

承诺书

我们完全清楚，在竞赛开始后参赛队员不能以任何方式，包括电话、电子邮件、“贴吧”、QQ 群、微信群等，与队外的任何人（包括指导教师）交流、讨论与赛题有关的问题；无论主动参与讨论还是被动接收讨论信息都是严重违反竞赛纪律的行为。

我们以中国大学生名誉和诚信郑重承诺，严格遵守竞赛章程和参赛规则，以保证竞赛的公正、公平性。如有违反竞赛章程和参赛规则的行为，我们将受到严肃处理。

我们参赛选择的题号（从 A/B/C/D/E 中选择一项填写）： A

我们的报名参赛队号（12 位数字全国统一编号）： 4321

参赛学校（完整的学校全称，不含院系名）： 你的大学

参赛队员 (打印并签名): 1. 成员 A

2. 成员 B

3. 成员 C

指导教师或指导教师组负责人(打印并签名): 指导老师

(指导教师签名意味着对参赛队的行为和论文的真实性负责)

日期: 2024 年 08 月 24 日

(请勿改动此页内容和格式。此承诺书打印签名后作为纸质论文的封面，注意电子版论文中不得出现此页。以上内容请仔细核对，如填写错误，论文可能被取消评奖资格。)

赛区评阅编号：
(由赛区填写)

全国评阅编号：
(全国组委会填写)

2024 高教社杯全国大学生数学建模竞赛

编 号 专 用 页

赛区评阅记录（可供赛区评阅时使用）：

评 阅 人						
备 注						

送全国评阅统一编号：
(赛区组委会填写)

(请勿改动此页内容和格式。此编号专用页仅供赛区和全国评阅使用，参赛队打印后装订到纸质论文的第二页上。注意电子版论文中不得出现此页。)

基于 ARIMA 模型的物料需求预测与生产计划优化

摘要

本片论文对于产业上的物料需求问题进行了模型的建立与求解。工业的物料需求是一种时间序列数据，我们首先是对不同物料的数据进行分类，将未有需求的空缺值补全为 0，即数据的预处理。然后，我们使用了 ARIMA 时间序列模型对数据进行拟合，发现模型对于处于时间序列后期的数据预测效果衰减的程度很大，所以我们对模型进行略微的修改，即在每次预测时只预测下一周的数据，下一周时再将下一周的真实数据放入训练集中，再次对模型进行拟合。通过对模型的不断更新，最大程度的保证模型的预测能力以及对时间序列数据的契合度。

然后，我们发现模型在预测时，会出现缺货量，库存量交替波动的情况，为了保证服务水平，我们引入了安全库存的概念，即在原有的生产计划上适当增加一些生产指标，以减少货品缺货的情况，进而提高服务水平。

然后，我们通过对安全库存的参数调整，以及对模型参数的再次网格遍历法，得出服务水平与库存资金相平衡的参数组合，降低了企业可能存在的最大库存积压资金。

最后，对于生产货物的滞后问题，我们对模型先前的循环预测部分进行了逻辑修改，使得新的预测部分可以正确计算库存值，正在生产值等。同时，我们也根据 ARIMA 模型，将其 p 参数设置为 2，使得模型中使用了 2 个单位时间滞后值，符合题目意图。为了使服务水平与库存积压资金相平衡，我们重复了第三问的做法，通过不断调整安全库存的参数，ARIMA 模型的 d , q 参数，找出最佳平衡参数组合。

关键字： 时间序列 ARIMA 模型 单步预测 安全库存

目录

一、问题重述	3
1.1 问题一	3
1.2 问题二	3
1.3 问题三	3
1.4 问题四	4
二、问题分析	4
2.1 问题一	4
2.2 问题二	4
2.3 问题三	4
2.4 问题四	4
三、模型的假设	5
四、符号说明	5
五、模型的建立与求解	5
5.1 问题一	5
5.2 问题二	9
5.3 问题三	10
5.4 问题四	12
六、总结	15
参考文献	15

一、问题重述

电子产品制造企业面临以下问题：在多品种小批量的物料生产中，事先无法知道物料的实际需求量。企业希望运用数学方法，分析已有的历史数据，建立数学模型，帮助企业合理地安排物料生产。

1.1 问题一

请对历史数据进行分析，选择 6 种应当重点关注的物料（可从物料需求出现的频数、数量、趋势和销售单价等方面考虑），建立物料需求的周预测模型（即以周为基本时间单位，预测物料的周需求量），并利用历史数据对预测模型进行评价。

1.2 问题二

如果按照物料需求量的预测值来安排生产，可能会产生较大的库存，或者出现较多的缺货，给企业带来经济和信誉方面的损失。企业希望从需求量的预测值、需求特征、库存量和缺货量等方面综合考虑，以便更合理地安排生产。

