冰面积与全球年平均气温呈负相关关系。。



采用最小二乘法对全球年平均气温数据和北冰洋年平均海冰面积进行线性回归拟合，结果如下：

线性回归分析结果 n=34									
	非标准化系数		标准化系数	t	p	VIF	R²	调整R²	F
	B	标准误	Beta						
常数	25.859	1.689	-	15.314	0.000***	-	0.69	0.68	F=71.106 P=0.000***
全球年平均温度	-1.546	0.183	-0.83	-8.432	0.000***	1.000			
因变量：北冰洋年平均海冰面积									

图 6 北冰洋年平均海冰面积对全球年平均气温数据线性回归分析结果表

$$S_{\text{年}} = 25.859 - 1.546T_{\text{年}} \quad (11)$$

拟合效果如图：



图 7 北冰洋年平均海冰面积对全球年平均温度的拟合曲线

得北冰洋年平均海冰面积与全球年平均气温呈负相关关系。

### 6.3.3 结果分析

根据北冰洋月平均海冰面积与全球月平均气温呈负相关关系和北冰洋年平均海冰面积与全球年平均气温呈负相关关系，综合得出北冰洋海冰面积变化和气候变化呈负相关关系：当气候变暖时，北冰洋海冰面积减少。当北冰洋海冰面积增加时，气候回冷。

## 七、问题三：海冰面积变化趋势与大气 $CO_2$ 浓度的一元线性回归模型

### 7.1 数据来源

大气  $CO_2$  浓度数据来自莫纳罗亚天文台，包含 1983 年至 2014 年的数据。

### 7.2 模型建立

#### 7.2.1 数据预处理

为研究海冰面积变化趋势和二氧化碳温室效应<sup>[5]</sup>的关系，本文对两者的数据进行清洗，排除其白噪声与季节性变化对其数据的影响。使用时间序列对清洗后的数据进行分解，提取其中的北冰洋海冰面积变化趋势数据。

#### 7.2.2 相关性分析

对北冰洋海冰面积变化趋势  $Y$  与大气  $CO_2$  浓度  $C$  之间进行相关分析，具体分析步骤如下：

- (1) 计算它们之间不存在相关性的概率  $p_{YC}$  值，判断它们之间相关关系的显著性水平。若  $p_{YC} < 0.05$ ，则说明两变量之间存在显著的线性关联，其中  $p_{YC}$  值越小，显著性水平越高。
- (2) 计算它们的相关系数：

$$r_{YC} = \frac{\sum_{i=1}^n (Y_i - \bar{Y})(C_i - \bar{C})}{\sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2 (C_i - \bar{C})^2}} \quad (12)$$

其中  $\bar{Y}_i$ 、 $\bar{C}_i$  分别为北冰洋海冰面积  $S$  和大气  $CO_2$  浓度的样本平均值， $r_{YC} \in [-1, 1]$ 。

- (3) 根据相关系数分析相关的正负向和相关性程度。当  $r_{YC}$  大于 0 时，说明  $Y$  与  $C$  呈正相关关系；若其小于 0，则说明  $Y$  与  $C$  呈负相关关系。 $r_{YC}$  向下越靠近 -1 或向上越靠近 1，那么说明  $Y$  与  $C$  之间的关系越紧密；若  $r_{YC}$  的值越趋向于 0，则说明  $Y$  与  $C$  之间的关系越小。

#### 7.2.3 一元线性回归

建立北冰洋海冰面积变化趋势  $Y$  与大气  $CO_2$  浓度  $C$  的一元线性回归模型，该模型假定因变量  $Y$  主要受自变量  $C$  的影响，它们之间存在着近似的线性函数关系，即：

$$Y = \beta_0 + \beta_1 C + \varepsilon_2 \quad (13)$$

其中  $\beta_0$  和  $\beta_1$  是回归系数， $\varepsilon_2$  是随机误差。 $R_{YC}^2$  代表曲线回归的拟合程度，越接近 1 效果越好，公式如下：

$$R_{YC}^2 = 1 - \frac{\sum_i (\hat{Y}_i - Y_i)^2}{\sum_i (Y_i - \bar{Y})^2} \quad (14)$$

其中  $Y_i$  为真实值,  $\hat{Y}_i$  为预测值,  $\bar{Y}$  为样本平均值。

### 7.3 模型求解及结果分析

#### 7.3.1 北冰洋月平均海冰面积变化和全球每月 $CO_2$ 浓度的关系

分别绘制北冰洋月平均海冰面积变化和全球每月  $CO_2$  浓度随日期变化的折线图：



图 8 北冰洋月平均海冰面积变化趋势

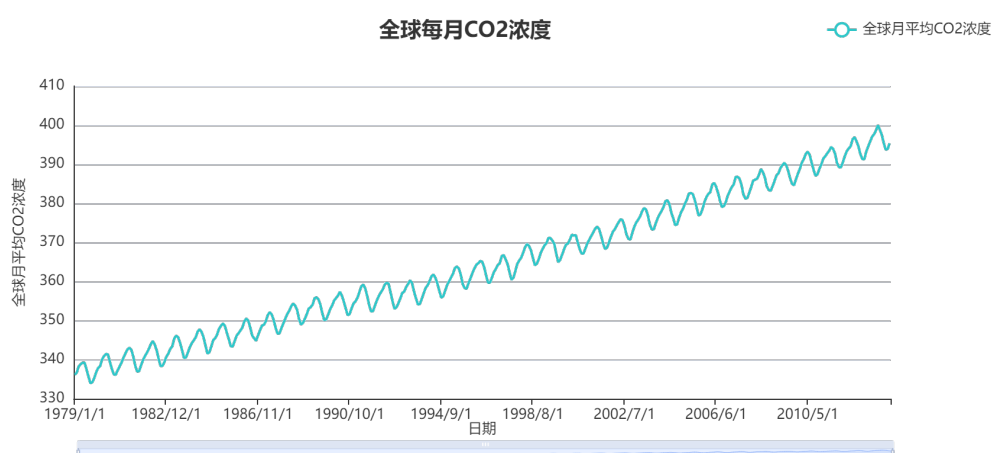


图 9 全球月  $CO_2$  浓度

由图大致判断北冰洋月平均海冰面积与全球月  $CO_2$  浓度具有负相关性, 计算二者的相关系数  $r_{YC} = -0.953$ ,  $p_{YC} = 0.0001$ , 为极强相关且显著性水平高, 故初步推断北冰洋月平均海冰面积与全球月  $CO_2$  浓度呈负相关性。

采用最小二乘法对北冰洋月平均海冰面积与全球月  $CO_2$  浓度进行线性回归拟合, 结果如下：

线性回归分析结果 n=440									
	非标准化系数		标准化系数	t	p	VIF	R²	调整R²	F
	B	标准误	Beta						
常数	22.822	0.172	-	132.579	0.000***	-	0.908	0.907	F=4300.01 P=0.000***
月平均CO2浓度	-0.031	0	-0.953	-65.574	0.000***	1.000			
因变量: 月平均海冰变化趋势									

