

# 基于粒子群算法的辛烷值损失减少过程的优化

丁佳铭

上海理工大学机械工程学院, 上海

收稿日期: 2023年2月21日; 录用日期: 2023年5月5日; 发布日期: 2023年5月12日

## 摘要

汽油是小型汽车的主要燃料, 但其燃烧产生的尾气排放对大气环境有着重要的影响, 将汽油清洁化已是迫在眉睫, 但同时也要保证其辛烷值的含量。本文运用神经网络的方法针对处理辛烷值(ON)损失预测模型的问题展开研究。对给定的原始采集数据中部分样本进行预处理, 通过拉依达准则( $3\sigma$ 准则)去除异常值, 采用最大最小的限幅方法剔除一部分不在此范围的样本; 然后对于残缺值大于20%的位点可以通过巴莱多定律进行删除, 为了选取对产品属性影响较大的操作变量与因素, 使用SPSS软件通过主成分分析法对原有的367个变量进行降维处理, 得到了具有典型性和独立性的22个主要变量, 通过BP神经网络建立预测模型。为了满足硫含量的条件, 利用粒子群算法对硫含量再次预测, 得出硫含量与主要变量的表达式, 能够得到使得辛烷值损失取得最小值的最优解。

## 关键词

神经网络, 主成分分析, 降维, 粒子群算法

# Optimization of Octane Number Loss Reduction Process Based on Particle Swarm Algorithm

Jiaming Ding

School of Mechanical Engineering, University of Shanghai for Science and Technology, Shanghai

Received: Feb. 21<sup>st</sup>, 2023; accepted: May 5<sup>th</sup>, 2023; published: May 12<sup>th</sup>, 2023

## Abstract

Gasoline is the main fuel for small cars, but the exhaust emissions from its combustion have an

important impact on the atmospheric environment. It is urgent to clean gasoline, but at the same time, it is necessary to ensure its octane number. In this paper, neural network is used to study the problem of dealing with the prediction model of octane number (ON) loss. Pre-process some samples in the given original collected data, and pass the Laida criterion ( $3\sigma$  Criterion) to remove the outliers, and use the maximum and minimum limiting method to remove some samples that are not in this range; Then, sites with incomplete values greater than 20% can be deleted through the Baredo's law. In order to select the operational variables and factors that have a greater impact on the product attributes, the original 367 variables are reduced by using the SPSS software through the principal component analysis method, and 22 main variables with typicality and independence are obtained. The prediction model is established through BP neural network. In order to meet the conditions of sulfur content, the particle swarm algorithm is used to predict the sulfur content again, and the expression of sulfur content and main variables is obtained, which can obtain the optimal solution to minimize the loss of octane number.

## Keywords

Neural Network, Principal Component Analysis, Dimension Reduction, Particle Swarm Algorithm

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

在当今这个时代，经济与需求都在快速发展，汽车已经成为社会生活和工作的必需品。尽管近几年电动汽车正在飞速发展，但是不可否认传统的内燃机汽车仍然是汽车行业的主流。这就导致给城市甚至世界环保带来很大压力。国家在不断加强对汽车行业的管控的同时对汽车尾气排放标准也在不断提高。影响汽车尾气排放因素有发动机、进气系统等，但其中一个相当重要的因素则是汽油。

汽油是当今世界上用量最大、用途最广的轻质石油产品之一，主要用于各类交通工具的动力燃料。汽油是通过重油分解出来的，大部分使用的是重油轻质化工技术。国家需要的成品汽油则需要再次进行精制处理。汽油标号代表的是汽油的抗爆性，在工程中则是指辛烷值。辛烷值越高，就代表汽油抗爆值越高。我国为了减少尾气中硫等有害气体的排放，大部分使用的是催化裂化技术来脱硫，但是会使汽油辛烷值的降低，从而导致经济效益大大减少。

命题宗旨是：利用数据挖掘技术[1]来建立降低辛烷值损失模型和辛烷值损失预测模型。所建立的模型区别于通过数据关联或机理建模的方法所建立的化工模型，充分利用到了其中大量的操作变量，实现对过程优化的及时响应。通过对主要操作变量的优化，实现减少硫含量的同时保证汽油辛烷值降低幅度最小，同时实现效益的增长。

