**C题：“双碳”目标下低碳建筑研究**

“双碳”即碳达峰与碳中和的简称，我国力争2030年前实现碳达峰，2060年前实现碳中和。“双碳”战略倡导绿色、环保、低碳的生活方式。我国加快降低碳排放步伐，大力推进绿色低碳科技创新，以提高产业和经济的全球竞争力。

**问题1：**现在有一间长4米、宽3米、高3米的单层平顶单体建筑，墙体为砖混结构，厚度30厘米(热导系数)，屋顶钢筋混凝土浇筑，厚度30厘米(热导系数)，门窗总面积5平方(热导系数)，地面为混凝土 (热导系数)。该建筑物所处地理位置一年(按365天计算)的月平均温度(单位：摄氏度）见下表。

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **月份** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| **平均温度** | -1 | 2 | 6 | 12 | 22 | 28 | 31 | 32 | 26 | 23 | 15 | 2 |

假设该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。请计算该建筑物通过空调(假设空调制热性能系数COP为3.5，制冷性能系数EER为2.7)调节温度的年碳排放量。这是典型的物理传热模型分析。

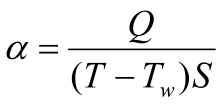
内侧：

外侧：

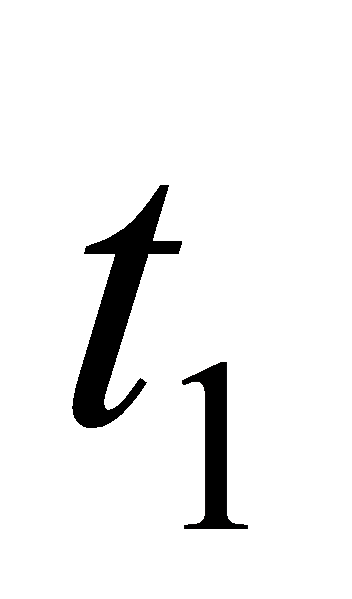
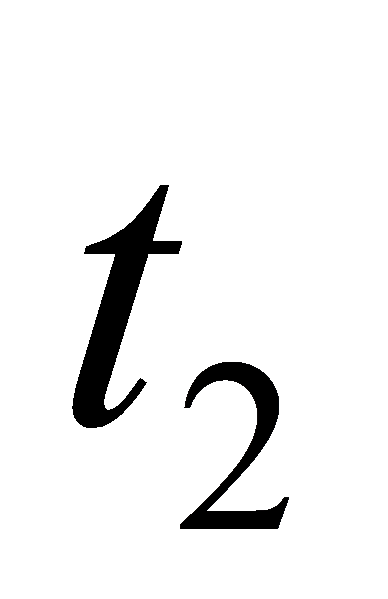
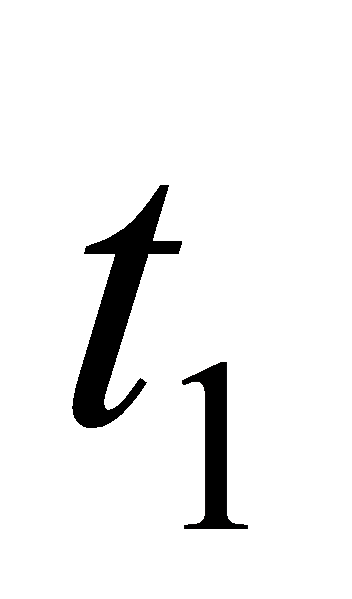
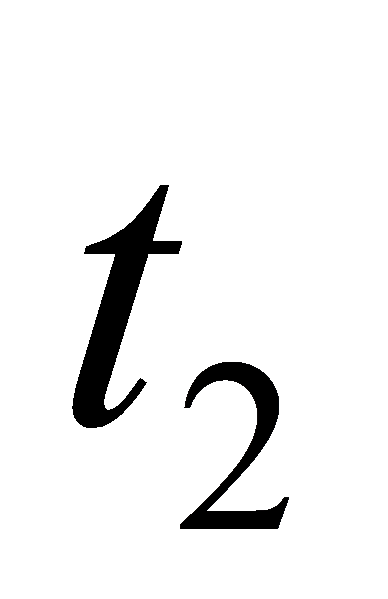
值得注意的是，由于流体温度沿传热面方向在不断变化，因此传热系数和壁温也在不断变化。

上式中的、、、均取平均值。

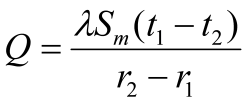
流传热系数

(单位：)

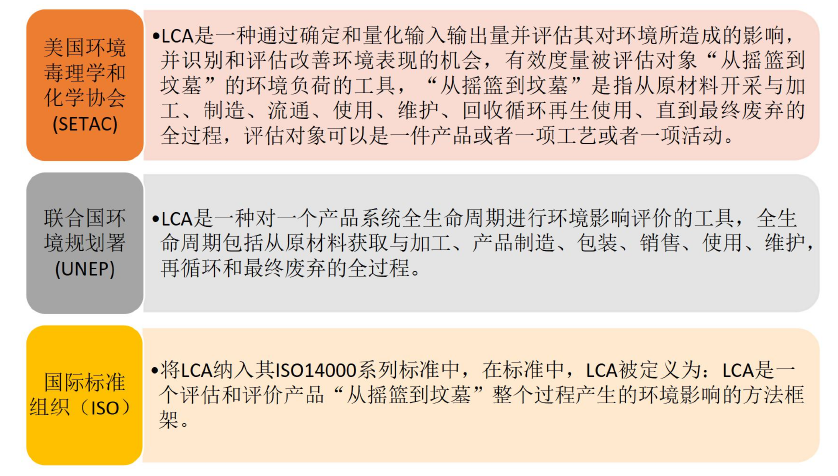
意义：表示温差为一度时，单位面积的传热量，是反应对热传热快慢的一个参数。

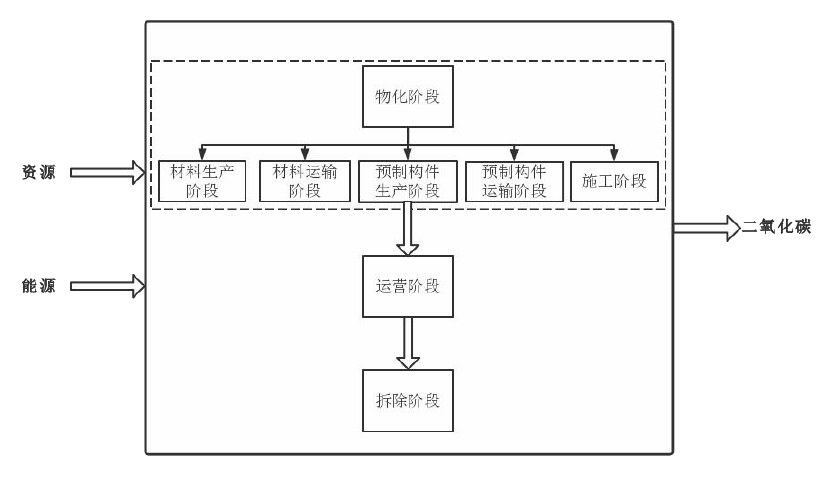
假设建筑物其导热系数为λ，长度为L，面积S，内、外表面的温度分别为和，并且>，沿着建筑物的xy方向进行热传导。由于导热面积沿着半径方向变化，因此需要平均导热面积来计算导热速率。

将温度梯度和平均导热面积代入傅里叶定律（    ）表达式，可得：



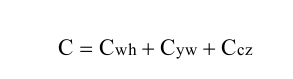
**问题2：**在居住建筑的整个生命周期 (建造、运行、拆除)中，影响碳排放的因素有很多，如建筑设计标准、气候、建材生产运输、地区差异、建造拆除能耗、装修风格、使用能耗、建筑类型等。请查找、分析资料，建立数学模型，找出与上述因素相关度大且易于量化的指标，基于这些指标对居住建筑整个生命周期的碳排放进行综合评价





采用化石能源消耗因子、电力消耗因子、运输方式碳排放因子、材料碳排放因子、材料生产阶段碳排放占全生命周期碳排放的比重较大，可见材料碳排放因子是影响建筑全生命周期碳排放量大小的关键因素。

常用的材料包括钢材、混凝土、铝材、水泥、砂子、砌块、涂料、建筑陶瓷、玻璃、防水卷材等。其中，钢材、铁和铝材可以回收循环再生利用，需要考虑其回收率，考虑回收率后的钢材和铁碳排放因子从文献中获取。



**问题3：**在问题2的基础上，分别考虑建筑生命周期三个阶段的碳排放问题，查找相关资料，建立数学模型，对2021年江苏省13个地级市的居住建筑碳排放进行综合评价，并对所建评价模型的有效性进行验证。查找江苏省各市的碳排放数据，可以直接采用模糊综合评价法解决。模糊聚类分析法是一种通过定量分析来确定样品亲疏关系的方法，模糊聚类分析是涉及事物之间的模糊界限时按一定要求对事物进行分类的数学方法。将江苏13市分为优劣，或者采用判别函数分析方法，根据多元回归、主成分分析得到主要的影响因素权重，判别函数法得到排名。

**问题4：**准确的碳排放预测能够为制定减排政策、优化低碳建筑设计提供重要的参考依据。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。根据历史数据，用时间序列ARIMA和支持向量机进行预测，并对多种模式进行组合，以提高预测精度，可以进行一个对比。灰色预测也可以使用。

**模型假设**

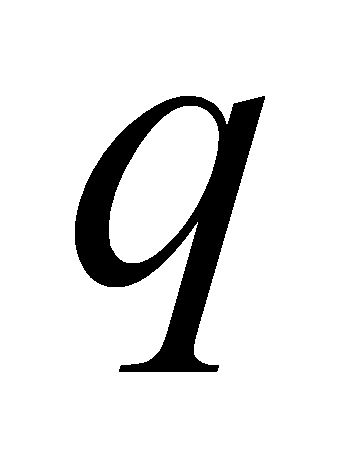
* 热损失主要是由热辐射和对流换热引起的。为了方便研究和计算，我们忽略了热辐射对热损失的影响，主要考虑对流传热。
* 碳排放只考虑摘要所述因素
* 历史数据准确

**三、符号含义说明**

|  |  |
| --- | --- |
| **符号** | **说明** |
| W  S | 距离  流量  横断面积  时间  导热速率  导热系数  长度  内、外表面的温度  影响权重  预测值 |

模型的建立与求解

问题一要假设该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。请计算该建筑物通过空调(假设空调制热性能系数COP为3.5，制冷性能系数EER为2.7)调节温度的年碳排放量。(尽量使用本题所给条件计算碳排放，不考虑其他损耗)

首先，我们认为空调的气体的流速对房间的散热有着密不可分的关系，若知道建筑物的横断面积以及流量，则由流速的定义我们不难得到：

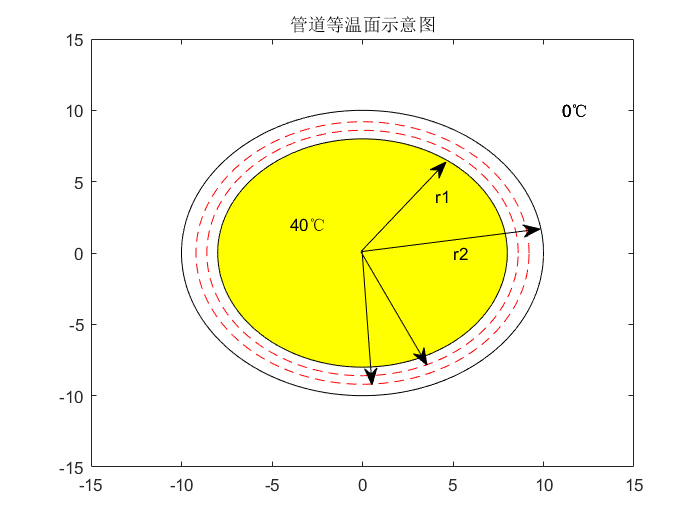
其次，我们开始着手于研究建筑物液体所蕴含的热能，可以按照吸热放热公式：



那么热能便可由以下公式求出



我们可以算出单位时间内调节室内外温差所需的电能（在常温常压下，将各项数据代入，可以算出，在单位时间内，。绘制等温截面图如下所示。、

、、

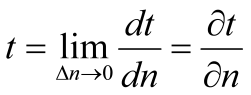
依据题目，我们发现该问题涉及高温物体传递给低温物体，所以它的特点为无宏观的位移。另外，对于固体的导热，我们假设的对象是一个“各向同性，质地均匀”的固体，即这一块物体它的转热性质是一样的。为方便解决问题，我们首先明确以下概念或做出以下假设：