图 10 北冰洋月平均海冰面积对全球月  $CO_2$  浓度线性回归分析结果表

$$Y = 22.822 - 0.031C \quad (15)$$

拟合效果如图：

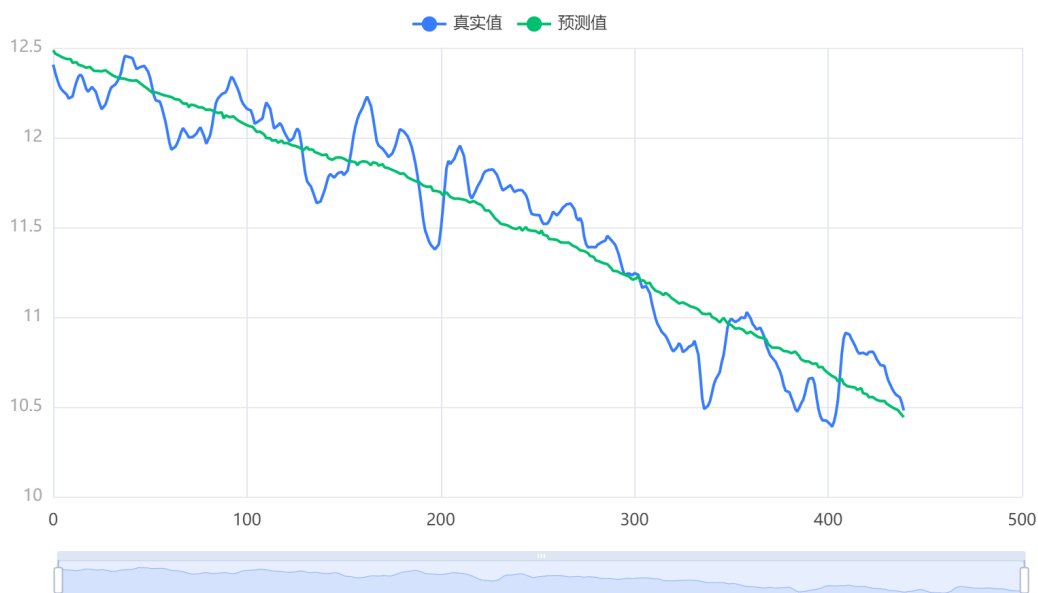


图 11 北冰洋月平均海冰面积变化趋势对全球月  $CO_2$  浓度的拟合曲线

得北冰洋海冰面积变化趋势对全球  $CO_2$  浓度呈负相关性：全球月  $CO_2$  浓度升高时，北冰洋海冰面积具有向下减少的趋势。

### 7.3.2 结果分析

温室效应：二氧化碳气体具有吸热和隔热的功能。大气中的二氧化碳浓度的不断增加，阻止了地球热量的散发，如同给地球盖上了保温毯，引起地表温度上升。这就是“温室效应”。

结合得北冰洋海冰面积变化趋势对全球  $CO_2$  浓度呈负相关性得海冰面积变化趋势与二氧化碳温室效应具有负相关性：当二氧化碳温室效应加剧时北冰洋海冰面积具有向下减少的趋势。

## 八、 问题四：全球海平面上升的多元线性回归模型

### 8.1 数据来源

全球平均绝对海平面高度数据来自美国环境保护署、美国国家海洋和大气管理局，包含 1880 年至 2014 年的数据。

### 8.2 模型建立

#### 8.2.1 流程图

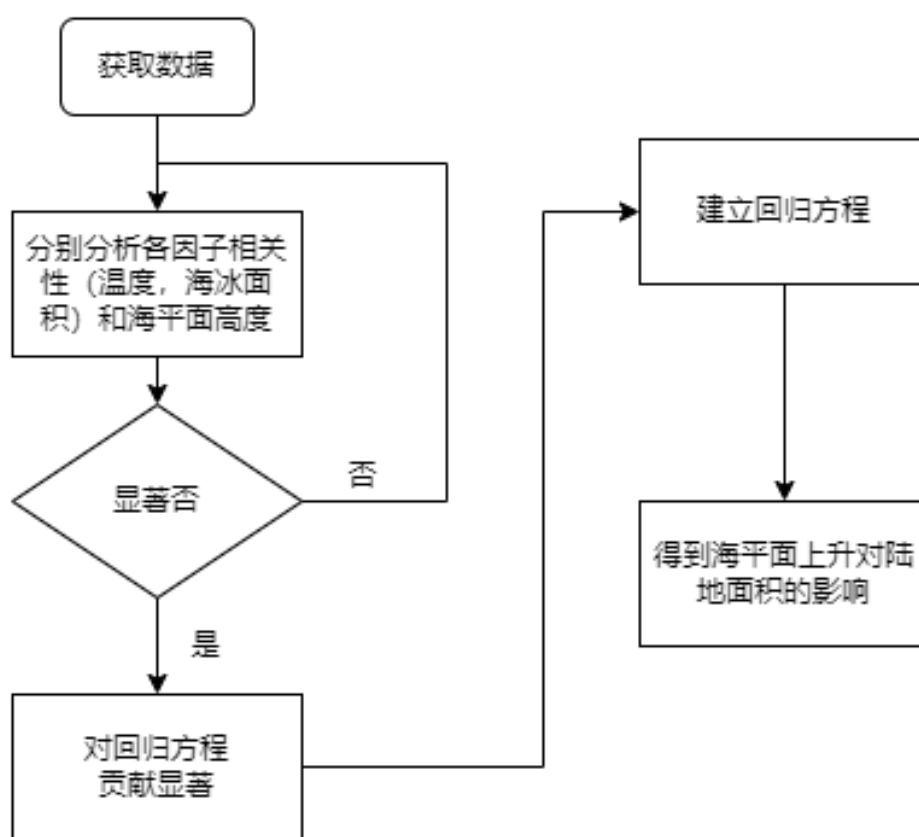


图 12 对海平面高度进行多元线性回归分析的流程图

#### 8.2.2 海平面上升对海岸线的侵蚀

北极海冰面积变化对地球陆地面积的影响具体可以反映为海平面高度的上升。当海平面上升时，它会产生两方面的影响。首先，部分较低的陆地和河岸会被海水淹没、侵占，海岸线会向内陆侵蚀。其次，随着海平面上升，海洋的波浪作用也可能对海岸造成巨大的额外侵蚀。