国内研究者主要集中在辛烷值损失模型的建模方面[2]，研究了基于统计学、机器学习和深度学习等技术的辛烷值损失模型。近年来，随着中国石化企业实验室信息管理系统的建成，积累了大量的实验数据。因此，以现有数据库中的汽油理化指标数据集为基础，开始探索和采用一些非线性的建模方法，进行辛烷值的定量计算，如秦玉翠利用误差反向传播人工神经网络(back propagation neural network, BP神经网络)，对近红外光谱仪测定数据进行处理分析，如光谱吸光度与汽油辛烷值，结果证明了 ANN 方法的可行性。另外也有研究以向量机回归法为基础，以分子结构为依据，构建基于烷烃马达法的辛烷值计算模

型,最后采用留一法验证模型,验证了数据的鲁棒性。这些研究工作说明了采用非线性的建模方法有利于进一步精确对辛烷值进行定量计算,进而精确对辛烷值的损失定量计算,有利于提高汽油清洁化的效率。

## 2. 寻找建立辛烷值损失预测模型的主要变量

### 2.1. 数据预处理

给定的原始样本数据采集于霍尼韦尔 PHD (中石化高桥石化的实时数据库)以及 LIMS 实验数据库,原始数据中大部分位点的变量数据正常,但每套装置采集的数据均有相当部分点位存在问题[3],因为数据在采集和传输这个过程中会存在着数据误传、漏传等一些误差现象,因此需要对原始样本数据进行预处理后方可进行后续分析使用。由于辛烷值的测定数据相对于操作变量数据而言相对较少,而且辛烷值的测定往往滞后,所以以辛烷值数据测定的时间点为基准时间,取其前 2 个小时的操作变量数据来进行数据预处理。

#### 不良数据处理

- 1) 异常值处理
  - a) 根据拉依达准则( $3\sigma$  准则)去除异常值
  - b) 通过变量操作范围(最大最小值的限幅方法)剔除异常值
- 2) 部分残缺及全部空值处理
  - a) 根据 80%法则对缺失值进行删除处理
  - b) 部分数据为空值的位点用平均值进行插值

数据处理流程图见图 1:

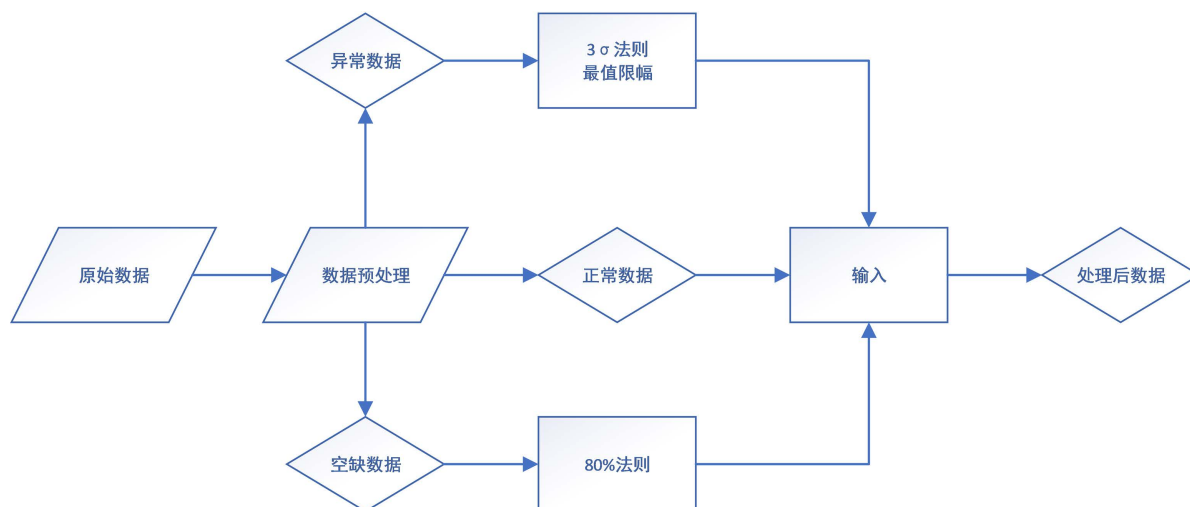


Figure 1. Data processing flow chart

图 1. 数据处理流程图

### 2.2. 寻找建模主要变量

#### 2.2.1. 问题分析

为了简化实际的生产问题。需要针对已经处理 325 个样本数据,将 367 个操作变量通过降维的操作筛选出建模主要变量,使其尽可能具有典型性,并拥有各自的独立性,方便应用于化工生产中。为了实现降低辛烷值损失的目标,就必须先通过数据降维,来减少不必要因素的影响。能实现数据降维的有以下方法:缺失比率、低方差滤波、高相关滤波、随机森林/组合树以及主成分分析法。本文使用主成分分析法(PCA)是一