1.温度场（ ）：指物体或空间各个点的温度在时空的分布



其中，，均为空间坐标。

2.指两个等温面的温差与其法线距离的比值，用公式表示即为：

值得注意的是，随着温度趋近于0，这意味着温度梯度是一个点的概念，而不是距离，并且该点具有方向，因此它是一个矢量，该方向垂直于等温面，其中 它指出，并且温度上升。 方向。

对于建筑物，在内部和外部都有两个传热表面。 然后可以写出相应的两个传热速率方程：

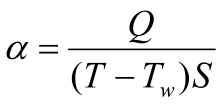
内侧：

外侧：

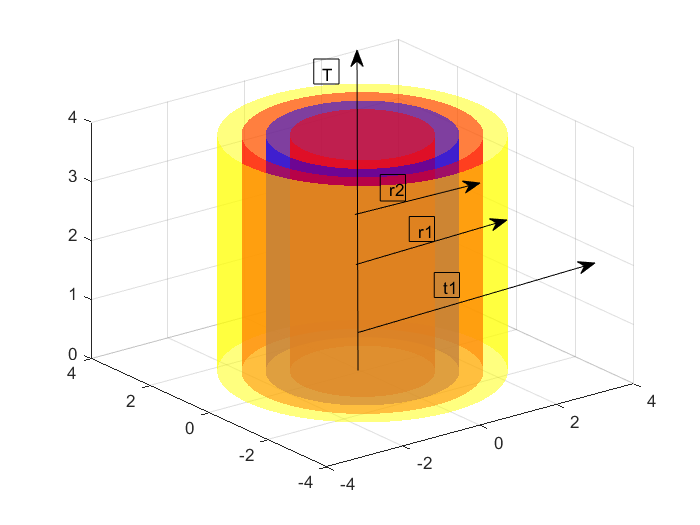
值得注意的是，由于温度沿传热面方向在不断变化，因此传热系数和壁温也在不断变化。

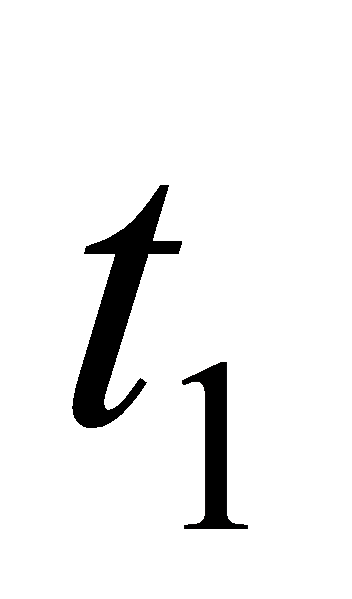
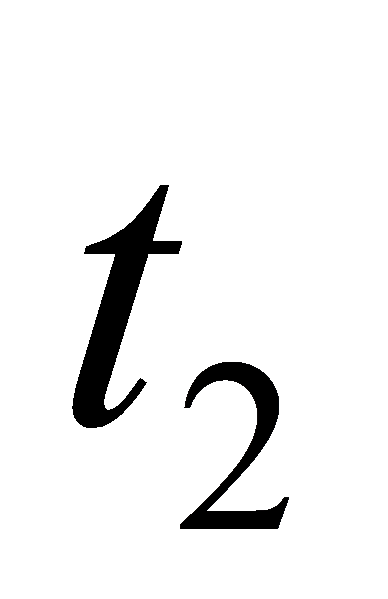
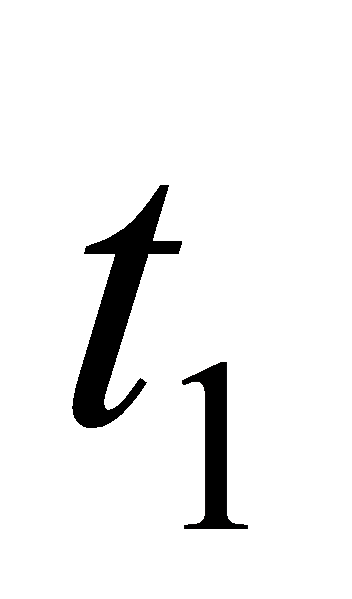
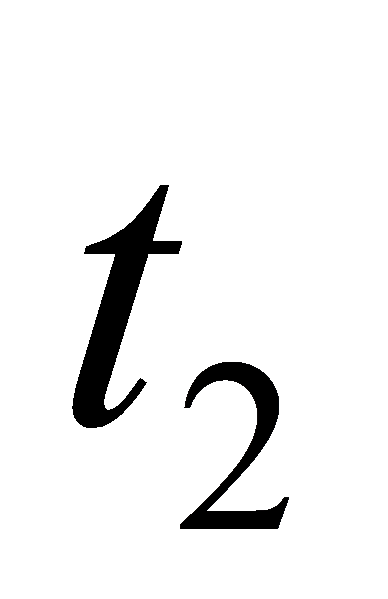
上式中的、、、均取平均值。

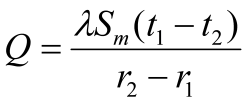
流传热系数

(单位：)

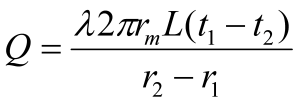
意义：表示温差为一度时，单位面积的传热量，是反应对热传热快慢的一个参数。



如图所示的建筑物，若其导热系数为λ，长度为L，内半径为，外半径为，内、外表面的温度分别为和，并且>，沿着建筑物的半径方向进行热传导。将温度梯度和平均导热面积代入傅里叶定律（    ）表达式，可得：



上式称为建筑物的热传导方程式，将平均导热面积代入式，则有：

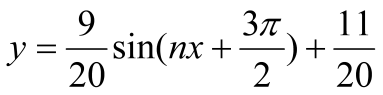


一度电相当于*0.28*千克碳排放

而室外的温度是随月份周期变化的，我们采用MATLAB进行模拟仿真温度热流的变化

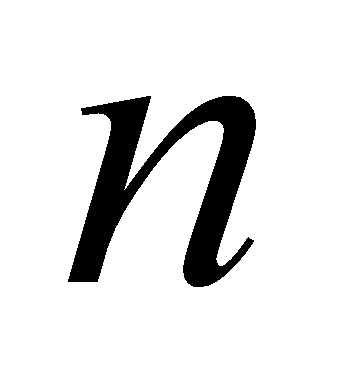
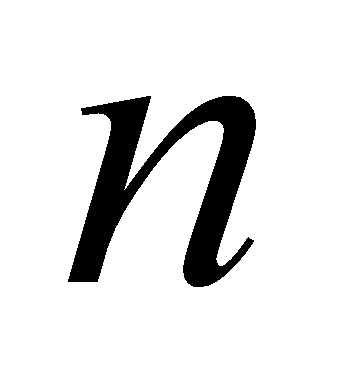
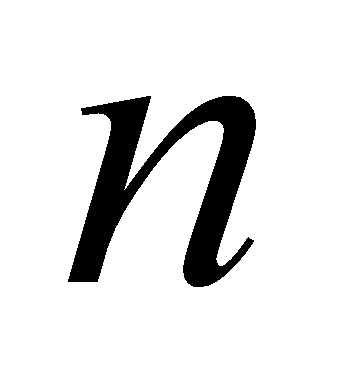
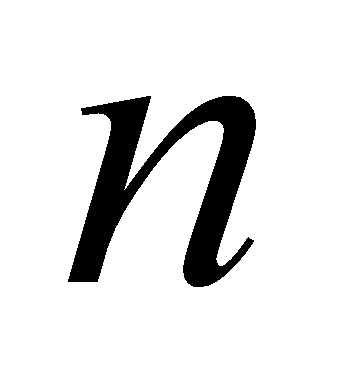
热流量从定量变化为变量，并且在（单位）之间变化，因为它是变量，所以在一定程度上考虑这种情况不便于研究问题 ，我们打算对流量进行一定的限制。 想象一下我们是否可以将两者与功能关系结合起来。 经过筛选后，我们决定将两者之间的关系设置为正弦函数，以便可以解决问题。

出于对问题的简便考虑，我们决定将函数的形式选为最简单的：



通过MATLAB绘制不同n值下的函数曲线，判断散热速度大小。



从图像中我们不难发现，的取值决定了正弦函数的收缩，换句话说，的值越大，图像越陡；的值越小，图像越平缓。由此，我们便可将的取值看作是反应流量变化快慢的一个参数。

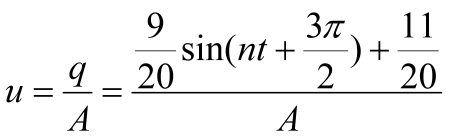
我们可以做以下规定：

当时，热流量变化速率较慢；

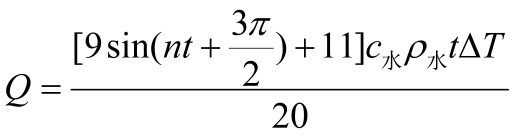
当时，热流量变化速率适中；

当时，热流量变化速率较快。

那么流速随时间的关系式我们就可以写作：

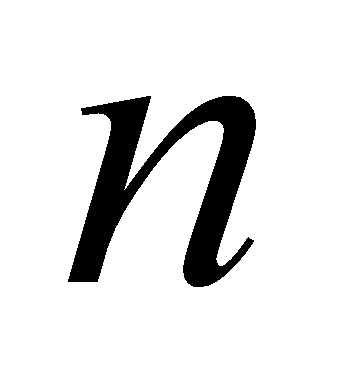


将式代入式我们就可以得到一个热量随时间变化的关系式：



该热量可以近似看作是建筑物所吸收的热量 绘制成曲线如下图所示：



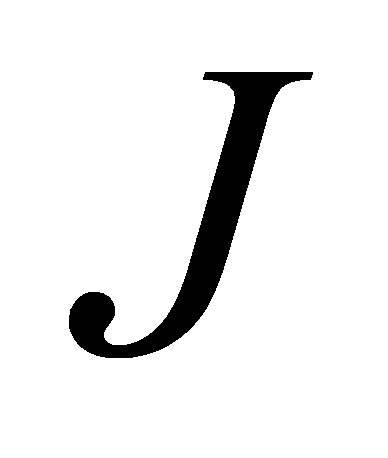
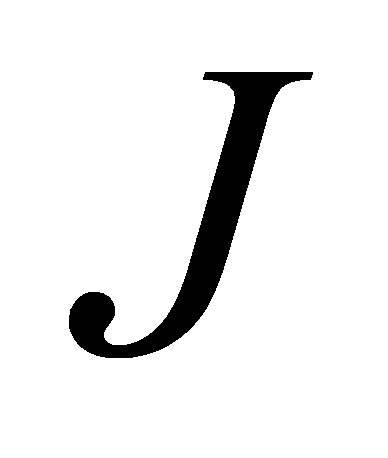
由图可看出，当 取不同值时，曲线的峰值与峰数都不同，这也就意味着，当流量变化速率较快时，产生的热量也就越多，那么相应地在同一坐标上，该点的温度应比之前的要高。

我们首先绘制标题中给出的室外温度表的折线图，目的是将室外温度与解除建筑物的外侧温度热流量联系起来，此外，我们还将拟合温度数据，该函数使用多项式：

进行拟合，并得到相应拟合残差图，结果如下所示。





题目要求我们将室内的温度维持在18-26度范围，这一区间，然后我们再根据前面的理论基础，算出要想达到这一要求所需的热流量。例如：当室内初始温度为时，室内空气所要吸收的热量范围，我们可以算得是（单位：），那么空调所需要产生的热量即为（单位；），接着我们便可以大概地算出此时空调热风所要达到的一个流量，我们算出来是在（单位：）这一区间。

但本题的问题在于：如何联系室外温度的影响，使得室内温度有一个很好的估计。

基于该问题，我们首先研究了室外空气的变化速率，绘制成以下折线图：



在图中我们能够直观地感受室外温度的变化速率。

关于该问题，我们重点研究的是温差，我们的想法是在室外不同时刻的温差等于的阶段，保持通入以该温度范围能够接受的最小流量，若超过了该温度范围就必须得对流量进行调整，这样做，既稳定了室内温度范围，又能够使耗电量最小。

现在，我们先假设一点：在某时间内，由于室内呈非闭合状态，空气流动是相通的，故可以将室内某温度近似地看作是对应该时刻的室外温度。

那么我们便可以这样认为：

假设在温度区间内

若，我们只需要得到升到18度时需要的流量即可；

若，我们只需要得到升到18度时需要的流量即可。

下表便是经过计算所得出的相应温度所需要的最小流量：

|  |  |  |
| --- | --- | --- |
|  |  | 产生到达18度时热所需要的流量 |
| -5 | -2 | 0.7 |
| -2 | 1 | 0.62 |
| 1 | 4 | 0.54 |
| 4 | 7 | 0.46 |
| 7 | 10 | 0.37 |
| 10 | 7 | 0.37 |
| 7 | 4 | 0.46 |
| 4 | 1 | 0.54 |
| 1 | -2 | 0.62 |
| -2 | -5 | 0.7 |