根据布容法则<sup>[6]</sup>，计算被侵蚀的海岸线距离：

$$R = \frac{L}{b+h}H \quad (16)$$

其中  $R$  表示海岸线被侵蚀的距离， $b$  表示海边沙滩的高度， $h$  表示近海的深度， $L$  表示从海滩到深度为  $h$  的近海的横向距离， $H$  为海平面高度。一般情况下，公式可以简化为：

$$R = 50H \text{ 或 } R = 100H \quad (17)$$

### 8.2.3 全球海平面上升和海岸线侵蚀距离的多元线性回归模型

众所周知， $CO_2$  等温室气体的排放以及海冰面积变化对全球气温的升高具有重要影响。同时，全球气温的升高和海冰面积变化对海平面的升高也具有重要的影响。因此，需要根据这些因素之间的相关性选取对海平面上升影响最显著的因素：

- 全球年平均温度与全球年平均海平面高度的相关性。
- 全球年平均温度与全球年平均大气  $CO_2$  浓度的相关性。
- 北冰洋年平均海冰面积与全球年平均温度的相关性。
- 北冰洋年平均海冰面积与全球年平均海平面高度的相关性。

相关性分析结果如下：

表 2 各因素之间的相关系数

因素	全球年平均温度
全球年平均海平面高度	0.818
全球年平均大气 $CO_2$ 浓度	0.847
因素	北冰洋年平均海冰面积
全球年平均温度	-0.830
全球年平均海平面高度	-0.922

由上述的相关因子分析，全球气温  $T$  受到大气  $CO_2$  浓度  $C$  和北冰洋海冰面积  $S$  的影响。因此，这里首先建立全球温度变化的多元线性回归模型：

$$T = a_0 + a_1S + a_2C + \varepsilon_3 \quad (18)$$

全球海平面的高度  $H$  受到北冰洋海冰面积  $S$  和全球气温变化  $T$  的直接影响。类似地，可以建立全球海平面高度的多元线性回归模型：

$$H = b_0 + b_1S + b_2T + \varepsilon_4 \quad (19)$$

结合全球温度变化的多元线性回归模型和全球海平面高度的多元线性回归模型可以得出全球海平面高度受北冰洋海冰面积  $S$  和大气  $CO_2$  浓度  $C$  影响的多元线性回归模型：

$$H = \gamma_0 + \gamma_1 S + \gamma_2 C + \varepsilon_5 \quad (20)$$

### 8.3 模型求解及结果分析

#### 8.3.1 全球海平面高度的多元线性回归

利用最小二乘法进行多元线性回归拟合，得到全球海平面高度受北冰洋海冰面积  $S$  和大气  $CO_2$  浓度  $C$  影响的多元线性回归模型：

$$H = -378.084 - 5.107S + 1.243C \quad (21)$$

$R^2 = 0.923$ ，拟合效果如图：

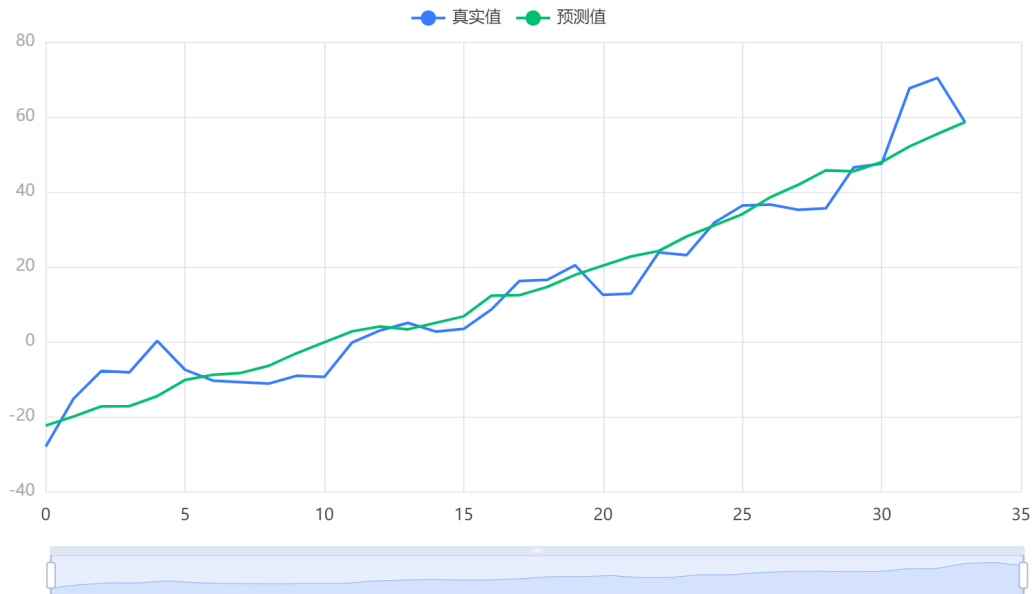


图 13 全球年平均海平面高度的拟合曲线

由于海平面上升对海岸线的侵蚀，可以得到相关模型：

$$R = -18904.20 - 255.35S + 62.15C \text{ 或 } R = -37808.4 - 510.7S + 124.3C \quad (22)$$

#### 8.3.2 结果分析

综上所述，当北冰洋海冰面积减少，海平面上升，导致海水淹没较低的陆地，海岸线向内陆侵蚀，地球的陆地面积减少。

## 九、问题五：气候变暖危害相关指标的逐步回归模型

### 9.1 模型建立

考虑到气候变暖的危害有：海平面上升、冰川融化、暴风雨和水灾等极端天气现象。本文采用了海平面上升和北冰洋海冰面积作为指标建立模型。

结合问题二和问题三的相关性分析可知，影响北冰洋海冰面积的因素有全球气候变化和全球二氧化碳温室效应。而基于问题四的相关性分析，可以得到影响全球海平面高度的因素有全球平均温度、北冰洋海冰面积和全球大气  $CO_2$  浓度。

故本文分别建立二者对应的逐步回归模型采用逐步法进行分析。

**定义 1 (逐步回归)** 以线性回归为基础的方法。其思路是将变量一个接着一个引入，并在引入一个新变量后，对已入选回归模型的旧变量逐个进行检验，将认为没有意义的变量删除，直到没有新变量引入也没有旧变量删除，从而保证回归模型中每一个变量都有意义。

**定义 2 (逐步法)** 结合向前法和向后法的优点，在向前引入每一个新自变量之后都要重新对已代入的自变量进行计算，以检验其有无继续保留在方程中的价值，并以此为依据进行自变量的引入和剔除交替进行，直到没有新的变量可以引入或剔除为止，得到的回归模型较为准确。

### 9.2 模型求解及结果分析

#### 9.2.1 北冰洋面积变化的逐步回归模型

逐步回归结果汇总	
方法	逐步
总变量情况	全球年平均温度、全球年平均 $CO_2$ 浓度
保留变量	全球年平均 $CO_2$ 浓度
舍弃变量	全球年平均温度

图 14 北冰洋海冰面积变化逐步回归结果汇总

利用逐步回归方法舍弃了变量全球年平均温度  $T$ ，保留了变量全球年平均  $CO_2$  浓度  $C$ ，得到北冰洋面积  $S$  的逐步回归模型如下：

线性回归分析结果n=34									
	非标准化系数		标准化系数	t	p	VIF	R <sup>2</sup>	调整R <sup>2</sup>	F
	B	标准误	Beta						
常数	23.677	0.745	0	31.775	0.000***	-	0.91	0.888	F=262.024, P=0.000***
全球年平均CO2浓度	-0.033	0.002	-0.944	-16.187	0.000***	1			
因变量：北冰洋年平均海冰面积									