请提供一种制定生产计划的方法，从第 101 周开始，在每周初，制定本周的物料生产计划，安排生产，直至第 177 周为止，使得平均服务水平不低于 85%。

这里假设：本周计划生产的物料，只能在下周及以后使用。为便于统一计算结果，进一步假设第 100 周末的库存量和缺货量均为零，第 100 周的生产计划数恰好等于第 101 周的实际需求数。

请在问题 1 选定的 6 种物料中选择一种物料，将其第 101 □ 110 周的生产计划数、实际需求量、库存量、缺货量和服务水平按表 1 的形式填写，放在正文中。

请将问题 1 中选定的 6 种物料的全部计算结果（第 101 □ 177 周）按表 1 的形式填写在 Excel 文件中，通过支撑材料提交。请将 6 种物料的综合结果（第 101 □ 177 周的平均值）按表 2 的形式填写。

1.3 问题三

考虑到物料的价格，物料的库存需要占用资金。为了在库存量与服务水平之间达到某种平衡，如何调整现有的周生产计划，并说明理由。

请根据新的周生产计划，对问题 1 选定的 6 种物料重新计算，并将全部计算结果以表 1 的形式填写在 Excel 表中，通过支撑材料提交，将综合结果按表 2 的形式填写，放在正文中。对问题 2 选择的 1 种物料，将其第 101 □ 110 周的生产计划数、实际需求量、库存量、缺货量和服务水平按表 1 的形式填写，放在正文中。

1.4 问题四

如果本周计划生产的物料只能在两周及以后使用，请重新考虑问题 2 和问题 3。能否将你们的方法推广到一般情况，即如果本周计划生产的物料只能在 $\square (\geq 2)$ 周及以后使用，应如何制定生产计划。

二、问题分析

2.1 问题一

对于所给的历史数据，我们首先需要对数据进行处理，去除一些总需求量较少的物料，以降低数据的噪声。

然后，我们对预处理的数据使用 `scikit` 库作出周数-需求量的图，并对每一种物料的需求数据进行快速傅里叶变换分析和自相关分析，根据所做出的图标，以及频数，数量，趋势，每周平均销售单价，周期性等选择出 6 种应当重点关注的物料。

2.2 问题二

为了保证服务水平不低于 85%，需要尽量减少缺货的情况，于是可以使用具有防范风险的安全库存模型来对模型进行补充，在每次生产时适当增加产量以保证大部分需求能被满足。

2.3 问题三

由于第二问只考虑了服务水平，没有考虑到库存的积压资金，所以我们先对库存积压资金做出定义，然后再对安全库存参数做出调整，以减少库存积压资金的同时保证服务水平。

2.4 问题四

我们先根据 ARIMA 模型的参数性质，对 p 参数进行调整，即模型可以使用 p 个单位时间滞后值，然后再对安全库存参数做出遍历分析，找出最佳的平衡参数组合。

三、模型的假设

我们假定：

- 对于第 i 周的数据未有提及，则第 i 周的需求量为 0，销售单价为 NULL。
- 对于第 i 周的缺货量，会加入原本第 $i+1$ 周的需求量成为第 $i+1$ 周的实际需求量。
- 所有物料的数量，供需数量都为 0 或正整数。
- 对于库存物料的积压资金，等于该周货物的多余量 * 平均单价。

四、符号说明

符号	意义
SS	安全库存（件）
Z	安全系数
σ_{LT}	需求量的标准差

五、模型的建立与求解

5.1 问题一

对于预处理好的数据，我们先通过总需求次数是否大于 20 来将需求量过于小的物料排除；然后对剩下的物料数据作出周数-需求量周数-销售单价周数-盈利的折线图，通过分析再进一步排除需求出现较大空窗期的物料；然后再对剩下的物料作出简单的快速傅里叶变换分析，自相关分析判断其是否具有周期性。同时，使用 python 库 statsmodels 对物料的需求量进行多项式回归、指数平滑模型、时间序列模型进行简单拟合，初步判断物料需求量函数拟合的可行性。

对于企业来说，肯定是希望做一些需求更稳定的，需求量少，而销售单较高的物料来进行生产。为此我们从以上的 12 中物料中选取了 6004010256、6004020656、6004100008、6004020918、6004020503、6004020900 这 6 种物料进行进一步分析。其中，6004010256、6004100008、6004020900 的单价较高；6004020656、6004010256、6004020503 的需求量大；6004020918、6004020900 的平均单价稳定。

我们先对数据进行了 ADF 测试，判断数据是否平稳，然后对进行了差分的具有平稳性的数据使用 BIC（贝叶斯信息准则），AIC（赤池信息准则）计算选出最佳的 p, d, q 参数组合。然后将 70% 的总数据作为训练集，30% 的数据作为测试集，在训练集上