有些温度经过对照图表，可以得到相应的时刻，而有些温度无明显的时刻，这时我们就需借助拟合函数进行求解，将（或）代入式中解得两个值，这两个值便为对应时刻。并计算全年的碳排放为79.38kg

问题二

随着全球气候的逐渐变暖，发展低碳建筑已成为必然趋势，对建筑碳排放的研究也势在必行。目前，在大多数关于建筑施工过程碳排放能源管理的研究中，人们无法用更准确的方法对其进行分析。为了寻找一种有效的碳排放分析方法，本文基于生命周期评价方法(LCA)，建立了碳排放模型，对不同区域的住宅建材、不同工业建筑的建材、空调冷暖方案、建筑全生命周期的碳排放量进行了计算。本文得出以下结论:不同地区住宅建筑材料碳排放的关系为:严寒地区>夏热冬冷地区>温和地区>夏热冬暖地区。在框架结构、砖混结构和钢混凝土结构的分析中，钢混凝土结构的碳排放量最多，框架结构的碳排放量最少。不同形式的空调冷源的碳排放量是不同的。其中，第二种选择是直燃式溴化锂吸收式冷水机组和热水机组。碳排放与能源结构有关。对于装配式建筑，建筑使用阶段的碳排放比例最高，达到75%-80%，建筑使用阶段和建材生产阶段的碳排放之和占总量的90%以上。

5.2 问题二模型的建立与求解

1. 问题分析

建筑物的能源消耗受到许多因素的影响，例如气候区差异，设计类型，施工方法，建筑材料，属性和能源管理.碳排放量的计算非常复杂.目前，基于生命周期评估(LCA) 的建筑物碳排放研究取得了一些成果.LCA方法主要用于计算产品在整个生命周期中的碳排放量，包括原材料，生产，使用和拆除。我们首先确定研究系统的边界,然后定量计算资源消耗和碳排放量，最后分析对环境的影响。施工过程分为三个阶段，确定每个阶段的碳排放边界。根据各阶段的特点，对碳源进行分类，根据油耗和燃料排放系数确定各机械团队的排放系数。可以计算机械设备在施工过程中的总碳排放量。

1. 评价指标

(1)建筑使用阶段低碳的评价用两个指标表示，见公式:



在公式中，Cm2——建筑使用过程中单位面积和单位时间的碳排放量(kgC02/m2a) ;P4——建筑使用过程中的碳排放量(公斤) ;S--建筑面积(m2) ;n--建筑使用时间，需要50年(一) ;CP单位时间，建筑使用期间人均碳排放量(千克C02/人a) ;建筑物用户的P数(人)。

(2)气候校正评价指标:

在比较不同地区的低碳建筑时，应添加气候因素。北方建筑物的供暖和南方建筑物的制冷消耗了大量的能源，带来了大量的碳排放，在进行比较时应该考虑到这一点。本文增加了供暖和制冷天数，以比较不同地方建筑物的人均碳排放量。



在公式中，C-人均碳排放指数(千克CO2/人) ;CA-单位面积能耗碳排放指数(千克CO2/m2) ;PA- 人均建筑面积(m2/人)

(3)建筑物各阶段(使用阶段除外)的碳排放量化指标如下:



在公式中，Ck.m2-建筑物每个阶段单位面积的碳排放量(kgCO2|m2) ;建筑物每个阶段的碳排放(公斤)。三个阶段的评价指标可以单独评价，也可以与建筑使用阶段的评价指标结合起来评价。

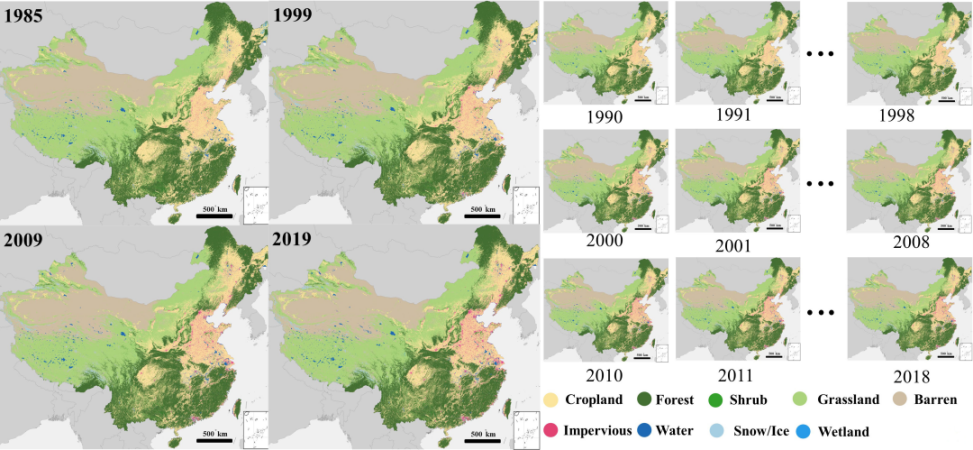
1. 指标分析

(1)不同地区主要建筑材料的碳排放分析

不同的地区有不同的建筑风格和形式，导致建筑材料的数量不同，因此建筑材料的碳排放量也会有所不同。过去，人们只关注建筑材料的数量，而忽略了建筑材料的碳排放。本文计算和分析了建筑材料单位面积的碳排放量。本文以不同气候区域的几个城市为例进行分析。寒区以哈尔滨为例，寒区以石家庄为例，以炎热的夏天和寒冷的冬天以南京为例，以温暖的夏天和温暖的冬天以广州为例。以昆明为例，各城市住宅建筑中主要建筑材料用量和比例

统计如表1所示:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 钢筋 | 木头 | 混凝土 | 砖瓦 | 沙土 | 石灰 | 玻璃 | 总计 |
| 哈尔滨 | 0.02 | 0.014 | 0.12 | 0.99 | 0.86 | 0.04 | 0.005 | 2.05 |
| 石家庄 | 0.05 | 0.017 | 0.1 | 0.72 | 0.675 | 0.051 | 0.004 | 1.57 |
| 南京 | 0.03 | 0.012 | 0.15 | 0.62 | 0.69 | 0.01 | 0.002 | 1.49 |
| 昆明 | 0.03 | 0.009 | 0.18 | 0.556 | 0.90 | 0.015 | 0.003 | 1.67 |
| 广州 | 0.04 | 0.006 | 0.14 | 0.352 | 1.30 | 0.2 | 0.003 | 2.00 |



从表1可以看出，每个城市的建筑材料量是不同的，这与每个地区的建筑形式和周围结构有关。从表1可以明显看出:

(1)在上述五个城市中，沙子、砖块和水泥的使用量最大，其使用量超过建

筑材料总重量的90%

(2)在上述五个城市中，砖块量从北到南依次减少。越冷的地方使用越多，

水泥被逆转;

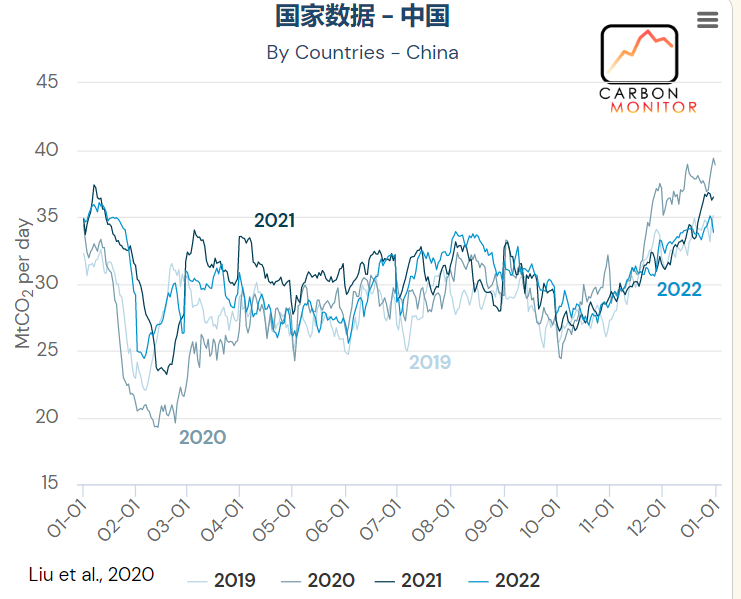
(3)使用较少的钢，石灰，木材和玻璃。

根据本文构建的碳排放计算模型，计算了建筑材料的碳排放量。可以看出，建材的碳排放量与建材的消耗量存在较大差异，这与单位建材的排放系数有关:

(1)以上五个城市水泥、砖、钢、石灰碳排放量较大，占统计总量的95%以上;玻璃、木材和沙子的碳排放量低于5%;

(2)从总排放量来看，重寒区>寒区夏季炎热、冬季寒冷地区>温和地区>>夏季炎热、冬季温暖地区，排放比为1: 0.9778: 0.8442: 0.7903:0.7854。

在统计的几种建筑材料中，沙子用量最大，超过40%，但沙子的碳排放量不到5%;钢和石灰的用量不大，两者之和小于5%，但碳排放量大，高于18%。通过以上对比可以发现，建材的碳排放量与建材用量和建材碳排放系数两方面都有关系。大量的建筑材料不一定有高碳排放量。从以上数据可以发现，如果要减少建筑材料的碳排放，就应该从几种碳排放量大的建材(砖、水泥、钢材)入手，通过改变建筑结构来减少高碳建材的用量，通过改进建材生产工艺来减少单位建筑材料的碳排放。调查发现，不同生产工艺中建筑材料的碳排放系数差异较大。以水泥为例，当今我国主流水泥生产工艺的能效为30%，新型干法水泥的能效可达55%。



5.3 问题三模型建立与求解

1. **基本思想与原理**

聚类分析法是一种通过定量分析来确定样品亲疏关系的方法，模糊聚类分析是涉及事物之间的模糊界限时按一定要求对事物进行分类的数学方法。[聚类分析](https://baike.baidu.com/item/%E8%81%9A%E7%B1%BB%E5%88%86%E6%9E%90" \t "https://baike.baidu.com/item/%E6%A8%A1%E7%B3%8A%E8%81%9A%E7%B1%BB%E5%88%86%E6%9E%90/_blank)是[数理统计](https://baike.baidu.com/item/%E6%95%B0%E7%90%86%E7%BB%9F%E8%AE%A1/408183" \t "https://baike.baidu.com/item/%E6%A8%A1%E7%B3%8A%E8%81%9A%E7%B1%BB%E5%88%86%E6%9E%90/_blank)中的一种多元分析方法，它是用数学方法定量地确定样本的亲疏关系，从而客观地划分类型。事物之间的界限，有些是确切的，有些则是模糊的。当聚类涉及事物之间的模糊界限时，需运用模糊聚类分析方法。模糊聚类分析广泛应用在[气象预报](https://baike.baidu.com/item/%E6%B0%94%E8%B1%A1%E9%A2%84%E6%8A%A5" \t "https://baike.baidu.com/item/%E6%A8%A1%E7%B3%8A%E8%81%9A%E7%B1%BB%E5%88%86%E6%9E%90/_blank)、[地质](https://baike.baidu.com/item/%E5%9C%B0%E8%B4%A8" \t "https://baike.baidu.com/item/%E6%A8%A1%E7%B3%8A%E8%81%9A%E7%B1%BB%E5%88%86%E6%9E%90/_blank)、[农业](https://baike.baidu.com/item/%E5%86%9C%E4%B8%9A" \t "https://baike.baidu.com/item/%E6%A8%A1%E7%B3%8A%E8%81%9A%E7%B1%BB%E5%88%86%E6%9E%90/_blank)、[林业](https://baike.baidu.com/item/%E6%9E%97%E4%B8%9A/84154" \t "https://baike.baidu.com/item/%E6%A8%A1%E7%B3%8A%E8%81%9A%E7%B1%BB%E5%88%86%E6%9E%90/_blank)等方面。

本文选取2021年江苏省13各地级市为样本，通过查找2021年各城市碳排放量数据，以全年碳排放量平均值作为这13个城市的评价指标，通过构建模糊聚类模型，将这13所城市进行优劣分类。