图 15 北冰洋海冰面积逐步回归模型结果表



$$S = 23.677 - 0.033C \quad (23)$$

拟合效果如图：

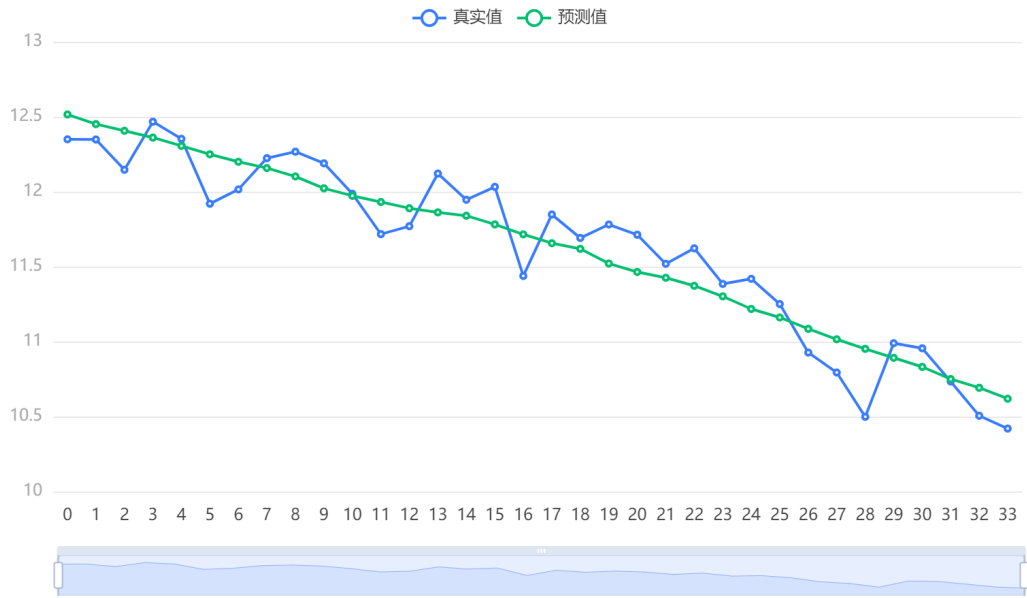


图 16 北冰洋海冰面积对  $CO_2$  浓度的拟合

北冰洋海冰面积与  $CO_2$  浓度呈负相关： $CO_2$  浓度升高，北冰洋海冰面积减少，即冰川融化加快。

### 9.2.2 全球海平面上升的逐步回归模型

逐步回归结果汇总	
方法	逐步
总变量情况	全球年平均温度、北冰洋年平均海冰面积、全球年平均 $CO_2$ 浓度
保留变量	全球年平均 $CO_2$ 浓度
舍弃变量	全球年平均温度、北冰洋年平均海冰面积

图 17 全球平均海平面高度逐步回归结果汇总

利用逐步回归方法舍弃了变量全球平均温度  $T$  和北冰洋海冰面积  $S$ ，保留了变量全球平均  $CO_2$  浓度  $C$ ，得到全球平均海平面高度  $H$  的逐步回归模型如下：

线性回归分析结果n=34									
	非标准化系数		标准化系数	t	p	VIF	R²	调整R²	F
	B	标准误	Beta						
常数	-498.996	25.61	0	-19.484	0.000***	-	0.926	0.924	F=403.028, P=0.000***
全球年平均CO2浓度	1.412	0.07	0.963	20.076	0.000***	1			
因变量: 全球年平均海平面高度									

图 18 全球平均海平面高度逐步回归模型结果表

$$H = -498.996 + 1.412C \quad (24)$$

拟合效果如图：

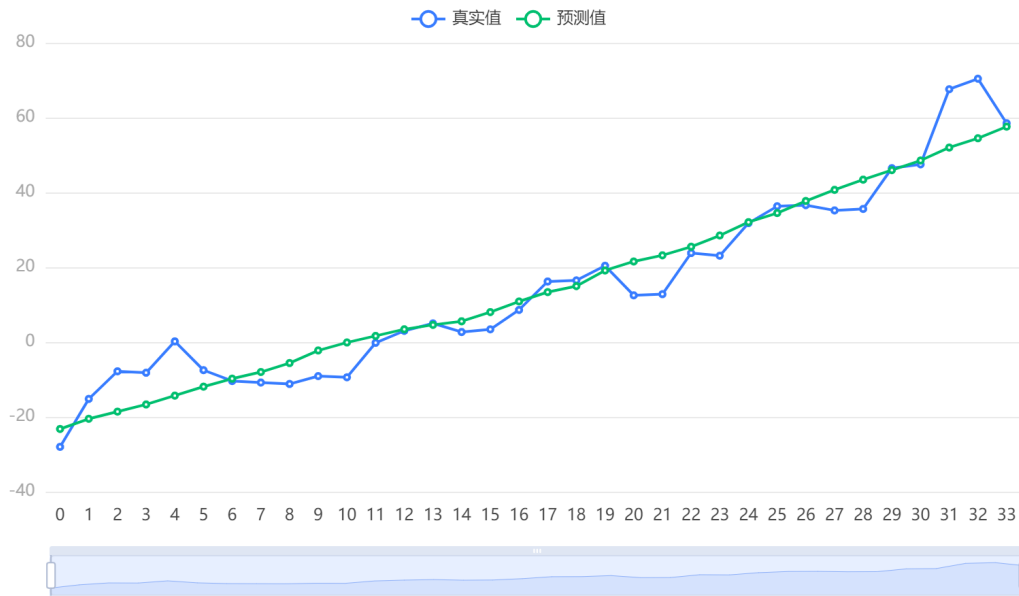


图 19 全球平均海平面高度对  $CO_2$  浓度的拟合

全球平均海平面高度与  $CO_2$  浓度呈正相关： $CO_2$  浓度升高，全球平均海平面高度上升。

### 9.2.3 结果分析

由上可知， $CO_2$  浓度导致气候变暖，从而使得北冰洋海冰融化加快，全球海平面升高。故人类要使有关气候变暖的危害性变小。就必须控制全球大气中  $CO_2$  的浓度。而人类的目前政策措施正是为了减少碳排放、实现碳中和，具体措施如下：