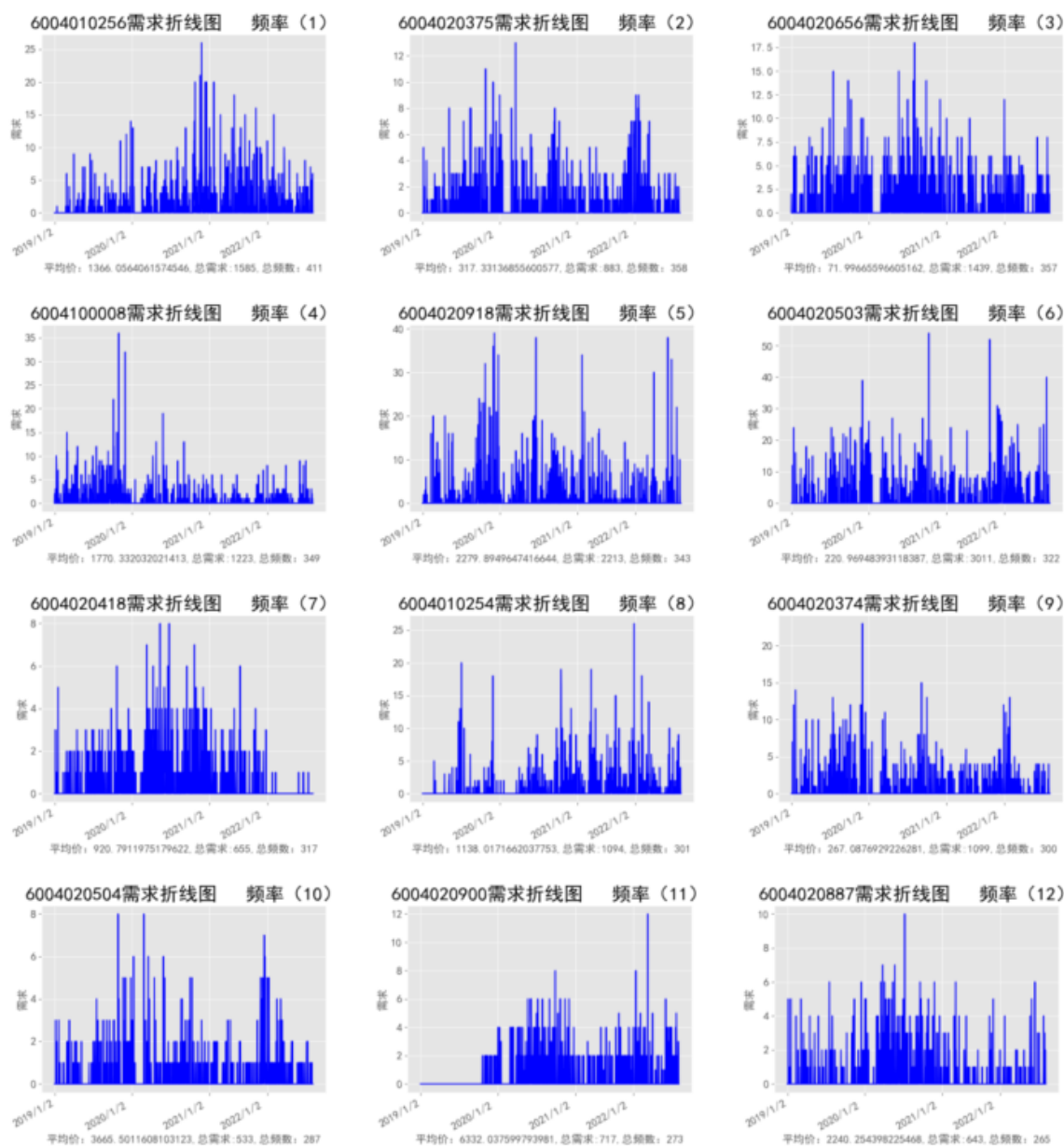


图 1 12 种预选的物料频率图

用得到的参数进行了基于最大似然估计的 ARIMA 模型的拟合，得到了最佳的 ARIMA 模型。接着使用模型对测试集进行预测，绘制出预测——周数、实际值——周数的二维图如下。

