1. 问题四：主要操作变量的优化方案设计
   1. 问题四分析

针对问题四，其要求是在保证产品硫含量不大于5μg/g的前提下，对主要变量中的操作变量进行优化，以达到辛烷值损失尽可能小的结果，并且在优化后筛选出辛烷值损失降幅大于等于30%的样本。

问题四实际上是一个多约束的最优化问题。因此我们需要针对该问题要求，确定决策变量、优化目标和优化条件，从而建立一个优化模型，并选择一个合适的算法对模型进行求解，最后对模型结果作一些定性、定量的分析和必要的解释。

根据题意，我们可以得知，本题的所涉及的决策变量为问题二所选取的操作变量，共17个操作变量。本题的优化目标为，通过对决策变量的调整使辛烷值损失降幅最大，也就是辛烷值损失通过优化后的数值尽可能小，其中我们需要通过问题三所筛选出的最优秀的预测模型对辛烷值进行预测，同时，我们还要保证控制硫含量尽可能小，因此同样需要使用预测模型对硫含量进行预测。本题的约束条件为两类。第一类为对决策变量的约束条件，决策变量不能在调整过程中出现超出附件四中所规定的操作变量范围，一共有17个约束条件。第二类为对硫含量的约束条件，产品的硫含量最终应不大于5μg/g。

在本题中，选取了优化算法中的粒子群算法对该模型进行求解。粒子群算法属于启发式算法，能有效处理变量组合优化的问题，粒子群算法的适用范围较广，优化效果好，且算法较为简单，容易理解。因此，本文使用粒子群算法来求解模型。

* 1. 基于粒子群算法的优化模型

由于本题中需要保证硫含量不大于5μg/g，因此在进行模型求解前，我们需要对样本中硫含量已经大于5μg/g的样本进行剔除，最后共留下了266个样本用来建立优化模型。

* + 1. 优化模型建立

1. 建立决策变量

将问题二中选取的操作变量设为决策变量。即{ S-ZORB.FT\_1504.TOTALIZERA.PV、 S-ZORB.PDI\_1102.PV、S-ZORB.PC 2401.PIDA.OP、…、S-ZORB.PT 7508.DACA、S-ZORB.SIS PT\_2703}设为{}，记为Z={}

1. 建立目标函数

我们记最开始的辛烷值损失为，优化结束后，辛烷值损失为，因此为了得到辛烷值损失降幅最大的目标，我们设定目标函数为：



由于在操作变量中，我们选取了原料中的辛烷值作为操作变量，而没有选择产品中的辛烷值作为操作变量，同时，根据题意，优化过程中原料、待生吸附剂、再生吸附剂的性质保持不变因此我们可以将此目标函数改为：



这在实现粒子群算法的优化中可以在一定程度上提高算法的效率

1. 约束条件

1、决策变量的约束条件：

决策变量应当满足以下约束条件：



其中和为主要操作变量的上下限。

2、产品硫含量的约束条件：

产品中的硫含量s应当满足以下条件：



* + 1. 粒子群算法

粒子群算法（PSO）是针对群体的演化算法，其主要思想来源于鸟类的飞行研究。鸟类在进行群体飞行时，其具体移动方位并不是固定的，每只鸟都会根据群体的总体移动方向以及自身范围内的总体移动方向来进行移动，就好像，鸟的飞行过程不仅仅受到单一因素的影响，而是不同因素叠加的影响。这里以鸟类觅食为例，当鸟群去觅食时，对其中一个鸟来说，就它会在一定范围内随机搜寻食物，当它发现一个食物时，但它只知道这个食物与它的距离，但不知道方向，因此对这只鸟而言，找到食物最好的方法就是去往与这个食物最近的鸟的方向进行飞行，这样每只鸟的距离越来越近，直到最终找到食物。

在PSO算法中，将每只鸟定义为一个粒子，这种鸟的搜索空间定义为一个解空间，当鸟在飞行过程中，鸟与食物的距离定位为粒子的适应度，而在实际应用过程中，粒子的适应度值由目标函数进行提供，一般来说，当求解最小值问题时，适应度值越小越好。而鸟总是往食物方向飞行，因此适应度值也会慢慢减小。PSO算法认为，每个粒子有两个特征，分别为速度和位置，速度主要由本身经验和群体经验决定，既可以随即得到，也可以通过赋予权重获得最好参数。而位置的定义就较为复杂，由于粒子本身的运动不仅受自身的惯性影响，它还受到全局最优和个体最优的影响，就好比鸟在飞行过程中，不仅受到食物的影响，也会受到离食物最近的鸟的影响。同时，每个粒子必须保证始终个体最优，所以每个粒子都需要记住自己历史适应度值，从而在每一次迭代后进行判断，保持个体最优。