- (1) 控制温室气体排放，如减少化石燃料的使用量；提高能源的利用效率；大力发展新能源；控制水田和垃圾填埋场甲烷的排放量等。
- (2) 利用生态系统固碳，增加碳汇。森林是陆地生态系统的主体，在全球碳循环和调节气候变化中具有重要作用。据华龙网报道，利用植物的光合作用，可以提高生态系统的碳吸收和储存能力，从而减少二氧化碳在大气中的浓度。因此，增加森林总量、提高森林质量、增强森林生态服务功能等都可以减缓全球变暖的趋势。
- (3) 倡导低碳生活，绿色出行。人们可以用步行、骑车等低碳、绿色、环保的出行方式来减少二氧化碳的排放，这样做既可以保持个人的身体健康，又能达到节能减排的目的。

## 十、模型的评价与改进

### 10.1 模型评价

- (1) 本文在时间序列分解的基础上加入了 MK 检验使得北极海冰面积变化趋势更具有依据。
- (2) 本文在回归分析时采用了逐步回归分析, 得到具有较大影响程度因子的回归方程。

### 10.2 模型改进

- (1) 本文对于气候只考虑了全球平均温度, 还可以考虑大气环流和降雨量等因素, 也许可以得到更完美的回归方程。
- (2) 本文对于海平面高度采用了全球平均值, 实际上不同地区的海平面高度存在较大的差异性, 可以根据不同地区的海平面高度数据进行具体分析。

## 参考文献

- [1] 魏立新. 北极海冰变化及其气候效应研究 [D]. 中国海洋大学,2008.
- [2] 任艳群, 刘苏峡. 北半球积雪/海冰面积与温度相关性的差异分析 [J]. 地理研究, 2018, 37(5): 870-882
- [3] 魏立新, 张占海. 北极海冰变化特征分析 [J]. 海洋预报.2007,(4).42-48.
- [4] 郝春江, 解思梅, 李同娟. 北极海冰的变化特征 [J]. 海洋湖沼通报,1994,(0):84-91.
- [5] 刘诚程, 祝英杰, 王连双, 等. 大气中二氧化碳浓度对夏末北极海冰范围影响的分析 [J]. 智库时代,2021(5):268-269.
- [6] Nielsen David Marcolino,Pieper Patrick,Barkhordarian Armineh,Overduin Paul,Ilyina Tatiana,Brovkin Victor,Baehr Johanna,Dobrynin Mikhail.Increase in Arctic coastal erosion and its sensitivity to warming in the twenty-first century[J].Nature Climate Change Volume 12,Issue 3.2022.PP 263-270

## 十一、附录

### 11.1 Python 实现问题一

#### 11.1.1 时间序列分解

```
import matplotlib.pyplot as plt
import pandas as pd
import pymannkendall as mk
import numpy as np
from statsmodels.tsa.seasonal import seasonal_decompose
from pylab import rcParams
import os

# 局部
file = '/Users/smh/Desktop/数学建模集训/数模七月短学期集训模型3/N_seaice/各地区/
N_Sea_Ice_Index_Regional_Daily_Data_G02135_v3.0.xlsx'
sheetNames = ['Baffin-Extent-km^2', 'Barents-Extent-km^2', 'Beaufort-Extent-km^2',
               'Bering-Extent-km^2', 'CanadianArchipelago-Extent-km^2', 'Central-Arctic-Extent-km^2',
               'Chukchi-Extent-km^2', 'East-Siberian-Extent-km^2', 'Greenland-Extent-km^2',
               'Hudson-Extent-km^2', 'Kara-Extent-km^2',
               'Laptev-Extent-km^2', 'Okhotsk-Extent-km^2', 'St-Lawrence-Extent-km^2']
DOM = [-1, 31, 29, 31, 30, 31, 30, 31, 31, 30, 31, 30, 31]
startYear = 1978
endYear = 2022
for sheetName in sheetNames:
    data: list = list()
    data1: list = list()

    sheet = pd.read_excel(file, sheet_name=sheetName, usecols=np.arange(2, 46, 1))
    sheet = sheet.to_numpy()
    for i in range(endYear - startYear):
        presentYear = startYear + i
        for m in range(1, 13):
            presentMonth = m
            sIdx = int(np.sum(DOM[:m]) + 1)
            eIdx = int(np.sum(DOM[1:m+1]))
            ymSlice = sheet[sIdx:eIdx, i]
            avgM = np.nanmean(ymSlice)
            # print(presentYear, presentMonth, avgM)

            # ymData = np.insert(ymData, -1, values=[presentYear, presentMonth, avgM], axis=0)
            data.append(avgM)
            data1.append((presentYear, presentMonth, avgM))

    data = [d for d in data if not np.isnan(d)]
    res = seasonal_decompose(data, model='additive', period=12)
    rcParams['figure.figsize'] = 10, 5
    fig = res.plot()
    fig.text(0, 0, sheetName)
```

```

fig.show()
# save imgs
if not os.path.exists('./imgs'):
    os.makedirs('./imgs')
fig.savefig('./imgs/%s.jpg'%(sheetName))
fig.clf()