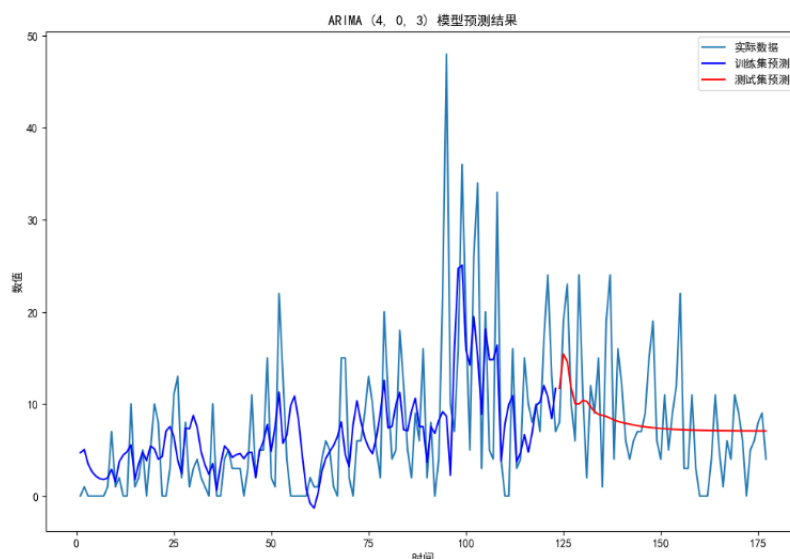


图 2 物料编码：6004010256

如上图片，我们发现，ARIMA 模型在训练集上的表现还比较好，但是在测试集上的表现出现较大偏差模型对于前期部分的测试数据的预测仍能基本跟上真实值的变化趋势，但是对于后期的数据预测效果较差，预测曲线基本趋平。于是，我们决定对 ARIMA 模型的预测部份进行修改，使它可以进行循环的“单步预测”，具体如下：

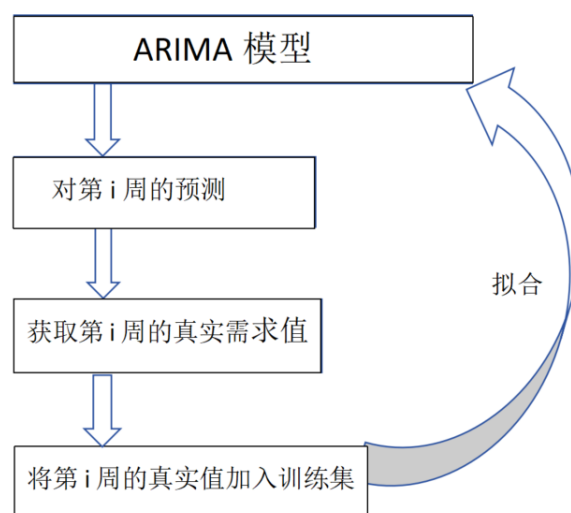
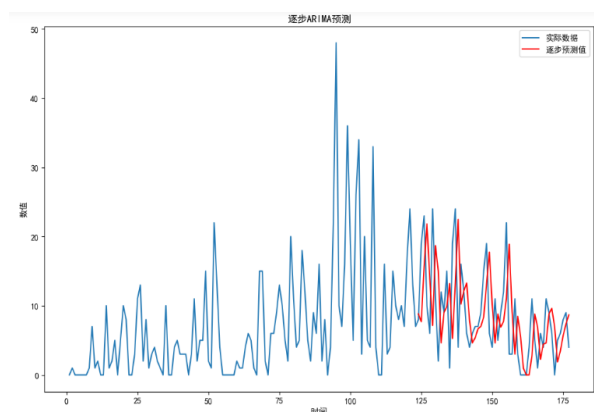


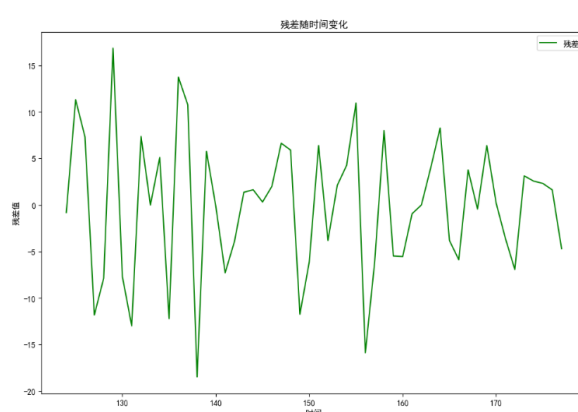
图 3 “循环单步预测”示意图

我们先用前 70% 的数据进行拟合，然后使用 ARIMA 模型对下一周进行预测，等到下一周获取到真实值后，将下一周的真实值加入到训练集中，再次使用 ARIMA 模型进行拟合，再对下周进行预测，如此循环往复。即每次模型只对下一周进行与预测，极大的保证了模型的性能优势与准确性。以下是我们使用了“循环单步预测”后的预测图