种常用于减少大数据集维数的降维方法。利用 SPSS 软件，使用主成分分析法实现变量的降维处理[4]。

### 2.2.2. 主成分分析法(PCA)

主成分分析法(PCA) [5]能够把大变量集转换为仍包含大变量集中大部分信息的较小变量集将相关性较高的特征变为相互独立且不相关的特征。优点是在减少数据集的变量数量的同时保留尽可能多的信息。

### 2.2.3. 主成分分析求解过程

1) 计算标准化样本的协方差矩阵，是反映标准化后的数据之间相关关系密切程度的统计指标，值越大，说明有必要对数据进行主成分分析[6]。

$$R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1p} \\ r_{21} & r_{22} & \cdots & r_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ r_{p1} & r_{p2} & \cdots & r_{pp} \end{bmatrix} \quad (1)$$

其中  $r_{ij} = \frac{1}{n-1} \sum_{i=1}^n (X_{id} - \bar{X}_i)(X_{ij} - \bar{X}_j) = \frac{1}{n-1} \sum_{k=1}^n X_{ik}X_{kj}$ 。

2) 得出样本矩阵的样本关系系数矩阵

$$R = \frac{\sum_{k=1}^n (x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j)}{\sqrt{\sum_{i=1}^n (x_{ki} - \bar{x}_i)^2 \sum_{k=1}^n (x_{kj} - \bar{x}_j)^2}} \quad (2)$$

3) 检验

进行 KMO 检验及 Barlett 球形检验，其中 KMO [7]检验是取样的适当性检验，主要是检验所选择的变量是否包含导致相关的根本性变量；Barlett 球形检验是判读相关系数矩阵与单位阵是否有显著差异。如果相关系数矩阵近似为单位矩阵，则变量之间彼此正交，进行主成分分析没有意义，说明两者之间有显著的差异，可以开展主成分分析。

4) 根据协方差矩阵，计算  $R$  的特征值和特征向量

特征值：  $\lambda_1 \geq \lambda_2 \geq \cdots \lambda_p \geq 0$  ( $R$  是半正定矩阵，且  $\text{tr}(R) = \sum_{i=1}^p \lambda_k = p$ )

$$\text{特征向量： } a_1 = \begin{bmatrix} a_{11} \\ c_{21} \\ \vdots \\ a_{p1} \end{bmatrix}, a_2 = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{p2} \end{bmatrix}, \cdots, a_p = \begin{bmatrix} a_{1p} \\ a_{2p} \\ \vdots \\ a_{pp} \end{bmatrix}$$

5) 计算主成分贡献率以及累计贡献率，根据选取的主成分个数的原则，来确定主成分个数。

$$\text{贡献率} = \frac{\lambda_i}{\sum_{k=1}^p \lambda_k} (i=1, 2, \cdots, p) \quad \text{累计贡献率} = \frac{\sum_{k=1}^i \lambda_k}{\sum_{k=1}^p \lambda_k} (i=1, 2, \cdots, p)$$

一般取累计贡献率超过 80% 的特征值对应的第一、第二、……、第  $m$  ( $m \leq p$ ) 个主成分，第  $i$  个主成分： $F_i = a_{i1}X_1 + a_{i2}X_2 + \cdots + a_{ip}X_p$  ( $i=1, 2, \cdots, m$ )。

## 6) 计算主成分载荷

$$p(z_k, x) = \sqrt{\gamma} e_k (i, k = 1, 2, \dots, p) \quad (3)$$

## 7) 通过主成分分析, 可以得到主成分得分矩阵

$$z = \begin{bmatrix} z_{11} & z_{12} & \cdots & z_{1m} \\ z_{21} & z_{22} & \cdots & z_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ z_{n1} & z_{n2} & \cdots & z_{nm} \end{bmatrix}$$