```

### 11.1.2 北冰洋相关地区的海冰面积变化的时间序列分解图

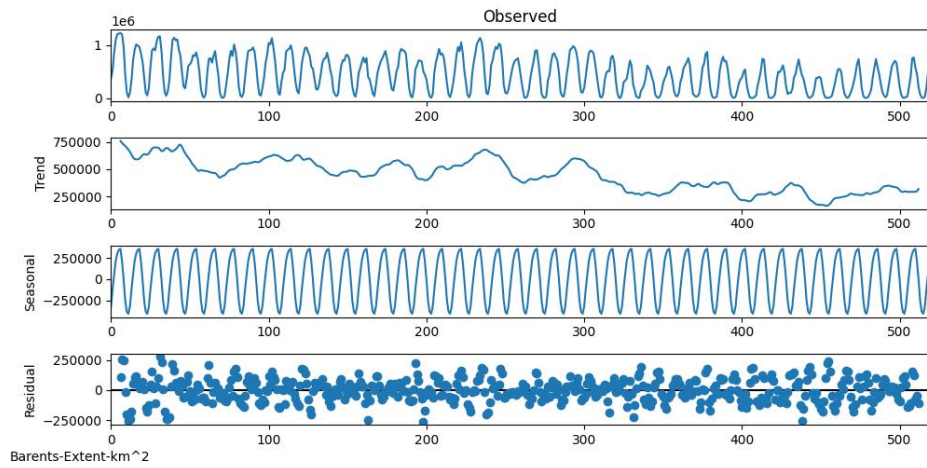


图 20 Barents 扩展地区

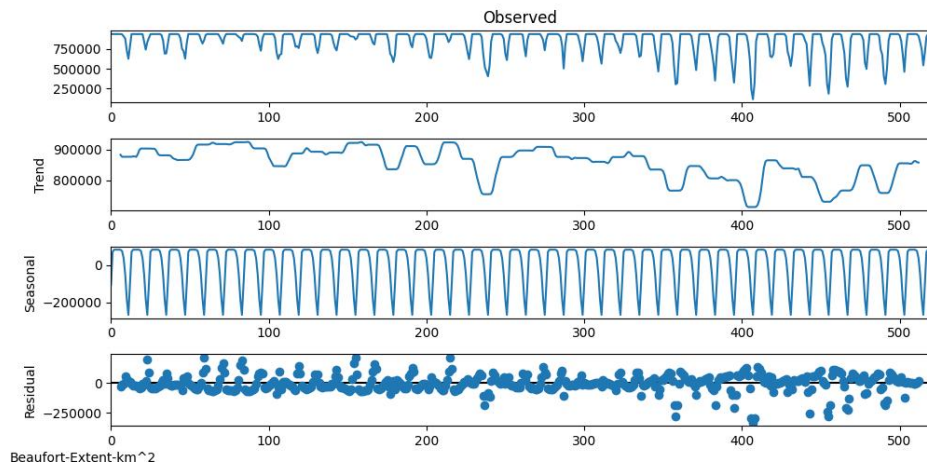


图 21 Beaufort 扩展地区

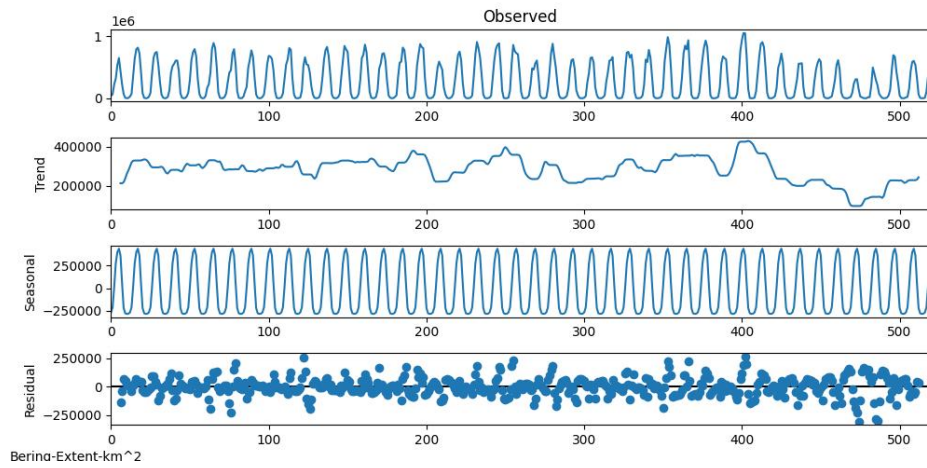


图 22 Bering 扩展地区

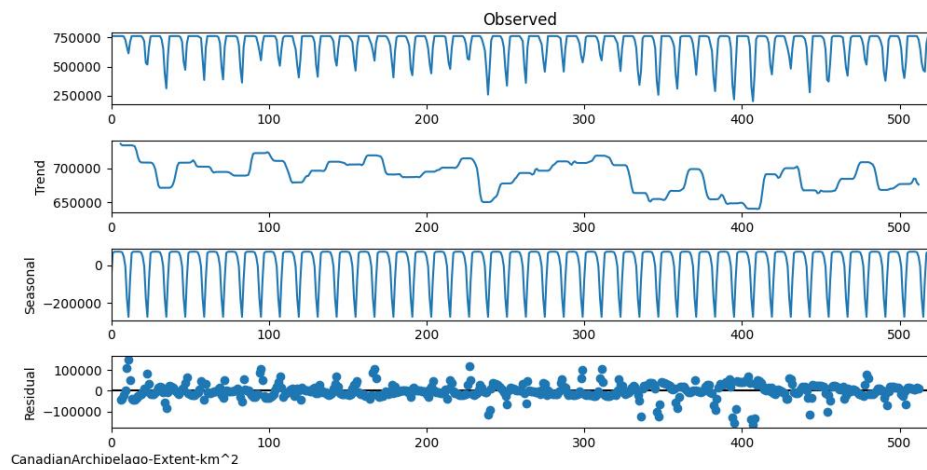


图 23 CanadianArchipelago 扩展地区

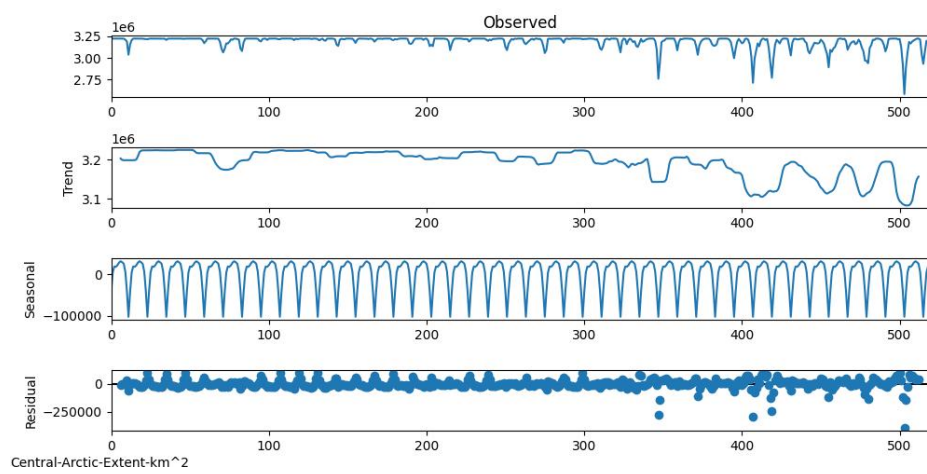


图 24 Central-Arctic 扩展地区



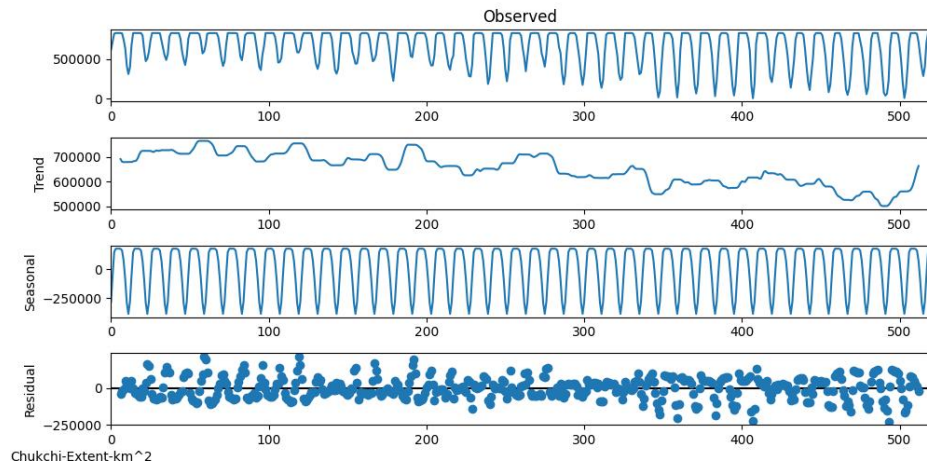


图 25 Chukchi 扩展地区

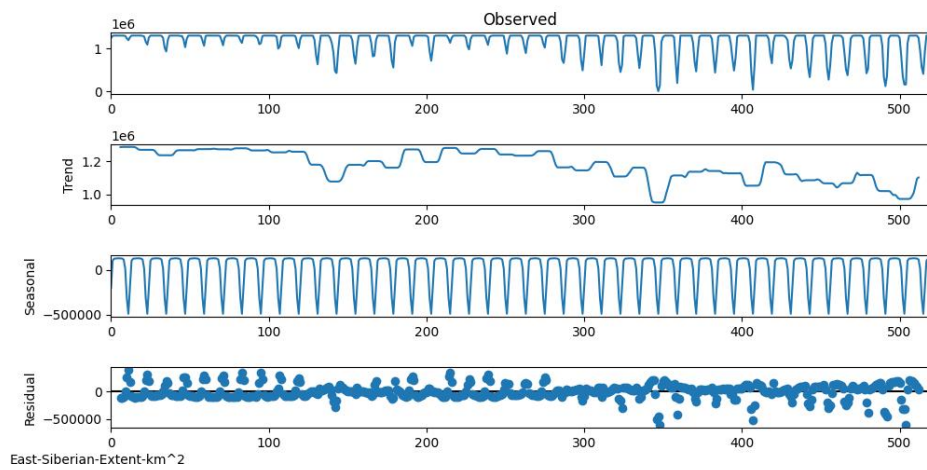


图 26 East-Siberian 扩展地区

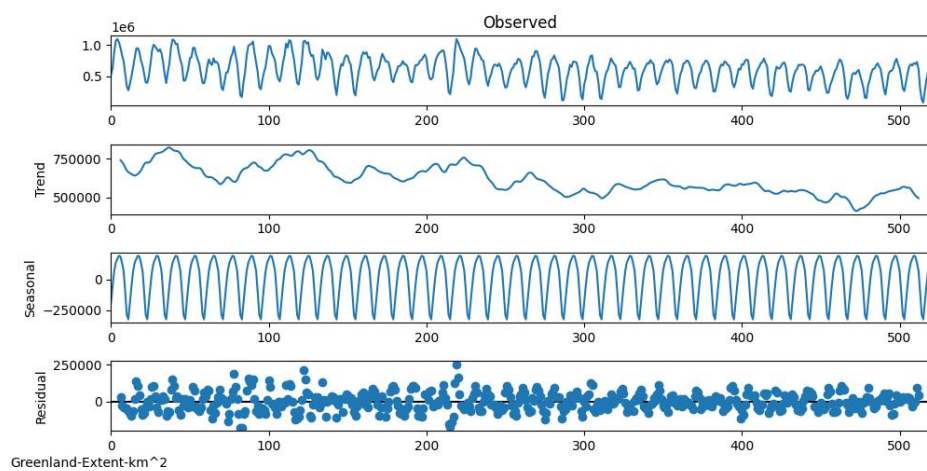


图 27 Greenland 扩展地区

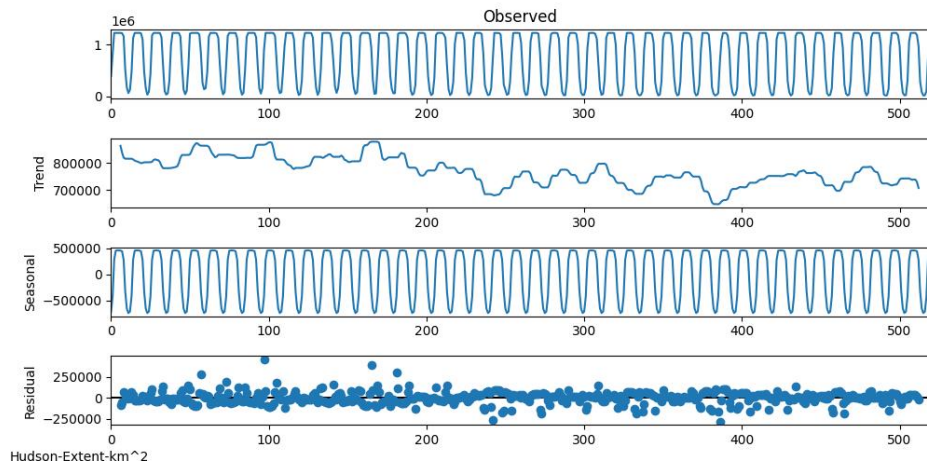


图 28 Hudson 扩展地区

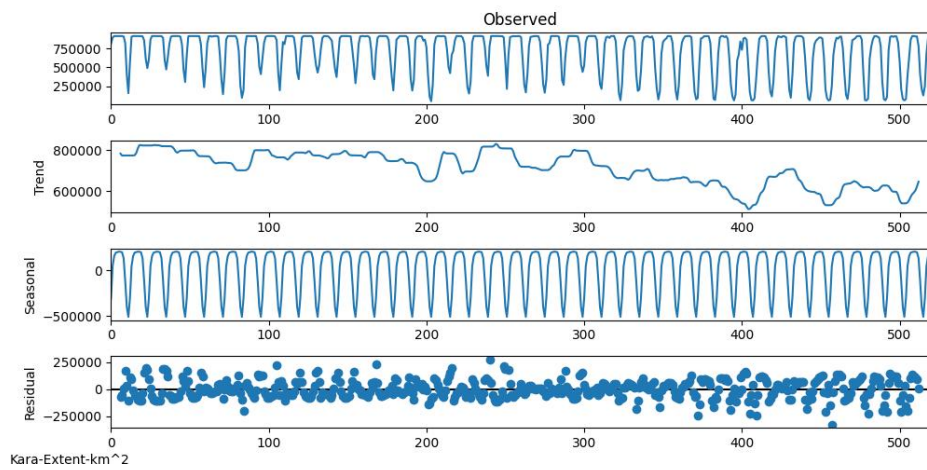


图 29 Kara 扩展地区

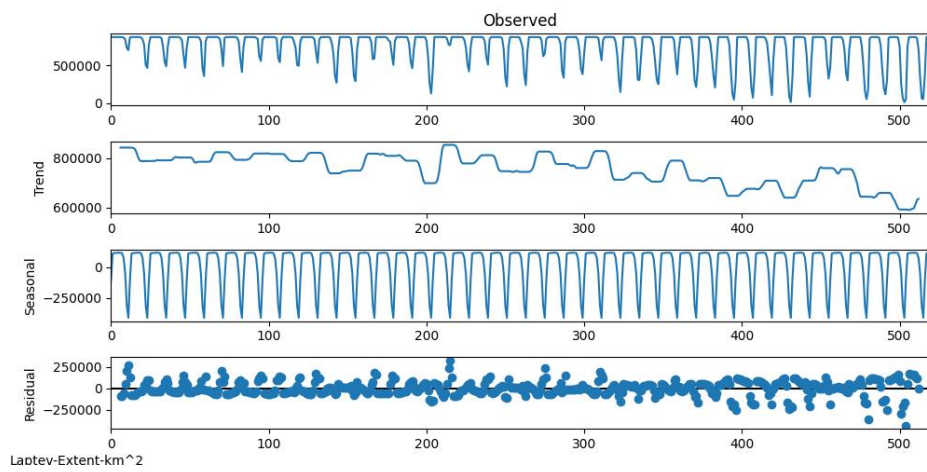


图 30 Laptev 扩展地区



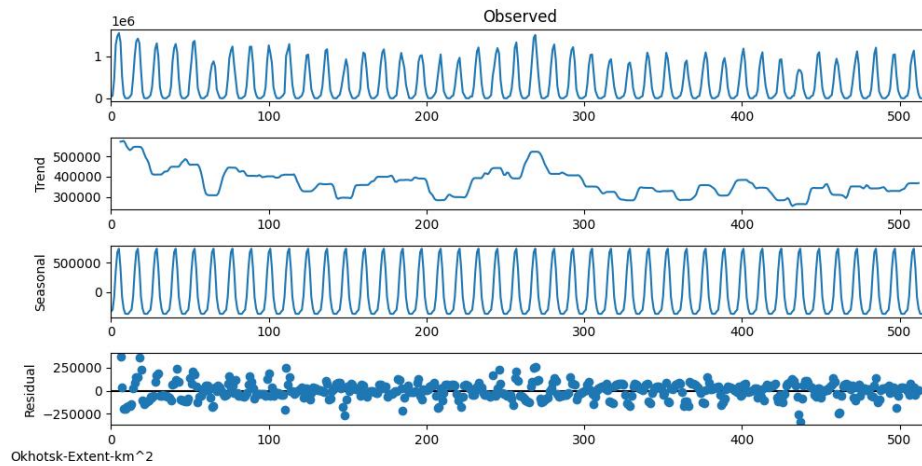


图 31 Okhotsk 扩展地区

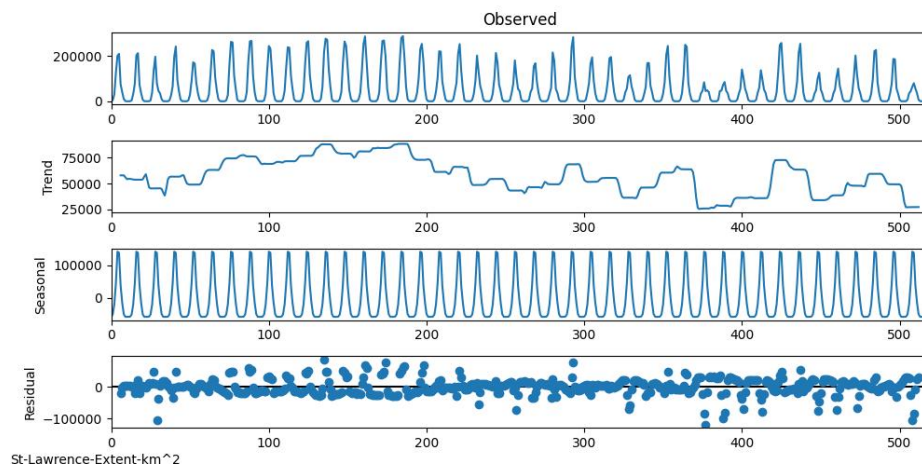


图 32 St-Lawrence 扩展地区

## 11.2 Python 实现问题三海冰面积变化趋势相关数据的提取

```
import matplotlib.pyplot as plt
import pandas as pd
import numpy as np
from statsmodels.tsa.seasonal import seasonal_decompose
import scipy.stats as stats