1. **模糊聚类分析的计算步骤[4]**

首先建立评价信息矩阵



**Step1：**对原始数据进行标准化处理。

采用极差变换：



其中为已知评价指标，为标准化数值。

**Step2：**建立相似矩阵。

采用最大最小法建立相似矩阵：



其中，



**Step3：**利用平方法求解相似矩阵的传递闭包，改造相似关系为等价关系。

**Step4：**利用传递闭包进行结果聚类。

1. **模型的建立**
2. **数据标准化**

由于分析范围为13所城市2021年碳排放量检测数据，本题不再对各影响因素，诸如建造阶段建筑物材料碳消耗、运行阶段燃料消耗、地区差异、气候因素导致的碳排放含量进行分析，而直接选用全年碳排放数值来衡量。

对数据进行标准化处理，采用极差变换：

 （4）

（4）式中：和分别为13所城市中碳排放的最小值和最大值；为第所城市第个月碳排放的标准化数值。

**2. 建立相似矩阵**

本题采用最大最小法建立相似矩阵。根据标准化数据建立各城市之间不同月份碳排放量指标的相似关系矩阵，其计算公式为：

 （5）

可得到13所城市2021年碳排放量的相似关系矩阵，如下表所示：

**表8.相似关系矩阵**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **南京** | **苏州** | **徐州** | **扬州** | **连云港** |
| **南京** | 1 | 0 | 0.1265 | 0.28 | 0.2544 |
| **苏州** | 0 | 1 | 0 | 0 | 0 |
| **徐州** | 0.1265 | 0 | 1 | 0.4517 | 0.4163 |
| **扬州** | 0.28 | 0 | 0.4517 | 1 | 0.7863 |
| **连云港** | 0.2544 | 0 | 0.4163 | 0.7863 | 1 |

1. **改造相似关系为等价关系**

由于相似矩阵满足自反性和对称性而不具备传递性，故为求其等价矩阵，需要对进行改造，只需求其传递闭包。

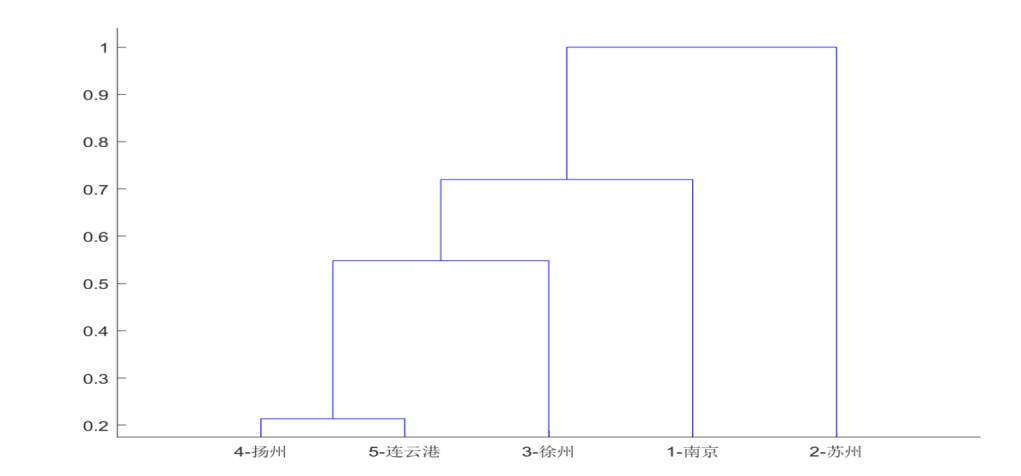
由平方法可求得传递闭包其具体数据如下所示：

**表7.传递闭包矩阵**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 1 | 0 | 0.28 | 0.28 | 0.28 |
| 0 | 1 | 0 | 0 | 0 |
| 0.28 | 0 | 1 | 0.4517 | 0.4517 |
| 0.28 | 0 | 0.4517 | 1 | 0.7863 |
| 0.28 | 0 | 0.4517 | 0.7863 | 1 |

1. **聚类结果分析**

利用传递闭包就是模糊等价关系矩阵。利用可对13所城市进行聚类分析，按不同的阈值水平对13所城市进行模糊聚类，将会得到不同的分类结果，聚类图如下所示：



**图2.聚类结果图**

**结果分析**：

情形1) 当时，将5所城市分为5类：，，，，。

情形2) 当时，将5所城市分为4类：，，，。

情形3) 当时，将5所城市分为3类：，，。

情形4）当时，将5所城市分为2类：，。

**表8.分类**结果**示意图**

|  |  |  |  |
| --- | --- | --- | --- |
| 空气质量 | 类别 | | |
| **优→→→劣** | **类别一** | **类别二** | **类别三** |
| 苏州 | 苏州 | 苏州 |
| 南京 | 南京 | 南京、连云港、扬州、徐州 |
| 连云港 | 连云港、扬州 |
| 扬州 |
| 徐州 | 徐州 |

5.4 问题四模型建立与求解

1.模型介绍

碳排放的日常监测数据存在波动性和周期性，但在预测研究中，很少有作者注意到数据波动对预测的影响。当数据量巨大且高度不稳定时，需要在预测初期对数据进行处理，并将不稳定的数据分解成稳定的序列进行分析。经验模态分解(EMD)方法出现在研究的早期。它是一种基于数据本身的时间尺度特征的信号分解方法，无需事先设置任何基函数。Li和Lu2015使用EMD模型将碳价格序列从高频到低频分解为独立的内在模态函数(IMFs)和残差项，然后使用不同的模型进行预测。但是，这种方法也有缺点。如果在分解过程中发生模态混叠，则特征模态函数没有物理意义。因此，提出集.成经验模态分解(EEMD)来克服EMD的缺点，得到了广泛的应用。EEMD可以有效提高预测性能，为了进一步提高预测性能，本文将EEMD应用于短期碳排放预测领域。

1. 建立EEMD模型

集成经验模态分解(EEMD)是一种针对EMD不足的噪声辅助数据分析方法。将白噪声引入信号。白噪声使其自动分布到适当的参考刻度。由于零均值噪声的特性，经过几次平均计算后噪声会相互抵消，因此积分均值的计算结果可以直接视为最终结果。积分均值与原始信号之差随积分平均次数的增加而减小。EEMD算法步骤如下:

(1)

将正态分布的白噪声添加到原始信号中;

x(t) = x0(t)十ni(t)

其中x (t)是添加白噪声后的新信号数据，xo(t)是原始信号数据，ni(t) 是白噪声。

(2) 将带有白噪声的信号作为一个整体，然后进行EMD分解。

m1(t)= 0.5(u1(t)一u2(t))

h1(t) =x(t)-m1(t)



重复(2)和(3)的工作，直到SD小于阈值(SD一般取0.2~0.3的值)即可得到第一个IMF。

(3) 重复步骤1和步骤2，每次添加新的正态分布白噪声序列;

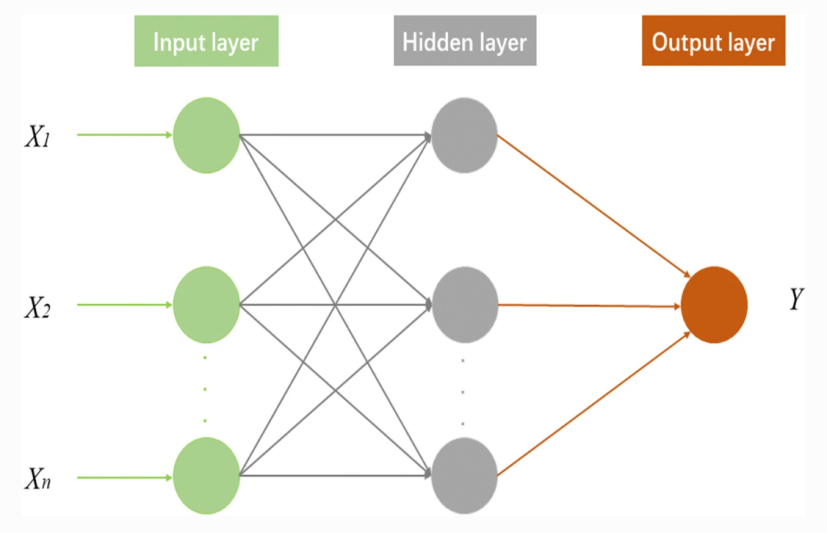
(4) 每次对IMF进行积分和平均后，得到最终结果



其中Cj(t)是最终得到的第j个分量，N是白噪声序列的数量，Ci(t)是添加第j个白噪声后的第i个分量。

3.BP神经网络

BP神经网络是一种人工神经网络，可以模拟人脑处理问题，模仿人脑神经系统的结构特征，分布式并行处理复杂问题。BP神经网络有输入层节点、输出层节点和隐藏层。BP神经网络的实现分为两个步骤。第一部分是前向输入传播。输入信号由输入层节点处理，然后传输到隐藏层节点。在隐藏层过程中，每个单元的神经元充当激活函数来处理信息。经过处理后，信息向前传输到输出层节点，得到输出结果。第二部分是反向传播。当输出层的输出未能达到预期的输出值时，错误信号沿原始连接路径返回。同时，反复校正每层神经元的权重，以最小化误差信号。当错误满足设定的要求时，网络的学习过程结束。下图显示了BP神经网络的结构。



4.基于粒子群优化的BP神经网络(PSOBP)

BP神经网络虽然具有较强的理论可行性，但在实际应用中仍存在局限性，例如，它容易陷入局部最小值、学习过程收敛速度慢和网络结构模糊等问题，因此采用粒子群优化算法(PSO)对其进行优化。PSO是一种进化的计算技术。该算法基于对鸟类捕食行为的研究。设计了一个无质量粒子来模拟鸟群中的鸟类。粒子只有两个属性:速度和位置，分别表示运动的速度和方向。每个粒子分别在搜索空间中搜索最优解，作为整个粒子群的当前全局最优解，并据此调整其速度和位置。由于实现简单，PSO已广泛应用于函数优化、神经网络训练、模糊系统控制等遗传算法。PSO初始化为一组随机粒子(随机解)。在每次迭代中，粒子跟踪两个“极端”(pbest， gbest) ，并使用以下公式进行自我更新。





其中vi表示第i个粒子的速度，cq和c2是学习因素，通常为C1= C2.在本文中，c1,c2都是1.49445，rand()是介于(0，1)之间的随机数。

5.实现PSOBP神经网络的步骤如下:

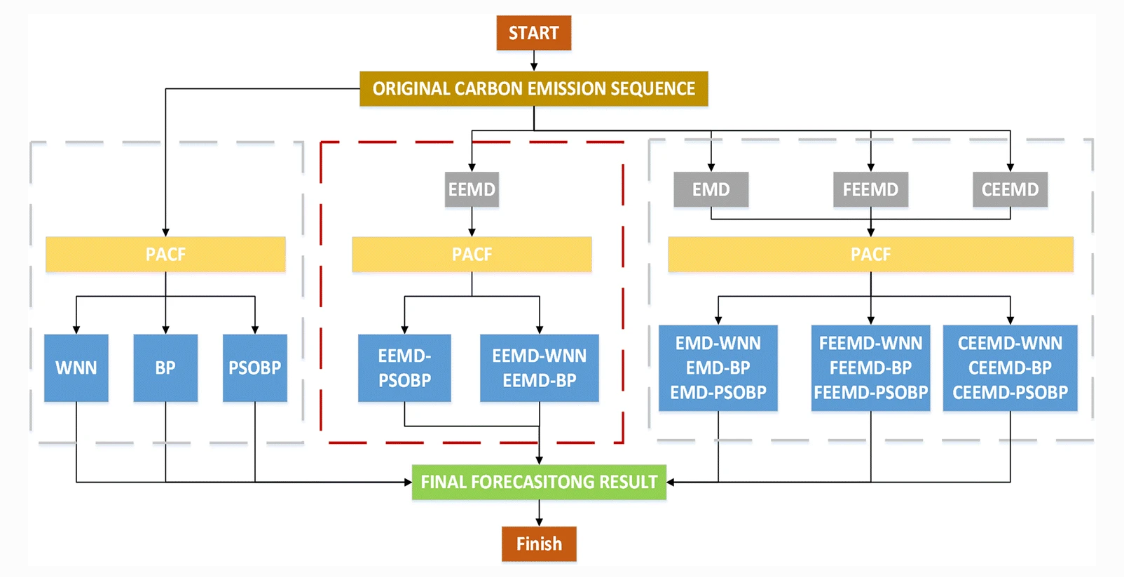
(1) 确定BP神经网络的结构和相关参数。

(2) 设置粒子的大小，粒子的初始飞行速度和相应的位置。选择粒子当前最佳光斑作为初始光斑，聚集的最佳光斑为最大最佳光斑。

(3)训练BP神经网络后， 使用得到的训练误差作为粒子当前位置的适应度值，将结果与之前最佳位置的值进行比较，如果优于之前的最佳位置，则替换它。否则不变。

(4) 根据公式(2.3.1) 和(2.3.2) ，重新规划粒子和相应点的飞行速度。

(5) 检查算法是否满足终止条件。如果满足条件，则输出最佳权衡值，然后模拟模型。否则，请跳过步进(3)