### 2.3. PCA 编码实现与结果

变量筛选原则: 控制变量对产品性质具有明显的响应关系, 变量间相互独立、不存在重复信息; 能够全面反映成品汽油的不同特征属性。利用主成分因子方法(PCA)分别对各个操作变量进行分析, 根据提取主成分个数累计方差超过 80% 的原则, 通过最大方差旋转法, 经因子载荷矩阵旋转后选择载荷值大于 0.8 的指标作为下一步待筛选指标适用性分析, 利用 SPSS 软件对操作变量进行因子分析。

通过观察表 1 各个主成分的特征值, 可以看出特征值的变化规律, 特征值的大小代表了主成分的方差贡献率大小和重要性程度。累计贡献率是因子分析中抽取出的因子特征值之和的比值, 取出的因子相较总体变量更具有代表性。

**Table 1.** Interpretation of total variance

**表 1.** 总方差解释

成分	特征值	方差百分比	累计贡献率
1	112.391	32.02	32.02
2	38.439	10.951	42.971
3	23.501	6.696	49.667
4	19.442	5.539	55.206
⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮
20	2.82	0.803	83.283
21	2.474	0.705	83.988
22	2.403	0.685	84.283

根据表 1 和表 2 的总方差和成分矩阵, 就可以计算指标权重, 根据要求解释的总方差需要超过 80%, 同时为了避免在之后的权重计算中出现负值, 就需要依靠 SPSS 对数据进行标准化。

#### 指挥权重计算流程

Step 1: 输入数据主成分的特征根和成分。

Step 2: 计算线性组合中的系数。

Step 3: 计算综合得分模型中的系数。

Step 4: 权重计算。

Step 5: 对指标权重按高到低排列, 选出其中由高到低的 22 个操作变量作为建模的主要变量。如下表 3。

**Table 2.** Composition matrix  
**表 2.** 成分矩阵

	1	2	...	39	40
硫含量	0.004	0.002	...	-0.058	-0.068
辛烷值	0.002	-0.001	...	-0.002	-0.078
饱和烃	-0.005	-0.008	...	0.020	0.044
烯烃	0.005	0.004	...	-0.020	-0.035
⋮	⋮	⋮	...	⋮	⋮
压缩入口	-0.001	-0.004	...	-0.81	-0.033
压缩出口	-0.006	-0.015	...	0.008	0.005
冲罐压力	-0.008	-0.006	...	0.008	0.001

**Table 3.** Main variables of modeling  
**表 3.** 建模主要变量表

主要变量	指标权重
D101 原料缓冲罐压力	0.04345997
S-ZORB.FT_9102.PV	0.043084445
S-ZORB.FT_1204.TOTAL	0.042648932
D-123 凝结水入口流量	0.042518765
反应过滤器压差	0.041835247
K-103A 进气温度	0.041811038
K-101A 进气温度	0.041694848
K-103A 进气压力	0.041642478
K-101A 排气压力	0.041641942
K-101A 进气压力	0.041641895
EH101 出口	0.041641592
D-110 底压力	0.041638707
K-103A 排气压力	0.041638701
原料缓冲罐液位	0.041633427
S-ZORB.FT_1202.TOTAL	0.041633427
K-103A 排气温度	0.041553295
R-102 底喷头压差	0.041098735
R-101 底格栅上下压差	0.041095436
D-123 压力	0.040581745
低压热氮气压力	0.040559523
还原器流化氢气流量	0.039573682

### 3. 建立辛烷值损失预测模型

#### 3.1. 问题分析

需要建立汽油辛烷值(ROM)损失预测模型。采用上文得到的建模主要操作变量,利用数据挖掘技术来建立辛烷值(ROM)损失预测模型,并进行模型验证。根据所得建模主要操作变量的数据分析,需要使用BP神经网络来建立预测模型,利用Matlab软件编写程序,实现对辛烷值(ROM)损失的预测,通过模型验证,来分析误差。

#### 3.2. BP神经网络的基本原理

神经网络又称为连接模型,是一种模仿动物神经网络行为特征,进行分布式并行信息处理的算法模型。模仿生物神经网络大量神经节点之间相互连接的关系,实现处理信息的目的。

BP神经网络[8]具有不断学习的功能,就是在模拟的过程中,通过不断的循环来收集系统所产生的误差,并且通过反馈到输出层,再通过输出层传递信号,最重要的是可以使用这些误差来调整神经元的权重,这样就生成一个可以模拟出原始问题并且能够很好地解决人工神经网络系统。本文就是基于BP神经网络对辛烷值损失的预测来建立模型[9]。