file = '/Users/smh/Desktop/数学建模集训/数模七月短学期集训模型3/N_seaice/北冰洋/每日/
        N_seaice_extent_monthly_mean_for_q3.csv'
fileco2 = '/Users/smh/Desktop/数学建模集训/数模七月短学期集训模型3/co2_mm_mlo_for_q3.csv'
sheet = pd.read_csv(file, usecols=[0, 1], header=None, encoding = "ISO-8859-1").dropna()
sheet = sheet.to_numpy()
sheet[:, 0] = sheet[:, 0].astype(int)

sheetco2 = pd.read_csv(fileco2, usecols=[0, 1, 4])
sheetco2 = sheetco2.to_numpy()
```

```

# # sea ice size figure seasonalized
# plt.figure(1)
y1 = sheet[:, 1]
x1 = np.arange(0, len(y1))
# plt.title("sea ice area (seasonalized)")
# plt.plot(x1, y)

# co2 figure
plt.figure(2)
x2 = x1
y2 = sheetco2[:, 2]
plt.title("co2 mm mlo (deseasonalized)")
plt.plot(x2, y2)

# classical seasonal decomposition
# decomposition figure
res = seasonal_decompose(y1, model='multiplicative', period=12)
fig = res.plot()
fig.show()

trend: np.ndarray = res.trend
# sea ice size figure
plt.figure(3)
plt.title("sea ice area (deseasonalized)")
# plt.plot(np.arange(0, len(trend), trend))
x3: list = []
y3: list = []
for i in range(len(trend)):