流程图

步骤1:数据预处理。利用EEMD平滑原始波信号，将日碳排放监测数据分解为一系列高度规则的内部模量函数。

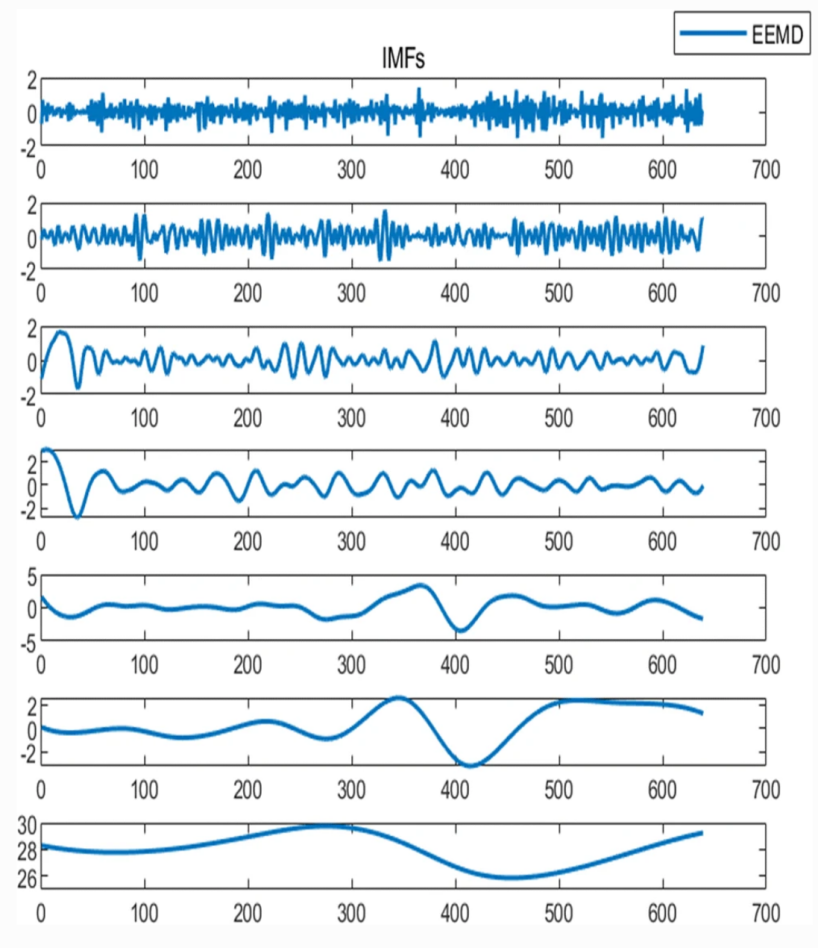
第2步:选择预测输入。偏自相关函数(PACF)用于确定每个内部模态函数与历史数据之间的关系。选择与要分析的序列高度相关的数据作为PSOBP的输入。

步骤3:内部模型函数。PSOBP用于预测内部模型函数。

步骤4:将预测结果集成。碳排放的最终预测值是通过将所有预测结果相加得到的。

5.1数据预处理

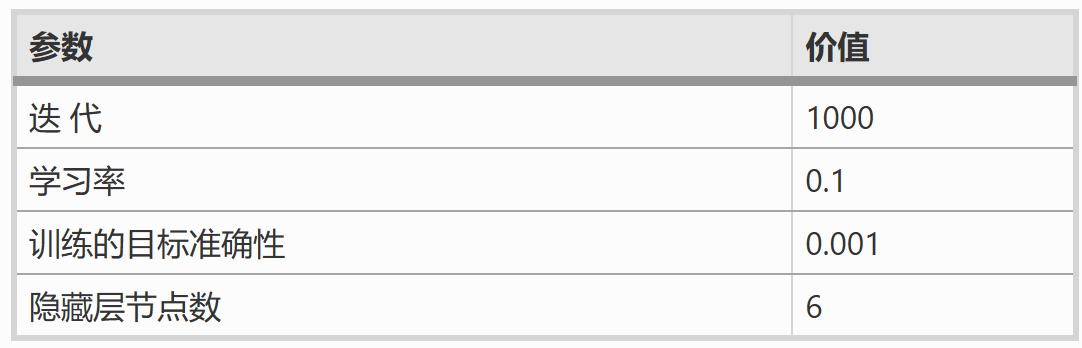
未处理的数据表现出明显的波动性，会影响预测效果。因此，EEMD是处理原始序列的一种手段，在此阶段之后，获得七个模态函数.与原始信号相比，分解后的内部模型函数结构更简单、更平滑。其次，利用PACF确定各模态函数的滯后周期数，选取与历史数据高度相关的滞后周期数对应的数据作为预测模型的输入。分解结果如图所示。确定各模态函数的预测输入后，通过PSOBP进行预测。



EEMD结果

5.2模型参数

在BP神经网络中，newff的作用是构建神经网络，然后用训练数据训练神经网络，sim用 来预测训练好的BP神经网络的输出。需要设置每个函数的参数。隐藏层节点数量太少，BP神经网络无法建立复杂的映射关系，网络预测误差大。但是，如果节点过多，网络学习时间会增加，可能会出现过拟合现象。本文拟合的非线性函数并不复杂，因此选择6作为隐藏层节点数，其余参数设置如表所示。



在PSO算法中，当粒子的速度和位置坐标趋于无穷大时，算法中控制参数的随机加权会导致爆炸。然而，目前的论文表明，可以通过限制步长或速度的参数并实施适当定义的收缩系数来防止爆炸。Popmax和Popmin表示未确定组合中每个分量的上限和下限，这隐含地限制了粒子的运动范围。此外，惯性权重的范围是[0.4,0.9]。PSOBP的参数设置在表中。

|  |  |
| --- | --- |
| 参数 | 值 |
| 迭代 | 40 |
| 人口规模 | 20 |
| C1 | 1.49445 |
| C2 | 1.49445 |
| Vmax | 1 |
| Vmin | -1 |
| popmax | 5 |
| popmin | -1 |

R2、RMSE和MAPE被广泛用作评估预测模型效果的指标。R2介于0和1之间。比率越大，模型越准确，预测效果越好。RMSE代表均方根误差，MAPE表示平均绝对百分比误差，其范围为[o,+∞]。值越大，误差越大。当预测值与真实值完全一致时，它等于0。他们的计算公式如下:

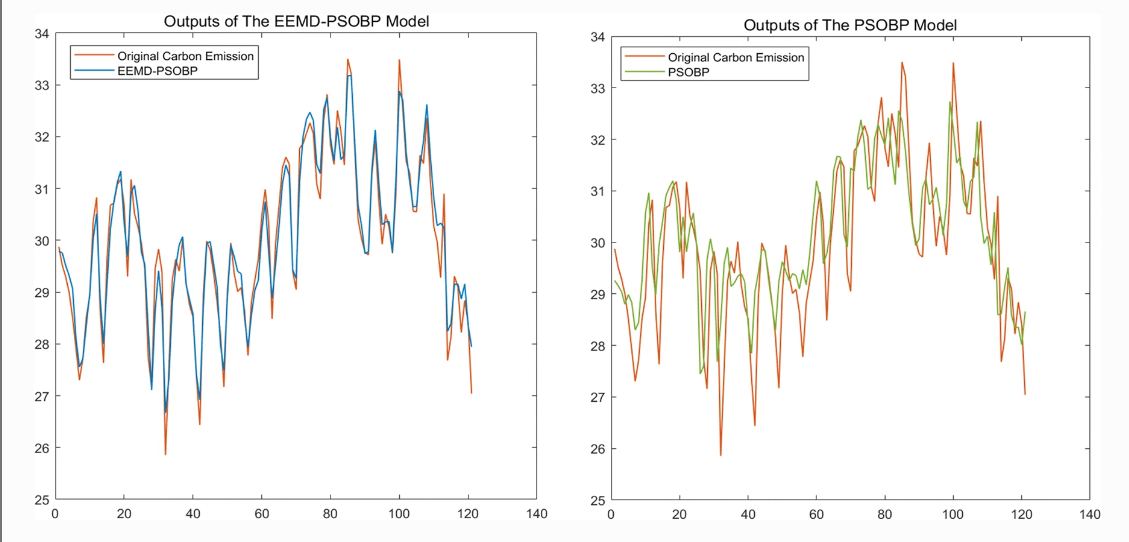




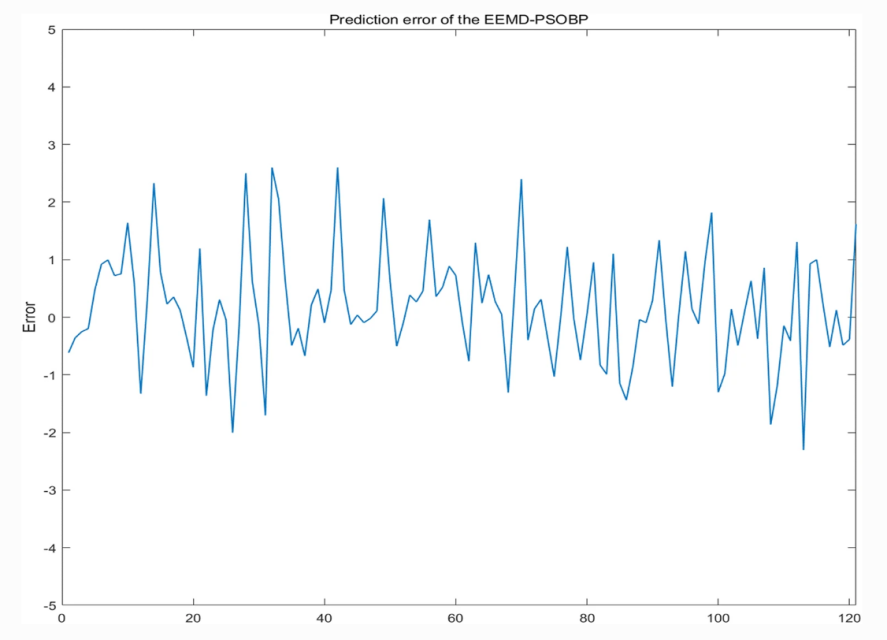


1. 结果与讨论

下图显示了模型的预测效果



最优模型（EEMD-PSOBP）最终预测结果



EEMD-PSOBP预测误差

通过预测结果可以发现:

(1)从任何评价标准 来看，本文采用的EEMD-PSO-BP模型的预测效果最好(R2= 0.5812，RMSE= 0.9583，MAPE =0.026) ,证明了利用这种组合模型预测短期碳排放的可行性和优越性。该模型可为短期碳排放预测提供参考方法，丰富短期碳排放预测的研究。上图显示误差在0附近波动，表明训练数据没有过度拟合，证明结果是有效的。

(2)优化后 的BP神经网络可以提高预测精度。就R2、RMSE和MAPE，而言，PSOBP的预测效果优于BP。优化后的模型都比原始模型具有更高的预测精度。

**灰色模型预测**

所给的历史数据时间悠久，让我们预测未来的江苏省碳排放进行预测，首先我们可以采用的是灰色预测，但是该方法预测的年限较短。据此我们可以采用GM(1,1)模型来进行预测。该模型操作简单，对于数据较少的情况下具有较好的优越性。这里我们对全球未来每年的气温均值进行预测。

GM（ 1,1）模型是生成模型， 在求解问题时需要生成相应的生成数列。设为原始数列，为做r次累加后的生成数列，即

=（ (1)，（2），··· ，（n） ） (6-1)

= ((1)，(2)，··· ，(n) ) （6-2）

则有累加生成式为：

（k）=(1)+(2)+···+(k) (6-3)

=

=+)

对于，其逆累加生成计算式为：

（（k））=（（（k）））—（（k-1）） （ 6-4）

=（k）—（k-1）

=)

利用GM(1,1)的十维灰色预测模型,为接下来的气温进行预测。

设=[(1)，（2），··· ，（5），（10）]

=[x1 x2 x3….]为样本数据序列,分别表示近十年全球气温情况均值情况,然后对序列进行一阶累加生成,生成序列，即（k）＝　（＝1,2,…,10），得到序列。

GM(1，1)预测模型是一阶单变量的灰色微分方程动态模型

+=b (k=1，2，...，10)

其中是的紧邻均值生成，上式的白化方程形式为+=b

对应用最小二乘法确定模型参数，求得:

==..，

其中B=

……

，

＝[(2)，（3），··· ，（10）]