PSO算法将速度公式定义如下：



其中，我们假定有N个粒子，每个粒子在M维空间求解，则在第n次迭代过程中，第i个粒子的位置为，第i个粒子的速度为，此时个体最优为，群体最优为。

PSO算法的流程图大致如下：



* + 1. 实证研究

在本题中，我们将266个样本，通过粒子群算法来进行优化求解，其中我们对模型的参数设定如下表：

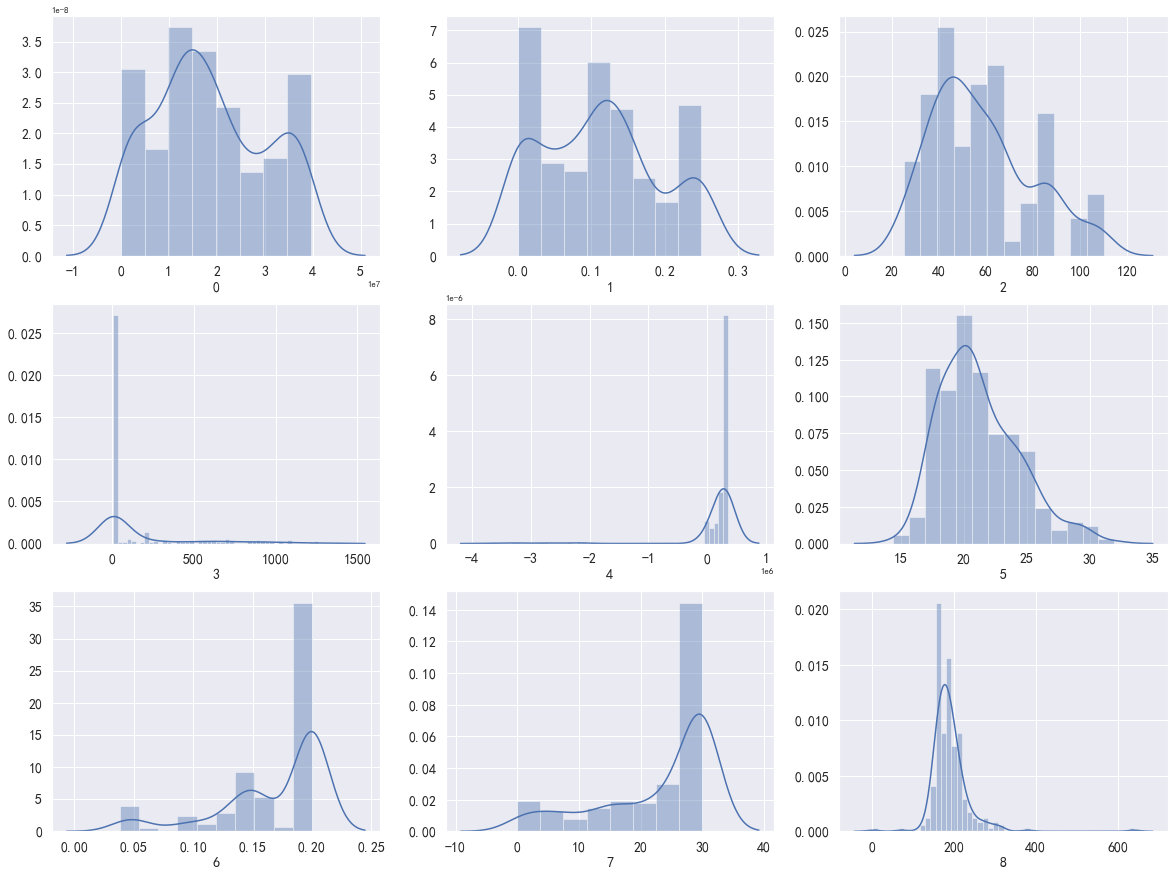
|  |  |  |
| --- | --- | --- |
| 参数 | 含义 | 值 |
| w | 惯性权重 | 0.5+0.4\*(i/I) |
| C1 | 学习常数 | 2 |
| C2 | 学习常数 | 2 |