    if not np.isnan(trend[i]):
        x3.append(i)
        y3.append(trend[i])
plt.plot(x3, y3)

# considering correlation
# data alignment
x: list = []
sx = x2 if len(x2) < len(x3) else x3
for i in range(min(len(x2), len(x3))):
    ssx = sx[i]
    if ssx in x3:
        x.append(ssx)

y2_1: list = []
for sx, sy in zip(x2, y2):
    if sx in x:
        y2_1.append(sy)

y3_1: list = []
for sx, sy in zip(x3, y3):

```

```

    if sx in x:
        y3_1.append(sy)

r, p = stats.pearsonr(y2_1, y3_1)
print("correlation coeff:%f, p: %f" % (r, p))

# sea ice area & co2 figure
# plt.figure(4)
# plt.title("sea ice & co2")
# plt.plot(x2, y2)
# plt.plot(x3, y3)
# plt.legend(['co2', 'sea ice size'])
# plt.text(0, 200, "r=%f, p=%f" % (r, p))
# plt.show()
plt.figure(4)
fig, ax = plt.subplots()
curve1 = ax.plot(x2, y2, color='#1f77b4')
ax.set_ylabel('co2')
ax2 = ax.twinx()
curve2 = ax2.plot(x3, y3, color='#ff7f0e')
ax2.set_ylabel('sea ice')
ax.legend(curve1+curve2, ['co2', 'sea ice'], loc='upper center')
plt.show()

```

### 11.3 问题二至问题五的模型

本文问题二至问题五的模型均用 SPSSPRO 实现。