#### 3.3. 模型建立

主要因素的确定:由上面的主成分分析分析计算可知,影响辛烷值损失(ROM)的主要因素是22个变量。BP神经网络的训练过程及流程图见图2:

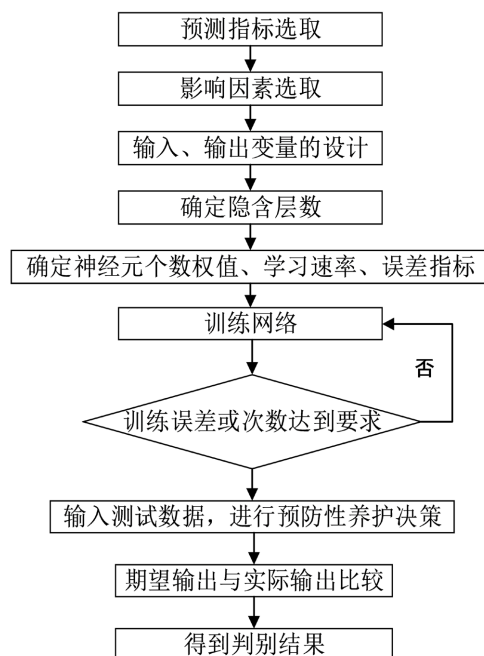


Figure 2. BP neural network flow chart  
图 2. BP神经网络流程图

- 1) 计算个层神经元的输入和输出;
- 2) 计算误差函数对输出层的个神经元的偏导数。
- 3) 计算误差函数对隐含层各神经元的偏导数。

- 4) 修正连接权值。
- 5) 计算全局误差。
- 6) 判断网络误差是否满足要求。

然后，我们选择辛烷值(ROM)损失相关数据为研究对象，利用 Matlab 软件进行预测分析，可以得到如图 3 所示结果。

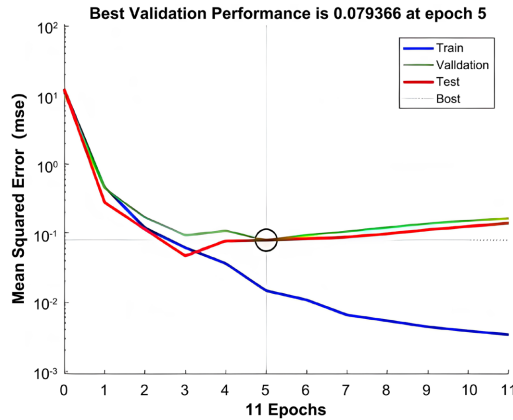


Figure 3. Relationship between gradient and mean square error of variables  
图 3. 变量的梯度和均方误差之间的关系图