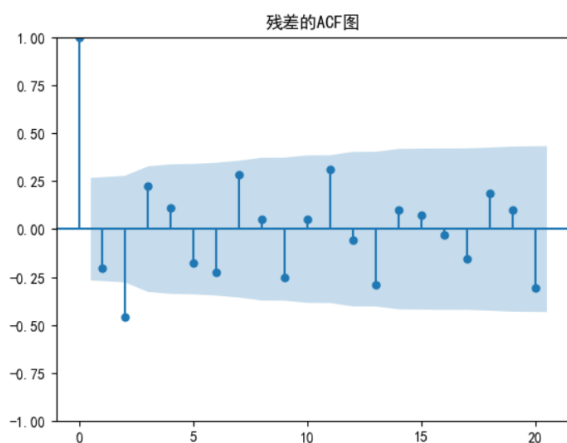
像：



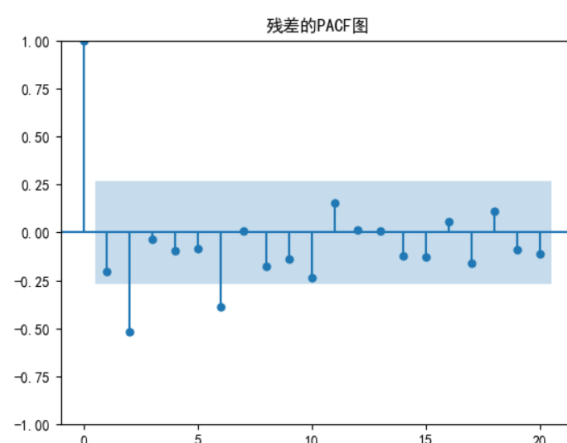
(a) 预测值与实际值对比图



(b) 模型残差图



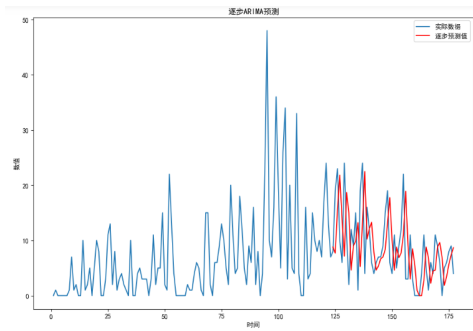
(c) 模型残差自相关 ACF 图



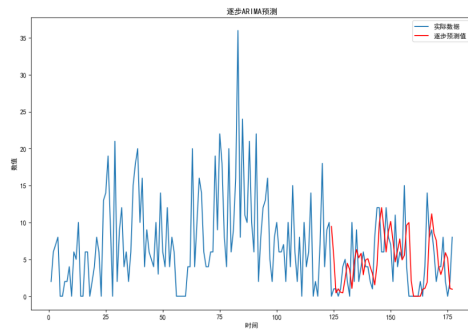
(d) 模型残差正态分布 PACF 图

可以看到，我们通过“循环单步预测”得到的 ARIMA 模型可以预测需求的趋势，且较为准确，预测曲线与实际需求曲线基本重合。同时，模型的残差图中的残差分布随机，没有明显的模式或者趋势。ACF 与 PACF 图中的值都是从 1.0 快速衰减到零区间附近，但也存在小部分“异常值”，尽管如此，所有的数值都是处于置信区间内，证明了“循环单步预测” ARIMA 模型在对物料需求上的预测准确性。

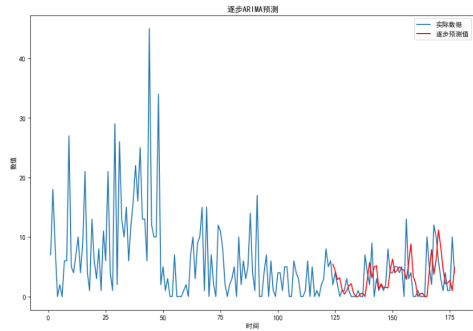
于是，我们将每个物料都进行了 ARIMA 模型的拟合，并使用循环的单步预测法进行预测，得到的结果如下：