通过MATLAB编程求出预测值。

灰色模型预测值

|  |  |  |  |
| --- | --- | --- | --- |
| **城市** | **预测值** | **城市** | **预测值** |
| **南京** | 22.5000 | 泰州 | 25.5077 |
| **无锡** | 18.4010 | 盐城 | 13.5156 |
| **徐州** | 24.1144 | 扬州 | 21.8197 |
| **常州** | 20.1807 | 南通 | 20.4127 |
| **苏州** | 16.5911 | 宿迁 | 19.2874 |
| **镇江** | 23.3373 | 淮安 | 18.4370 |
| **连云港** | 20.4108 |  |  |
| **8** | 17.8035 |  |  |

**时间序列分析法**

我们采用时间序列分析法中的移动平均法，来预测未来的温度情况。

Step1： 设观测序列为时间序列，取移动平均的项数。一次移动的平均值计算公式为

Step2：二次移动平均值计算公式

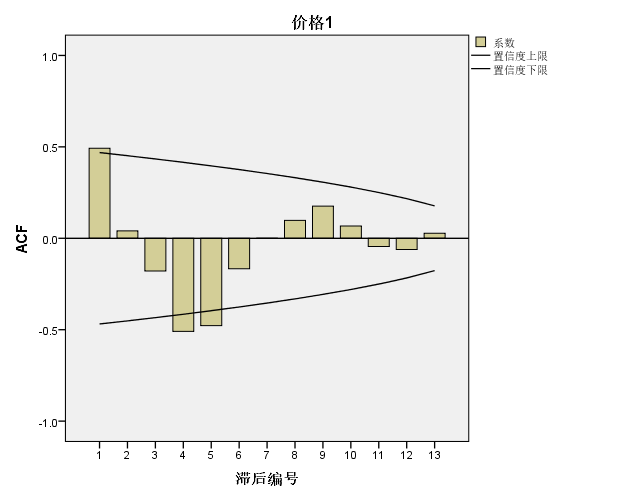
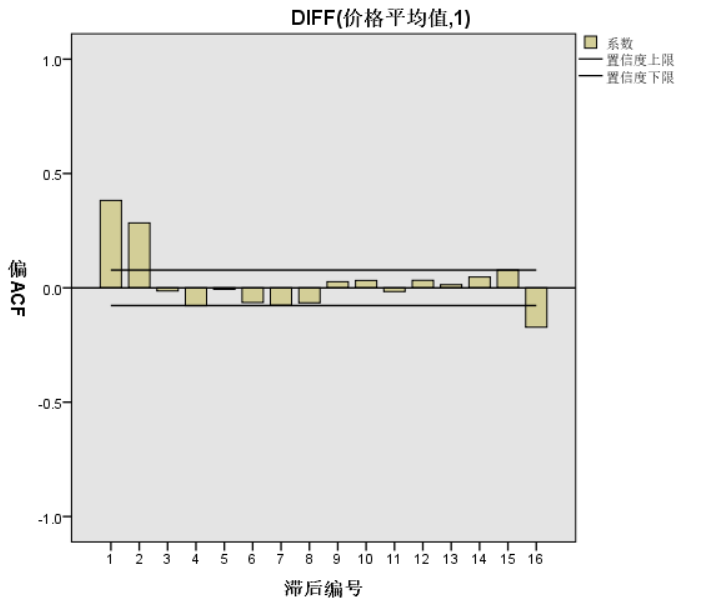
Step3： 当预测目标的基本趋势是在某一水平上下波动时，我们可以用一次移动平均方法建立模型进行预测分析，即

，

Step4:其预测标准误差为

我们以最近N期序列值的平均值作为未来各期的预测结果。当历史序列的基本趋势变化不大且序列中随机变动成分较多时,N的取值应该较大一些，否则N的取值应该小一些。在有确定的季节变动周期的资料中，移动平均的项数应选取周期长度。选择最佳N值的一个有效方法是,比较若干模型的预测误差，其中预测标准误差最小者为好。

借助SPSS软件进行求解，首先进行序列图分析，由于P值大于0.05我们对其做残差处理。



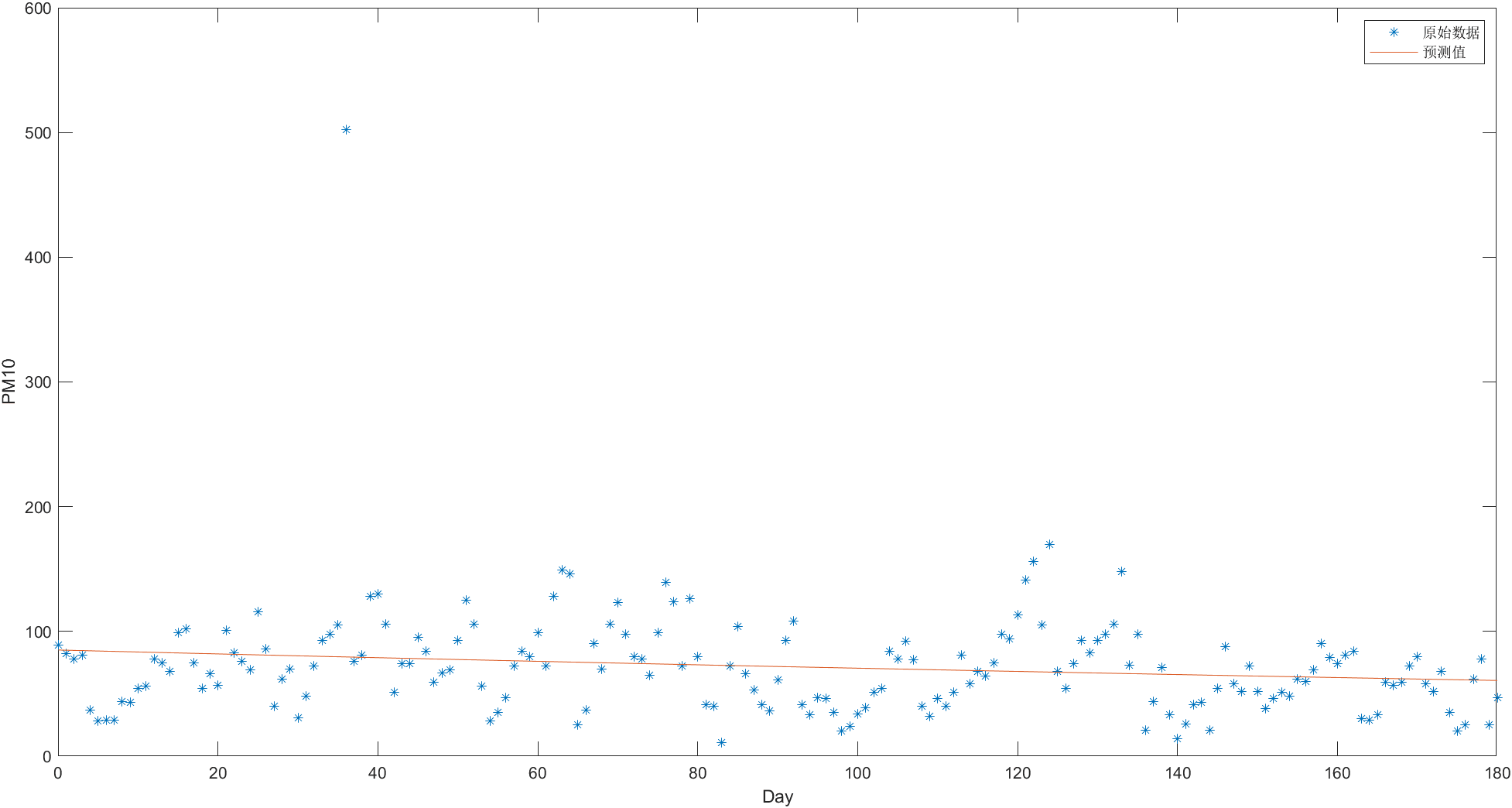
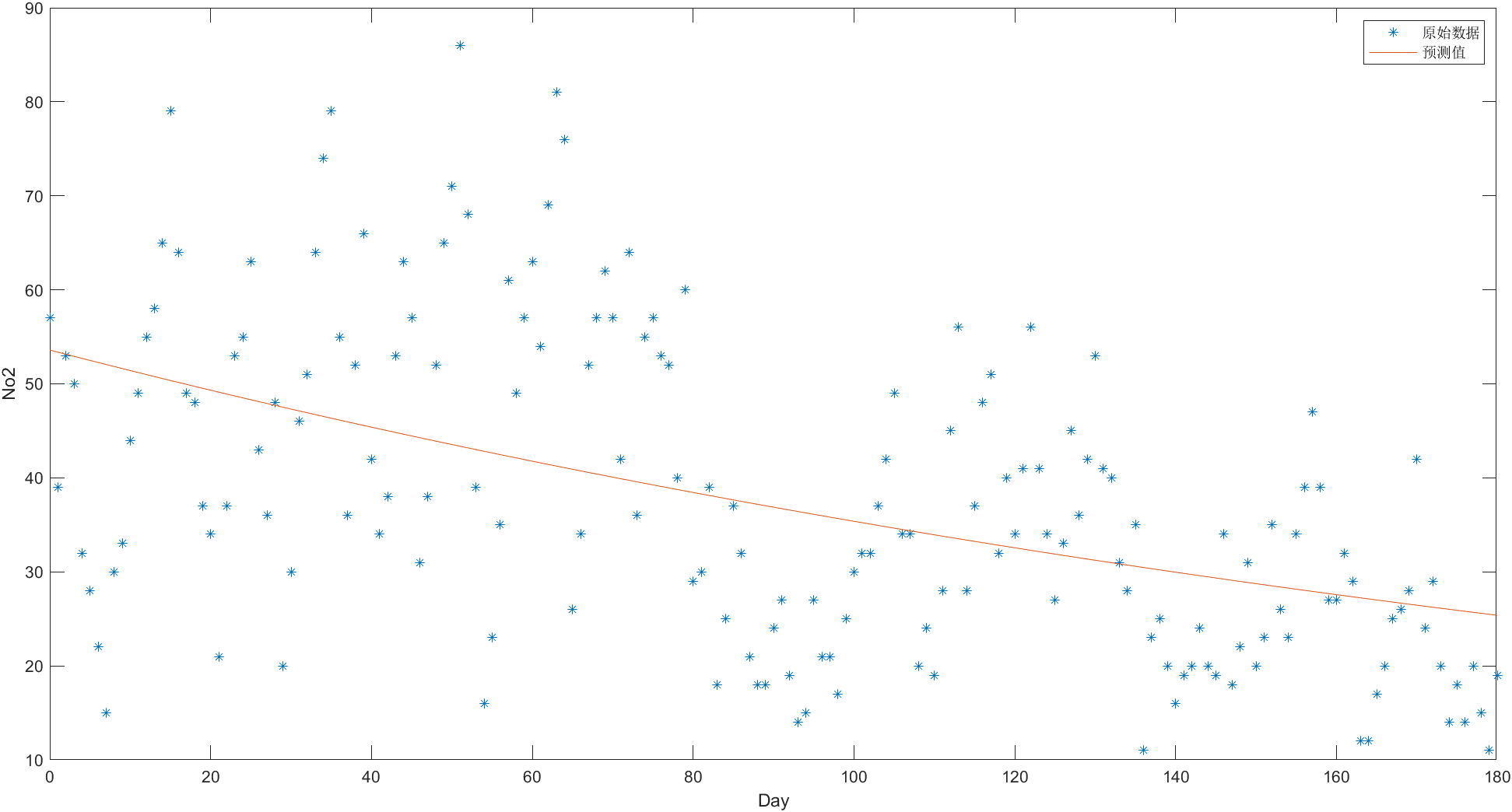
处理前、后自相关系数