一般来说，MSE 的值越小，R 值越接近 1，模型拟合的越好。通过分析图 4 可以知道，训练效果比较好。

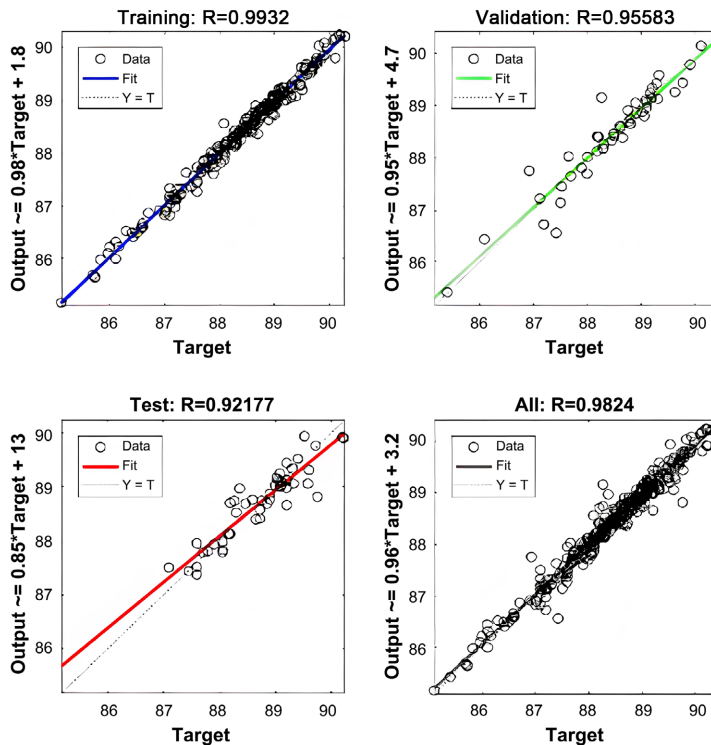


Figure 4. Final result diagram  
图 4. 最终结果图



将拟合值对真实值回归，拟合优度越高，说明拟合的效果越好。由上图拟合数据可知，使用的神经网络预测模型误差很小，说明该模型符合要求。

## 4. 优化主要变量操作方案

### 4.1. 问题分析

问题四要求在保证产品硫含量不大于 5  $\mu\text{g/g}$  的前提下，求得辛烷值( RON) 损失减少的各主要变量的优化条件。即求得使得辛烷值损失最少的，且满足硫含量不大于 5  $\mu\text{g/g}$  的各个主要变量的值。也就是求解最优解的问题：辛烷值损失值的最小为目标，22 个变量的取值范围以及硫的含量范围作为约束条件，求解最优条件。

其中辛烷值损失量的预测模型在上一章已经求出，此预测模型也即为辛烷值损失量与各个主要变量的关系式。而硫含量与筛选出的主要变量的变化相关，即随着主要变量的变化而变化，因此根据神经网络算法对硫含量再进行一次预测，可以得出硫含量与主要变量的表达式。

### 4.2. 解决方案

先用 BP 神经网络算法求出硫含量关于 22 个主变量的预测模型，然后选定辛烷值损失值的最小为目标，22 个变量的取值范围以及硫的含量范围作为约束条件，用粒子群算法对优化问题进行求解。

#### 4.2.1. 粒子群算法求解最优化问题

是一种基于群体智能理论的并行算法，利用群体中的个体对信息的共享使整个群体的运动在问题求解空间中产生从无序到有序的演化过程，从而获得最优解[10]。它是通过适应度来评价解的品质，但它比遗传算法规则更为简单，它没有遗传算法的“交叉”和“变异”操作，它通过追随当前搜索到的最有值来寻找全局最优[11]。

假设  $X_i = (x_{i,1}, x_{i,2}, \dots, x_{i,n})$  是微粒  $i$  的当前位置， $V_i = (v_{i,1}, v_{i,2}, \dots, v_{i,n})$  是微粒  $i$  的当前飞行速度，那么，基本粒子群优化算法[12]的进化方程如下：

$$v_{i,j}(t+1) = v_{i,j}(t) + c_1 \text{rand}_1() (p_{i,j}(t) - x_{i,j}(t)) + c_2 \text{rand}_2() (p_{g,j}(t) - x_{i,j}(t)) \quad (4)$$

$$x_{i,j}(t+1) = x_{i,j}(t) + v_{i,j}(t+1) \quad (5)$$

$$P_i(t) = (p_{i,1}(t), p_{i,2}(t), \dots, p_{i,n}(t)) \quad (6)$$

$$P_g(t) = (p_{g,1}(t), p_{g,2}(t), \dots, p_{g,n}(t)) \quad (7)$$

式中， $t$  表示迭代代数；

$P_i(t)$  表示微粒  $i$  迄今为止经历过的历史最好位置；

$P_g(t)$  是当前粒子群搜索到的最好位置(全局最好位置)；

$C_1$ 、 $C_2$  分别为认知学习因子和社会学系因子。

#### 4.2.2. PSO 的基本算法流程

Step 1: 在允许的范围内，初始化搜索点的位置及其速度；

Step 2: 计算粒子的适应度值，即目标函数值；

Step 3: 根据以上公式，对每一个粒子的速度和位置进行更新；

Step 4: 检验是否符合结束条件，若当前的迭代次数达到了预先设定的最大次数(或达到最小错误要求等)，则停止迭代，输出最优解，否则转到 Step 2。算法流程见图 5。