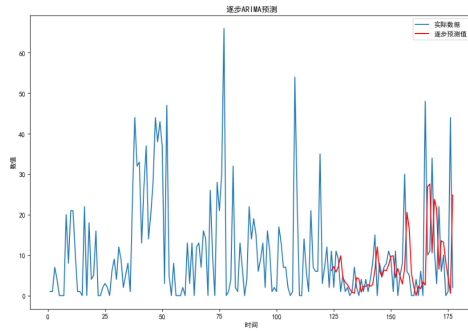
(a) 物料编码: 6004010256



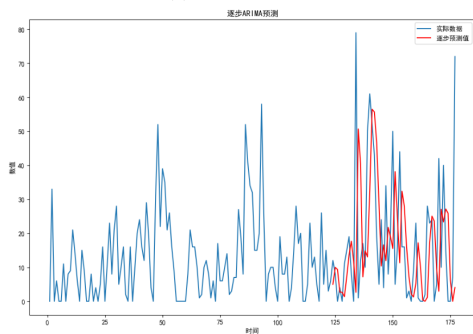
(b) 物料编码: 6004020656



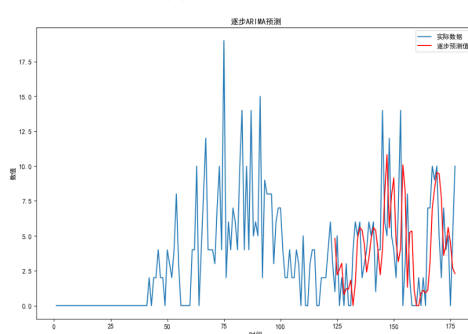
(c) 物料编码: 6004100008



(d) 物料编码: 6004020918



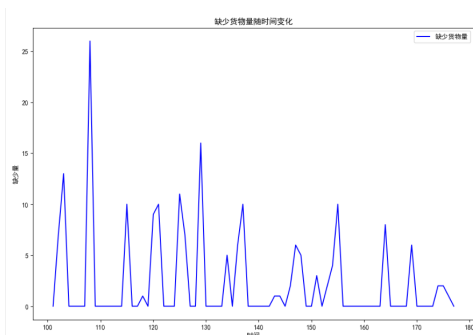
(e) 物料编码: 6004020503



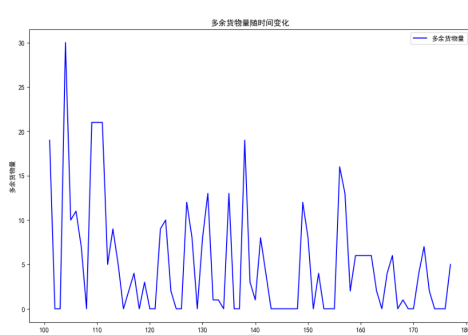
(f) 物料编码: 6004020900

5.2 问题二

题目要求从第 101 周开始预测，于是我们更改训练集与测试集，继续使用第一问得到的循环单步预测法进行预测。同时记录下缺货量，库存量等信息，如下：



(a) 缺失值



(b) 库存值

我们发现，对于大部分物料来说，缺货量与库存量总是处于较为随机的交替起伏的情况下，即使平均服务的水平达到了 85%，但是缺货量与库存量的波动仍然较大，这对于企业的信誉来说是不利的。于是，我们在预测的同时加入了安全库存的概念，可以在原有预测生产的基础上，适当增加一些生产指标，以增加企业面对突变的需求量的抗风险能力。

$$SS = Z \times \sigma_{LT}$$

其中， SS 为安全库存， Z 为安全系数， σ_{LT} 为需求量的标准差。

周数	生产计划/件	实际需求量/件	库存量/件	缺货量/件	服务水平
101	33	5	28	0	1.0
102	0	26	2	0	1.0
103	28	34	0	4	0.88
104	46	7	39	0	1.0
105	0	20	19	0	1.0
106	6	5	20	0	1.0
107	0	4	16	0	1.0
108	0	33	0	17	0.48
109	51	21	30	0	1.0
110	0	0	30	0	1.0
111	0	0	30	0	1.0

表 1 6004010256 物料第 101~110 周的生产计划、实际需求、库存、缺货量及服务水平

于是，在使用了安全库存之后，我们的平均满意度可以达到较高的水平。

5.3 问题三

对于题目所说的物料积压资金，我们认为可以用最大库存量时的所需最大库存积压资金来衡量模型预测在物料积压资金上面的水平。

$$\text{物料数量} \times \text{平均单价} = \text{所需最大库存积压资金} \quad (1)$$

因为对于企业来说，一般生产的预备资金是一直用于某个物料上的，不会流动，即增加或者减少。所以我们认为所需最大库存积压资金就代表了企业可以只放置这么多资金在

这个物料上，就能保证货物正常生产以及积压。

为了计算物料库存的积压资金，我们先对无物料的前 100 周的平均价格数据进行了分析，以下以物料编码 6004010256 的物料为例：

我们对物料前 100 周的价格数据进行了简单的统计分析和画图，发现价格在前 55 周处于相对波动的范围，在后 56 周至 100 周则相对稳定，由于我们计算的是第 101 周至 177 周的预算与货物的积压资金，所以我们选择使用后 45 周的平均价格作为计算货物积压的平均价格，即 1483.1 元。