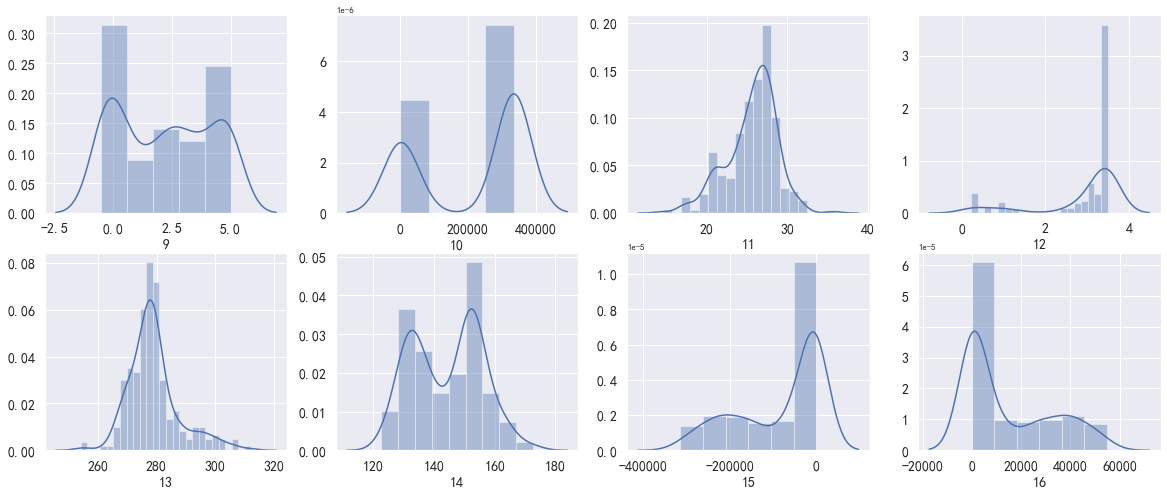
其中对于惯性权重w，将其值设定为0.5+0.4\*(i/I)，其中i为第i次迭代，I为总迭代数，这种设定是为了使惯性权重从0.5逐步递增到0.9，使该模型在前期有较好的全局搜索能力，后期有较好的局部搜索能力。

在266个样本中，有168个样本经过优化算法后，辛烷值损失降幅超过了30%。统计各个辛烷值（RON）损失降幅的频数和频率，如下所示：

|  |  |  |
| --- | --- | --- |
| 降幅范围(%) | 频数 | 频率 |
| (80,100] | 0 | 0.00% |
| (60,80] | 14 | 5.26% |
| (50,60] | 54 | 20.30% |
| (40,50] | 50 | 18.80% |
| (30,40] | 43 | 16.17% |
| (20,30] | 35 | 13.16% |
| (10,20] | 11 | 4.14% |
| (0,10] | 6 | 2.26% |
| 异常样本 | 0 | 0.00% |
| 总计 | 266 | 1 |

表中的异常样本表示该样本经过优化后硫含量超标，大于5μg/g。通过柱状图，可以明显发现辛烷值损失降幅超过30%，达到60.53%，只有少数样本最终损失降幅未能达标。其中，辛烷值损失降幅超过60%的达到了25.56%，接近样本的1/4，这是由于其原始辛烷值损失本身就比较高，导致其优化过程中，效果会比较好。





根据操作变量的分布图，可以发现部分变量呈正态分布，集中在均值分布。

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 样本编号 | 1 | 2 | 3 | 4 | 5 | 6 |
| 均值 | 18565612 | 0.111719 | 58.15046 | 148.8417 | 168656.2 | 21.17981 |
| 标准差 | 11538537 | 0.080178 | 21.27562 | 289.7407 | 539616.7 | 3.12508 |
| 最小值 | 80000 | 0 | 25 | 8 | -3664580 | 14.35065 |
| 最大值 | 39558757 | 0.25 | 110 | 1261.995 | 350000 | 31.96805 |
| 样本编号 | 7 | 8 | 9 | 10 | 11 | 12 |
| 均值 | 0.167761 | 22.98737 | 189.6031 | 2.170147 | 210488.1 | 25.60416 |
| 标准差 | 0.045396 | 9.366758 | 46.45738 | 1.971345 | 161316 | 3.103966 |
| 最小值 | 0.03822 | 0 | 1.582905 | -0.5 | 75.1712 | 14.47561 |
| 最大值 | 0.2 | 30 | 640.5102 | 5 | 335357.6 | 35.80647 |
| 样本编号 | 13 | 14 | 15 | 16 | 17 |  |
| 均值 | 2.919271 | 278.9364 | 144.2994 | -83100.6 | 14513.36 |  |
| 标准差 | 1.011305 | 8.655094 | 11.46113 | 98873.73 | 17754.71 |  |
| 最小值 | 0.209332 | 254.0271 | 122.5975 | -314601 | 0.15 |  |
| 最大值 | 3.5 | 312.4448 | 172.6794 | 75 | 54809.38 |  |

通过对操作变量的描述性分析，可以观察到各个操作变量优化后的取值范围，可以为未来操作变量的调整提供参考建议。

* 1. 结果结论

通过建立基于粒子群算法的组合优化模型，我们成功优化了266个样本的辛烷值损失降幅，并分析了经过优化后，操作变量的分布情况。