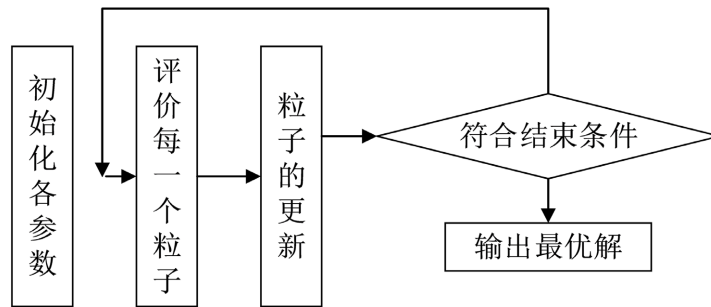


Figure 5. Particle swarm algorithm process  
图 5. 粒子群算法流程

### 4.3. 粒子群算法的初始值设定与结果展示

#### 4.3.1. 粒子群算法的初始参数设定

根据经验对粒子群算法初始值进行设定，如表 4 所示：

Table 4. Initial parameter list

表 4. 初始参数列表

设置参量名称	参量数值
群体粒子数	30
最大迭代次数	50
每个粒子的个体学习因子	2
每个粒子的社会学习因子	2
惯性权重	0.9
初始化粒子所在的位置	随机在定义域内

#### 4.3.2. 粒子群算法结果

经过粒子群算法计算后，可得到一组比较理想的使得辛烷值损失取得最小值的最优解，则优化结果见表 5。

Table 5. Optimization results

表 5. 优化结果

变量	取值
辛烷值	0.2
硫含量	3.2
D101 原料缓冲罐压力	0
S-ZORB.FT_9102.PV	0.021
D-123 凝结水入口流量	2.001
反应过滤器压差	0.011
K-103A 进气温度	0
K-101A 进气温度	0.0105

## Continued

K-103A 进气压力	0
K-101A 排气压力	0
K-101A 进气压力	0
EH101 出口	0
D-110 底压力	0.0011
K-103A 排气压力	0.001
原料缓冲罐液位	0.041
S-ZORB.FT_1202.TOTAL	0
K-103A 排气温度	0
R-102 底喷头压差	0
R-101 底格栅上下压差	0.006
D-123 压力	0.0003
低压热氮气压力	0.0003
还原器流化氢气流量	0.601
EH-102 出口管	0.066
热氮气过滤器 ME-113 差压	0

根据最优化结果可知，硫含量满足要求，且辛烷值损失值的结果为 0.2 个单位，远小于该厂辛烷值损失值平均值 1.37 的 30%，符合优化要求。

## 5. 模型可视化

要求以图形展示其主要操作变量优化调整过程中对应的汽油辛烷值和硫含量的变化轨迹。而由于主要操作变量只能逐步调到位，故需要使一个操作变量进行单独的有步长的变化，其他操作变量不变，展现汽油辛烷值和硫含量的变化情况。

### 5.1. 变量筛选

部分变量的可调整幅度值大于操作变量最优值与 133 号样本的差值，表示这些变量已经在进行人工操作后，也无法降低辛烷值损失量，故将其剔除。

如表 6 所示，操作变量 D101 原料缓冲罐压力不符合上述条件，进行剔除，不予考虑。

**Table 6.** Filter main variables

**表 6.** 筛选主要变量表

主要变量	D101 原料缓冲罐压力	D-123 凝结水入口流量	……	反应过滤器压差
133 样本值	0.3506	1990.0110	……	22.6288
最优解值	0	2.001	……	0.011
样本值与最优解的差值	0.3506	1990.0110	……	22.6288
操作变量每次允许调整幅度值 $\Delta$	0.5	200	……	1

## 5.2. 可视化分析

绘制图，以某操作变量的值为横坐标，以辛烷值与硫含量的值为纵坐标。操作变量以相应可调整幅度值进行递增或递减，解出的最优值 133 号样本的值进行逼近。辛烷值与硫含量的值由上文的预测模型和操作变量值进行确定。在筛选出的主要变量中，以 D-123 凝结水入口流量作为典型例子，进行绘图分析。