在此基础上，以物料 6004010256 为例子，我们计算了物料库存的积压资金，如下：

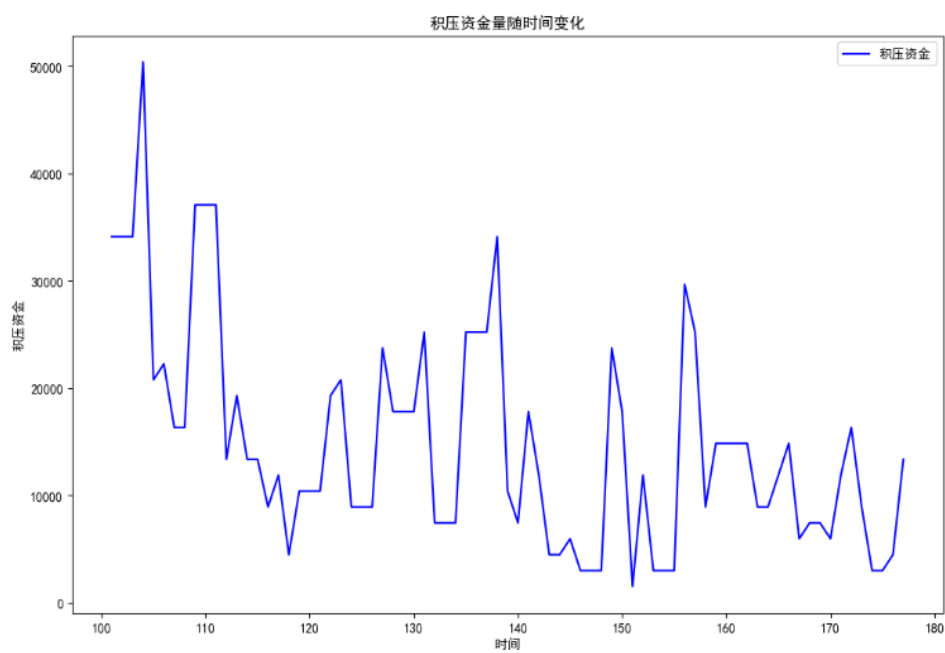


图 7 6004010256 物料在使用第二问模型下的库存积压资金-时间图

通过如上的图片，我们发现由于安全库存的设定导致我们的库存量基本保持在 1 以上，而且峰值也远高于大部分需求的值，这也就使得物料的库存积压资金过高，不符合企业的低资金积压目标。所以我们重新通过网格遍历法以及调整安全库存的服务满意度大小来调整生产计划。具体来说，我们重新通过网格遍历法搜索计算，先在每一种参数组合下求出最大的资金挤压值，然后对比每一组参数的最大资金积压值，选取最小的那一组参数作为最终的参数组合。

但是我们发现这样还不够，由于安全库存的影响过于强烈，导致我们的库存量过高，所以我们对安全库存的服务水平参数进行了调整，使得安全库存的影响减弱，同时保证题目中随定义的服务水平。

表 2 物料：6004010256 的安全库存与服务水平和最大库存积压资金的关系

安全库存	服务水平	最大库存积压资金
0	0.8946935172146779	28178.9
1	0.900971134024425	19280.3
2	0.9167137896136947	19280.3
3	0.9271877899399739	20763.4
4	0.9391972779102511	22246.5
5	0.9483656639510848	23729.6
6	0.9568084614732336	25212.7
7	0.9636698282973934	26695.8

可以看到，对于任意的安全库存值来说，服务水平都是在 85% 以上，但是随着安全库存的增加，最大库存积压资金会有些许不同，于是我们在保证服务水平尽可能高的情况下，选择安全库存的货物数量为 2。对于剩下的物料，我们也是先观察他们每周价格的波动情况，求出相对客观的平均价格；然后通过网格搜索求出最大库存资金中最小的的 p, d, q 参数值，最后遍历安全库存的大小，选择最大库存资金中最小的安全库存量即可。

5.4 问题四

我们使用了基于平均需求和供应提前期的安全库存的双步预测模型，即将原本模型的经典安全库存模型更改为基于平均需求和供应提前期的安全库存模型，将原本的每次只对下一周进行预测变为对下两周进行预测。并且对第 101 周和第 177 周的预测与计算进行了特殊处理。以此来适应题目的要求。

具体来说，我们在每一个单步预测的循环内：

1. 对第 i 周和第 $i+1$ 周的需求量进行预测。
2. 通过预测计算第 $i-1$ 周要生产的物料数量。
3. 获取第 i 周的实际需求量，计算第 i 周的缺货量，存货量。
4. 更新库存，加上第 $i-1$ 周的上产量。
5. 重复 1-4 步骤。

第 $(i - 1)$ 周要生产的物料数量

$$= \text{第}i\text{周的预测值} + \text{第}(i + 1)\text{周的预测值} - \text{安全库存} - \text{库存 (现有的可以用的货物)} \quad (2)$$