拟合度

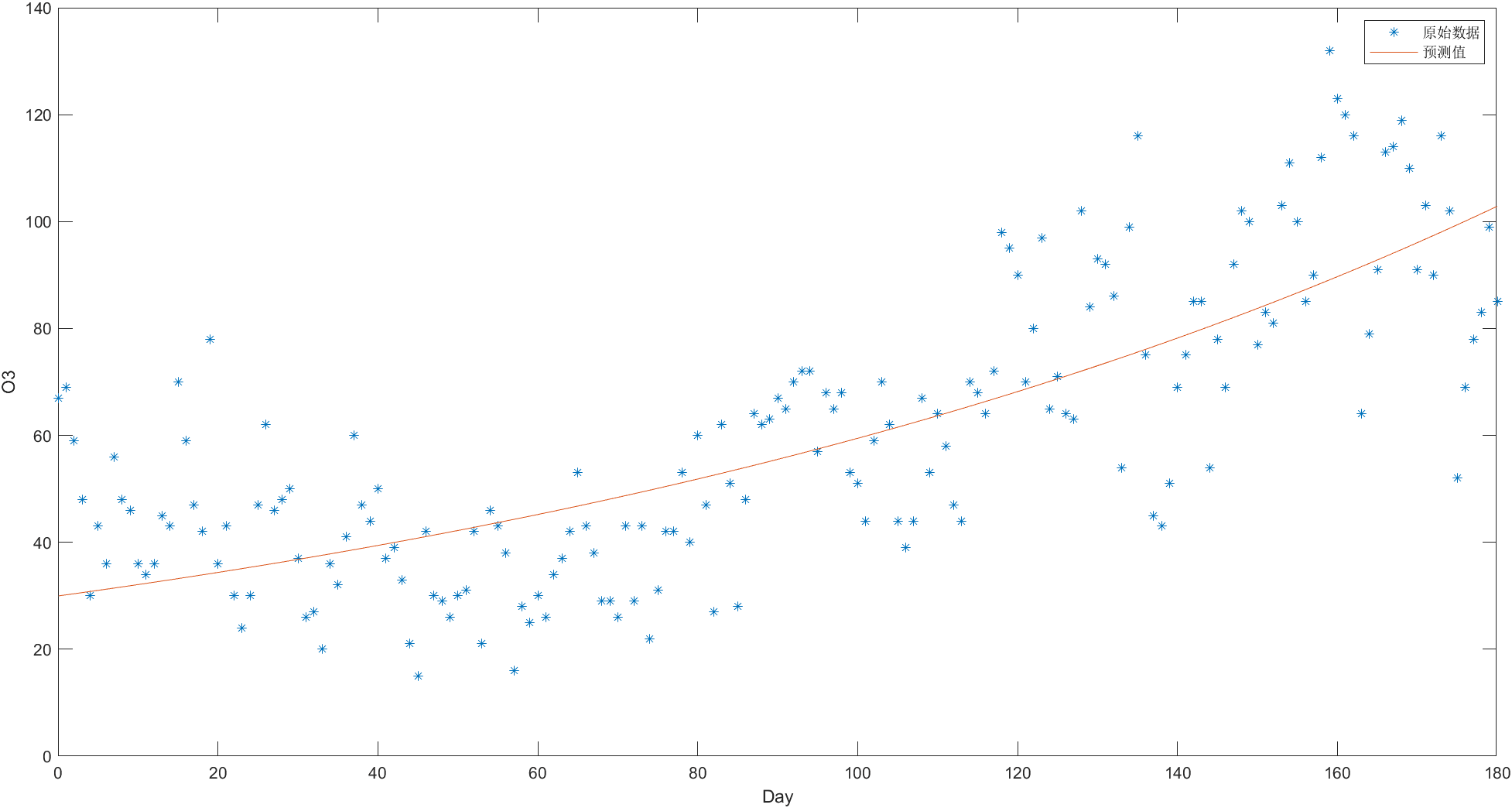
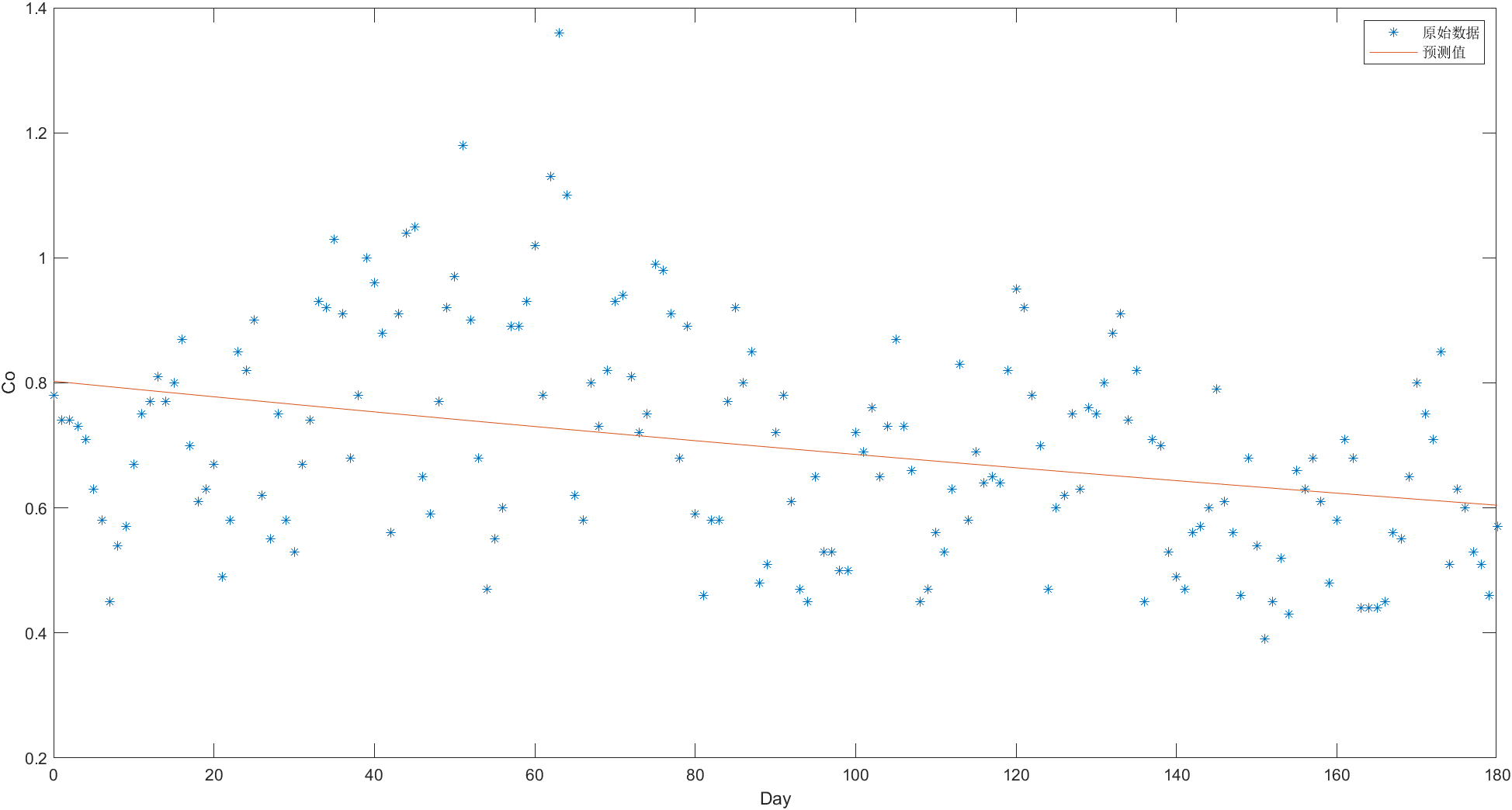


对比预测值与实际值，我们以2004年到2021年的历史数据作为验证，分别画出历史记录数值与预测温度的曲线，从图中我们可以看出模型预测效果较好。

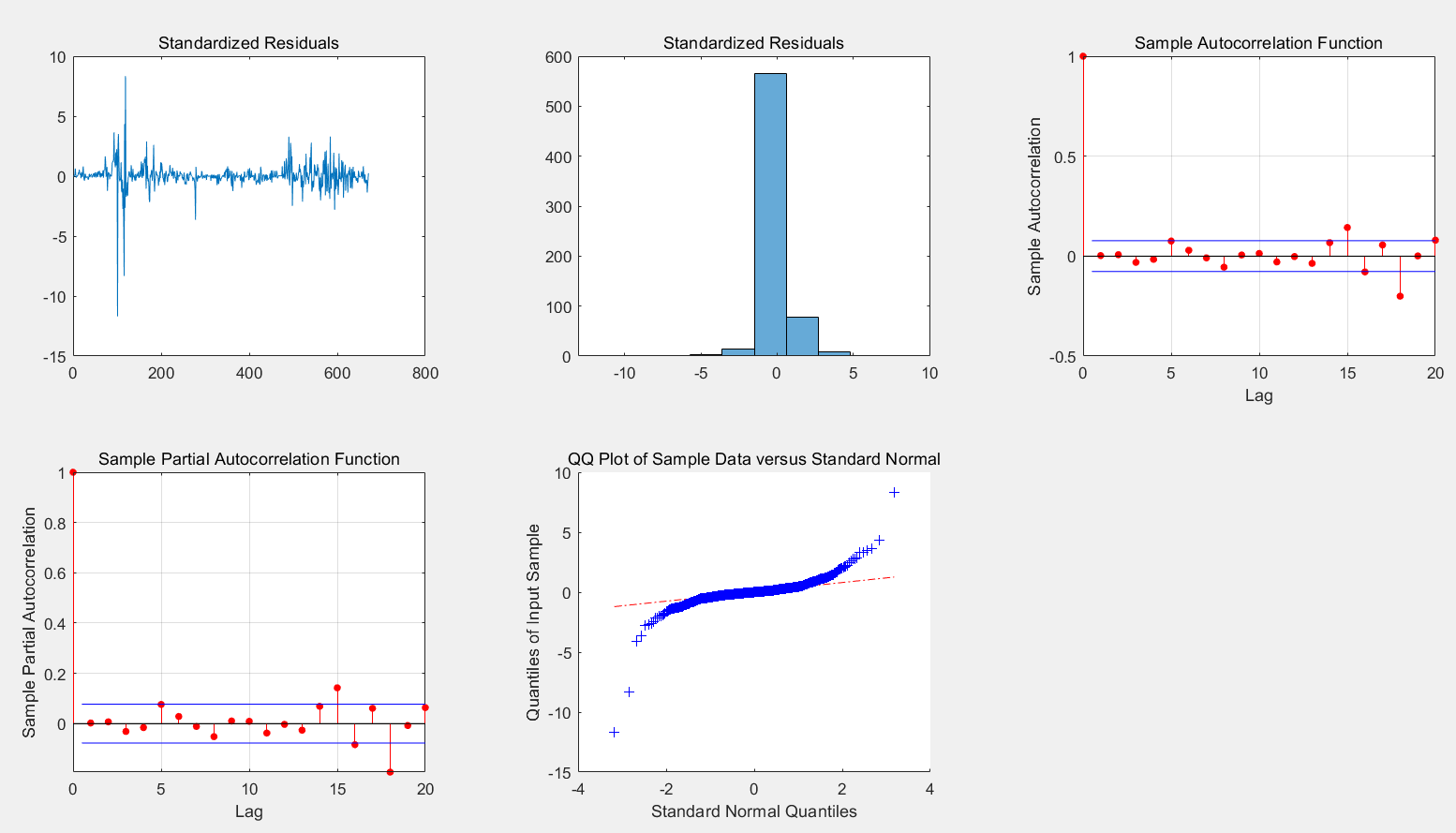
利用MATLAB绘制出估测曲线与真实值的对比图如下所示：（这里只选取一些更为重要的影响因素，其重要性大小已由第一问给出。）

**南京估测图** 徐州**估测图**

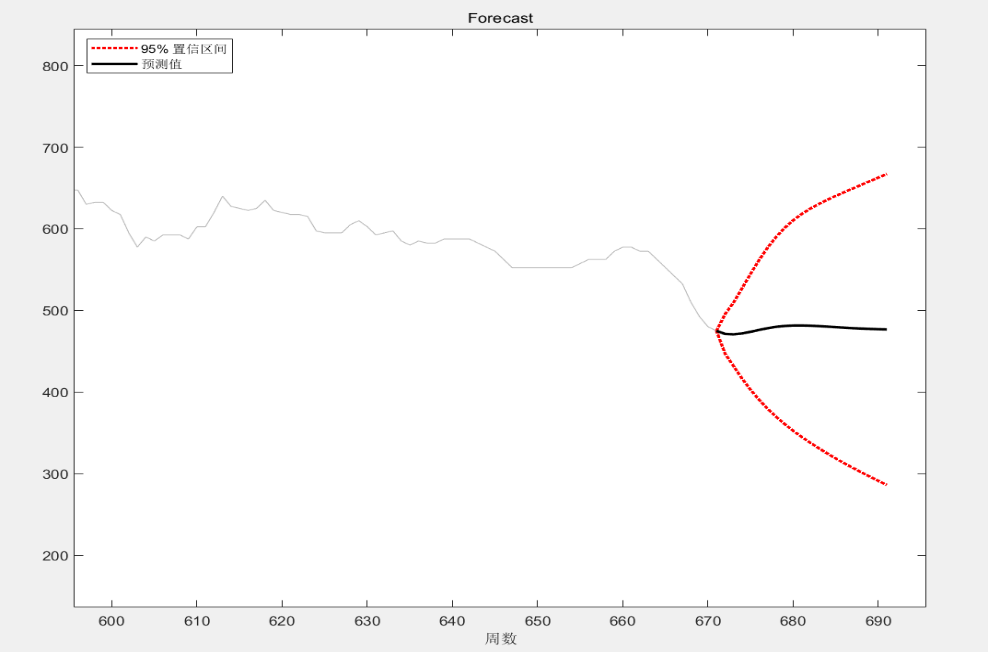


**扬州估测图** **连云港估测图**



残差检验

经残差检验，模型符合要求，可以进行预测



1. 模型评价

(1)