由图 6 可见，D-123 凝结水入口流量的值从 2000 个单位到 0 时，也即从 133 样本的值到最优解时，辛烷值损失值逐渐减小，且硫的含量满足要求。

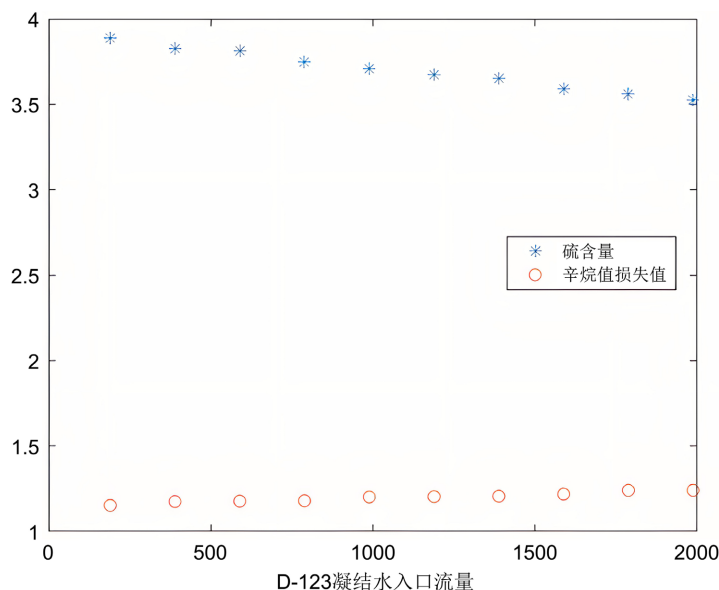


Figure 6. Relation among sulfur content, octane number loss value and D-123 condensate inlet flow  
图 6. 硫含量、辛烷值损失值关于 D-123 凝结水入口流量关系图

## 6. 结论

本文首先基于对辛烷值损失影响的大小，利用主成分分析筛选主要影响变量，得到 22 个主要变量；然后通过得到的主要操作变量利用 BP 神经网络建立辛烷值损失预测模型，模型验证后误差很小，说明构建了可靠的计算模型；为了优化变量，满足辛烷值和硫含量的要求，使用高精度的粒子群算法，使实现的损失值远小于优化目标。

## 参考文献

- [1] 方洪鹰. 数据挖掘中数据预处理的方法研究[D]: [硕士学位论文]. 重庆: 西南大学, 2009.
- [2] 吴新玲, 毋国庆. 基于数据变换的维数消减方法[J]. 武汉大学学报(理学版), 2006(1): 73-76.
- [3] 徐林明, 李美娟. 动态综合评价中的数据预处理方法研究[J]. 中国管理科学, 2020, 28(1): 162-169.
- [4] 童其慧. 主成分分析方法在指标综合评价中的应用[J]. 北京理工大学学报(社会科学版), 2002(1): 59-61.
- [5] 陈佩. 主成分分析法研究及其在特征提取中的应用[D]: [硕士学位论文]. 西安: 陕西师范大学, 2014.
- [6] 李艳双, 曾珍香, 张闯, 于树江. 主成分分析法在多指标综合评价方法中的应用[J]. 河北工业大学学报, 1999(1): 3-5.
- [7] Parhizkar, T., Rafeipour, E. and Parhizkar, A. (2021) Evaluation and Improvement of Energy Consumption Prediction Models Using Principal Component Analysis Based Feature Reduction. *Journal of Cleaner Production*, **279**, Article

---

ID: 123866. <https://doi.org/10.1016/j.jclepro.2020.123866>

- [8] 李萍, 曾令可, 税安泽, 金雪莉, 刘艳春, 王慧. 基于 MATLAB 的 BP 神经网络预测系统的设计[J]. 计算机应用与软件, 2008(4): 149-150+184.
- [9] 柳小桐. BP 神经网络输入层数据归一化研究[J]. 机械工程与自动化, 2010(3): 122-123+126.
- [10] 郑秋梅, 商振浩, 王风华, 等. 基于深度神经网络和支持向量机的海底管线水合物生成预测模型[J]. 中国石油大学学报(自然科学版), 2020, 44(5): 46-51.
- [11] 王道明, 鲁昌华, 蒋薇薇, 肖明霞, 李必然. 基于粒子群算法的决策树 SVM 多分类方法研究[J]. 电子测量与仪器学报, 2015, 29(4): 611-615.
- [12] 居凤霞. 粒子群优化算法的改进及应用[D]: [硕士学位论文]. 广州: 华南理工大学, 2014.