按照上述的步骤，我们更新了代码模块，重新进行了模型的拟合，得到了新的预测结果，如下图：

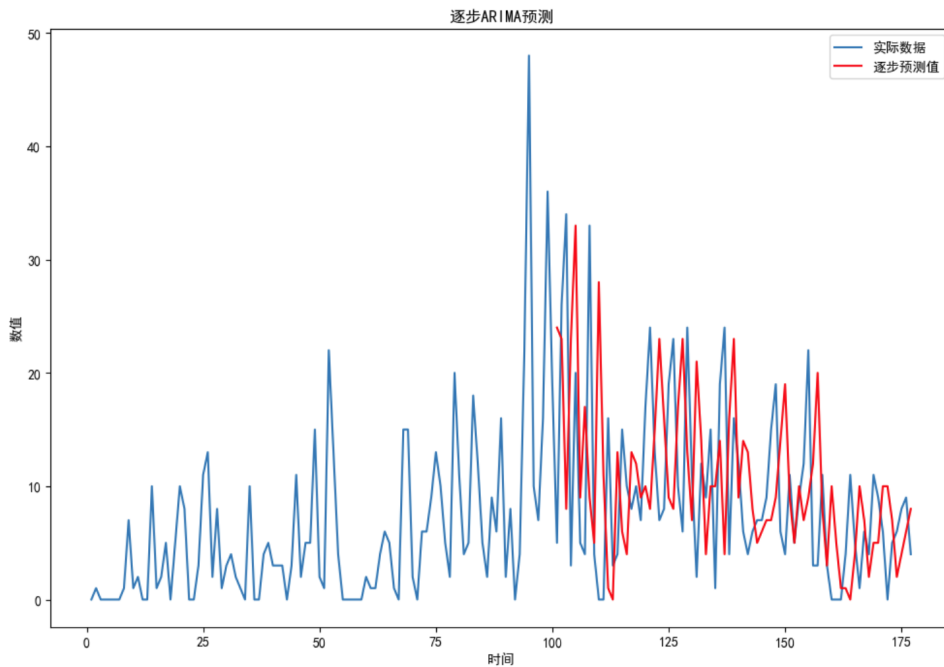


图 8 6004010256 物料在考虑生产滞后性的预测-时间图

通过如上的图片，我们发现，我们的模型在考虑了生产滞后性后，预测的效果依然较好，预测曲线与实际需求曲线基本重合，并且计算得出其服务水平参数为：88.9%，满足第二问对服务水平的要求。

同时，我们将我们先前的 ARIMA 模型的 p 参数进行了修改，将其值设定为 2。因为在 ARIMA 模型中， p 参数表示模型中使用了多少个时间滞后值。例如，如果 $p=1$ ，则模型会考虑前一期的值作为当前值的预测因素。我们将它设为 2 则可以符合题目条件要求“本周计划生产的物料只能在两周及以后使用”。但是实验数据却显示对于相同的 d, q 参数， $p=3$ 的模型的平均服务水平和最大物料积压资金值要优于 $p=1, p=2, p=4$ 的模型。此时，我们也对安全库存的数量进行了遍历，与 p 参数一起进行。最终我们选择了表现最佳的 $p=3$ ，与安全库存数量 $=6$ ，（注：3 倍平均需求量的库存积压值为 31145.1）具体如下：

表 3 物料 6004010256 在不同安全库存条件, 不同 p 值下的最大库存积压资金与服务水平

	安全库存数量	最大库存积压资金	服务水平
p,d,q=(2,1,1)	0	57840.9	0.8396
	1	45976.1	0.8342
	2	43009.9	0.8559
	3	43009.9	0.8725
	4	41526.8	0.8894
	5	41526.8	0.9042
	6	43009.9	0.9110
	7	44493.0	0.9196
p,q,d=(3,1,1)	0	60807.1	0.8374
	1	48942.3	0.8517
	2	34111.3	0.8604
	3	34111.3	0.8768
	4	35594.4	0.8893
	5	37077.5	0.8997
	6	38560.6	0.9075
	7	40043.7	0.9132
p,q,d=(4,1,1)	0	60807.1	0.8374
	1	48942.3	0.8472
	2	37077.5	0.8615
	3	37077.5	0.8738
	4	38560.6	0.8886
	5	37077.5	0.8977
	6	38560.6	0.9054
	7	40043.7	0.9135

在确保物料 6004010256 在生产滞后情况下的模型求解后，我们依照此思路对剩余的物料进行了模型求解，得到的结果如下：

表 4 六种物料的示意表格

物料编码	安全库存量	最大库存积压资金	平均服务水平
6004010256	6	38560.6	0.9075
6004020656	3	25212.7	0.9018
6004100008	4	17797.2	0.9004
6004020918	13	56357.8	0.9039
6004020503	16	75638.1	0.9010
6004020900	2	22246.5	0.9012

对于不同的物料，使用相同的我们的基于 ARIMA 时间序列模型的代码进行求解，最终获得低于 4 倍的平均需求量的最大积压资金，并且平均服务水平均达到 90% 以上，我们认为我们的预测模型具有一般性，即可以将我们的模型扩展到一般情况。对于不同的物料我们可以通过修改 ARIMA 模型的相关参数，或者对安全库存进行修改来对实际的情况进行拟合，最终达到较好效果。

六、总结

本文首先对物料的需求量等数据进行了数据预处理，然后进行了 ARIMA 模型的拟合，接着通过循环单步预测法对需求量进行了预测，运用了 ADF 验证等方式，说明了预测的准确性。然后我们引入了安全库存的概念，调整安全库存使得平均服务水平达到 85% 以上。接着通过调整安全库存的大小，以及对参数的调整，我们可以使得物料的最大库存积压资金达到最小，同时保证服务水平，达到了服务水平与积压资金的平衡。最后，我们对先前的 ARIMA 模型预测部份进行微调，使得模型可以适应生产滞后的情况，并借用第二三问思路，对安全库存量与参数进行遍历，得到服务水平与积压资金平衡的参数组合。最终在 6 种物料上进行了拟合，均达到很好效果，也说明了我们模型对于一般情况的适应性，即具有一般性。

参考文献

[1]

附录内容。