EEMD首次用于碳排放预测。考虑到样本数据的波动性，创新性地将“先分解后预测”的研究思路应用于碳排放预测领域。实验结果表明，EEMD分解后预测精度将大大提高，丰富了EEMD研究，也为碳排放预测精度的提高提供了改进方向。

(2)

创新性地提出了EEMD-PSOBP模型。该模型实现了碳排放的短期预测，丰富了短期碳排放预测的研究，有助于政府更清晰地了解碳排放的变化趋势，为制定短期碳排放目标提供了有力

的证据。该模式有潜力推广到其他碳排放量高的地区、行业或行业，制定更精准的减排措施。

参考文献

自己找几篇相关的加上

**附录**

|  |
| --- |
| function rhat=hecheng(r);  k=length(r);  for i=1:k  for j=1:k  rhat(i,j)=max(min([r(i,:);r(:,j)']));  end  end  end  clc  clear  close all  a=xlsread('指数.xlsx','B2:F7');  [m,n]=size(a);%求矩阵的行数和列数  amin=min(a,[],2);%计算每一行的最小值  amax=max(a,[],2);%计算每一列的最大值  b=(a-amin)./(amax-amin)%进行极差标准化  for i=1:n  for j=1:n  r(i,j)=sum(min([b(:,i)';b(:,j)']))/sum(max([b(:,i)';b(:,j)']));  end  end  r(2,2)=1;  r%显示相似矩阵  r1=hecheng(r);%进行一次合成运算  r1=hecheng(r1);%进行第二次合成运算  r1=hecheng(r1);%进行第三次合成运算  ur=unique(r1);%求等价矩阵中的所有不同元素  ur=sort(ur,'descend')%把等价矩阵中的元素按照从小到大排列  R2=(r1>=ur(2))%求关于ur(2)的lamda截距阵  R3=(r1>=ur(3))  R4=(r1>=ur(4))  d=1-r1;d=tril(d);%根据相似系数计算距离矩阵，并截取矩阵的下三角元素  d=nonzeros(d);d=d';%提取linkage需要的距离数据  S={'1-南京','2-苏州','3-徐州','4-扬州','5-连云港'};  z=linkage(d);dendrogram(z,'label',S)%画聚类图 |

|  |
| --- |
|  |
| clc  clear  close all  %% 导入数据  D=xlsread('主成分的分值.xlsx','Sheet1','B2:E182');  AQI=D(:,4);  y=D(:,1:3);  y1=y(:,1);  y2=y(:,2);  y3=y(:,3);  %% 多元线性回归预测模型  n=181;%样本数  m=3;%变量数  Y=[ones(n,1),y];  [b,bint,r,rint,state]=regress(AQI,Y);  b%回归系数  bint%回归系数的区间估计  state%[相关系数R^2,F值,与F对应的概率p,误差方差]  rcoplot(r,rint)%r:残差 rint:残差置信区间  clc  clear  close all  %% 去除异常点37、71、72、77、86、92导入数据  D=xlsread('主成分的分值.xlsx','Sheet2','B2:E176');  AQI=D(:,4);  y=D(:,1:3);  y1=y(:,1);  y2=y(:,2);  y3=y(:,3);  %% 第二次拟合  n=175;%样本数  m=3;%变量数  Y=[ones(n,1),y];  [b,bint,r,rint,state]=regress(AQI,Y);  b%回归系数  bint%回归系数的区间估计  state%[相关系数R^2,F值,与F对应的概率p,误差方差]  rcoplot(r,rint)%r:残差 rint:残差置信区间  clc  clear  close all  %% 在第一次剔除异常点的基础上，再去除异常点26、81、114、117、120、156导入数据  D=xlsread('主成分的分值.xlsx','Sheet3','B2:E170');  AQI=D(:,4);  y=D(:,1:3);  y1=y(:,1);  y2=y(:,2);  y3=y(:,3);  %% 第三次拟合  n=169;%样本数  m=3;%变量数  Y=[ones(n,1),y];  [b,bint,r,rint,state]=regress(AQI,Y);  b%回归系数  bint%回归系数的区间估计  state%[相关系数R^2,F值,与F对应的概率p,误差方差]  rcoplot(r,rint)%r:残差 rint:残差置信区间  clc  clear  close all  D=xlsread('.xlsx','C2:J182');  AQI=D(:,1);  X=D(:,3:8);  P=[-0.451 -0.007 0.462;  -0.392 0.210 0.510;  -0.310 0.557 -0.593;  -0.485 -0.061 -0.329;  -0.501 -0.052 -0.005;  0.245 0.800 0.258];  y=X\*P;  y1=y(:,1);y2=y(:,2);y3=y(:,3);  pred=12.1990-0.3960\*y1-0.0888\*y2+0.6762\*y3;  plot([1:181],AQI,'ro-')  hold on  plot([1:181],pred,'b+-')  title('估测值与观测值对比') |

clc

clear

close all

X=xlsread('1.xlsx');

X=X(:,5:10);

z=zscore(X); %数据标准化

M=cov(z); %协同方差

[V,D]=eig(M); %求出协方差矩阵的特征向量、特征根

d=diag(D); %为一组的贡献率

eig1=sort(d,'descend'); %将贡献率按从大到小元素排列

v=fliplr(V); %对特征向量重新排列

S=0;

i=0;

while S/sum(eig1)<0.85

i=i+1;

S=S+eig1(i);

end %计算对贡献率大于85%的部分

NEW=z\*v(:,1:i); %输出从而产生的数据

W=100\*eig1/sum(eig1)

figure(1)

pareto(W); %对直方图的绘画

clc

clear

close all

x=xlsread('1.xlsx');

x=x(:,5:10);

[n,p] =size(x); % n为样本，p为指标

%%一、对数据X进行标准化处理

X=zscore(x); % mat1ab内置的标准化函数(x-mean(x))/std(x)

%%二、对所运行的样本进行协方差矩阵的计算

R=cov(X);

%% 上面的两步可合并下面的一步直接对相关系数矩阵进行计算

R=corrcoef(x);

disp('样本相关系数矩阵为: ')

disp(R)

%% 三、对特征值和向量的计算机（R）

%R是半正定矩阵， R同时是对称矩阵，

[V,D]=eig(R); % 特征向量矩阵V 特征值构成的对角矩阵D 对特征值和特征向量进行自动计算

%% 四、对贡献率和累计贡献率的计算

ll=diag(D); %本函数可以计算矩阵的主对角线元素值

ll=ll(end:-1:1);

contribution\_rate=ll/sum(ll); % 对贡献率进行计算

cum\_contribution\_rate=cumsum(ll)/ sum(ll); % 对累计贡献率进行计算

figure(1)

pareto(100\*contribution\_rate); %绘画贡献率的直方图(前95%)

disp('特征值为: ')

disp(ll')

disp('贡献率为: ')

disp(contribution\_rate' )

disp('累计贡献率为: ')

disp(cum\_contribution\_rate' )

disp('与特征值对应的特征向量矩阵为: ')

%特征向量要和特征值相互对应

V=rot90(V)';

disp(V)

%% 对所需要主成分的值进行计算

m=input('需要保存的主成分的个数: ');

F=zeros(n,m); %初始化保存主成分的矩阵

for i=1:m

ai=V(:,i)'; %取出第i个特征向量，并且将其转化为行向量

Ai=repmat(ai,n,1); %行向量重复n次，构成n\*p的矩阵

F(:,i)=sum(Ai.\*x,2); %需要计算每一行的和（标准化数据处理）

end

**附录一 灰色预测**

%原始数列 A

A = [572.50 572.50 562.50 552.50 542.50 532.50 510.00 492.50 480.00 475.00

];

n = length(A);

%对原始数列 A 做累加得到数列 B

B = cumsum(A);

%对数列 B 做紧邻均值生成

for i = 2:n

C(i) = (B(i) + B(i - 1))/2;

end

C(1) = [];

%构造数据矩阵

B = [-C;ones(1,n-1)];

Y = A;

Y(1) = [];

Y = Y';

%使用最小二乘法计算参数 a(发展系数)和 b(灰作用量)

c = inv(B\*B')\*B\*Y;

c = c';

a = c(1);

b = c(2);

%预测后续数据

F = [];

F(1) = A(1);

for i = 2:(n+5);

F(i) = (A(1)-b/a)/exp(a\*(i-1))+ b/a;

end

%对数列 F 累减还原,得到预测出的数据

G = []; G(1) = A(1);

for i = 2:(n+5);

G(i) = F(i) - F(i-1); %得到预测出来的数据

end

disp('预测数据为：'); G

%模型检验

H = G(1:5);

%计算残差序列

epsilon = A - H;

%法一：相对残差 Q 检验

%计算相对误差序列

delta = abs(epsilon./A);

%计算相对误差 Q

disp('相对残差 Q 检验：')

Q = mean(delta)

%法二：方差比 C 检验

disp('方差比 C 检验：')

C = std(epsilon, 1)/std(A, 1)

%法三：小误差概率 P 检验

S1 = std(A, 1);

tmp = find(abs(epsilon - mean(epsilon))< 0.6745 \* S1);

disp('小误差概率 P 检验：')

P = length(tmp)/n

**主成分分析**

function result=cwfac(vector);

fprintf('相关系数矩阵:\n')

std=CORRCOEF(vector) %计算相关系数矩阵

fprintf('特征向量(vec)及特征值(val)：\n')

[vec,val]=eig(std) %求特征值(val)及特征向量(vec)

newval=diag(val) ;

[y,i]=sort(newval) ; %对特征根进行排序，y 为排序结果，i 为索引

fprintf('特征根排序：\n')

for z=1:length(y)

newy(z)=y(length(y)+1-z);

end

fprintf('%g\n',newy)

rate=y/sum(y);

fprintf('\n 贡献率：\n')

newrate=newy/sum(newy)

sumrate=0;

newi=[];

for k=length(y):-1:1

**时间序列**

close all

clear all

Data\_SCIdx=xlsread('C:\Users\Nix\Desktop\2020 51MCM\2020-51MCM-Problem A\ARIMA时间序列\附件1-郑州港郑州历史降雨量.xlsx','平均','b2:b672');

data = Data\_SCIdx;

plot(data)

y\_h\_adf = adftest(data)

y\_h\_kpss = kpsstest(data)

Yd1 = diff(data);

yd1\_h\_adf = adftest(Yd1)

yd1\_h\_kpss = kpsstest(Yd1)

Yd2 = diff(Yd1);

yd2\_h\_adf = adftest(Yd2)

yd2\_h\_kpss = kpsstest(Yd2)

Y = diff( diff(data) );

figure

autocorr(Y)

figure

parcorr(Y)

max\_ar = 3;

max\_ma = 3;

[AR\_Order,MA\_Order] = ARMA\_Order\_Select(Y,max\_ar,max\_ma,2);

Mdl = arima(AR\_Order, 2, MA\_Order);

EstMdl = estimate(Mdl,data);

[res,~,logL] = infer(EstMdl,data);

stdr = res/sqrt(EstMdl.Variance);

figure('Name','残差检验')

subplot(2,3,1)

plot(stdr)

title('Standardized Residuals')

subplot(2,3,2)

histogram(stdr,10)

title('Standardized Residuals')

subplot(2,3,3)

autocorr(stdr)

subplot(2,3,4)

parcorr(stdr)

subplot(2,3,5)

qqplot(stdr)

diffRes0 = diff(res);

SSE0 = res'\*res;

DW0 = (diffRes0'\*diffRes0)/SSE0

step = 20;

[forData,YMSE] = forecast(EstMdl,step,'Y0',data); [forData,YMSE] = forecast(EstMdl,step,data);

lower = forData - 1.96\*sqrt(YMSE); %95置信区间下限

upper = forData + 1.96\*sqrt(YMSE); %95置信区间上限

figure()

c = 1;

plot(data(c:671),'Color',[.7,.7,.7]);

hold on

h1 = plot(length(data)-c+1:length(data)-c+1+step,[data(end);lower],'r:','LineWidth',2);

plot(length(data)-c+1:length(data)-c+1+step,[data(end);upper],'r:','LineWidth',2)

h2 = plot(length(data)-c+1:length(data)-c+1+step,[data(end);forData],'k','LineWidth',2);

legend([h1 h2],'95% 置信区间','预测值',...

'Location','NorthWest')

title('Forecast')

xlabel('');

ylabel('');

hold off