联合站工艺基础参数

联合站工艺流程介绍

该联合站是一座集油气分离、原油脱水、原油稳定、原油外输多功能的联合站。站内拥有原油外输泵2台，稳定泵2台，压缩机2台，加热炉5台，油罐（5000m3）4座，重油罐（2000m3）2座，三相分离器3台，稳定塔2座，轻油罐2座。联合站设计原油处理能力为200×104 t/a，原油稳定能力200×104 t/a，目前实际原油处理及稳定能力120×104t/a。工艺流程如图 4-2 所示为：井排来油→三相分离器→加热炉→稳定塔→一次沉降罐→二次沉降罐→净化油罐→外输[73]。

三相分离器

三相分离器

井排来液

三相分离器

水套炉

外输泵

水套炉

一次沉降罐

二次沉降罐

净化油罐

稳定塔

**图4-2 联合站工艺流程**

**Fig4-2 The process of the centralized processing station**

联合站数据来源

该联合站有来自五个采油队的井排来液以及某联合站的来油，进站后由站内安置的流量计、压力、温度测量仪表等分别对井排来液及某联合站来油进行实时数据计量；外输油品出站前也会进行计量，记录当日外输液量及含水率、温度等实时数据。针对联合站的三相分离器、加热炉、稳定塔和外输泵都装有相关的流量计、温度计和压力表，数字化油田的实现为这些参数的计量提供了便利的条件。

本文采集到的联合站内生产参数有：1#三相分离器、2#三相分离器、3#三相分离器的处理液量、处理气量、油水界面、工作压力、工作温度，1#加热炉和2#加热炉的进口温度、出口温度、进口压力、出口压力、燃油用量、燃气用量、通过油量、通过液量、运行效率，稳定塔的压力和进油量以及外输泵的泵压力、干压、排量、工作电流、工作电压、运行频率、进口压力以及运行效率。

联合站能耗指标计算

根据采集的油品化验数据和实时监测数据，计算出**联合站单位燃料气（油）消耗**、**联合站单位原油电能消耗**以及**联合站热能利用率**，具体的联合站能耗指标计算公式[74]如下所示：

集油系统的边界应界定为从井口出油到计量间、转油站、最终进入集中处理站（联合站、脱水站）之前的全部流程。涉及井口、计量间、转油站及所有连接管线，加热炉、泵等主要耗能设备。

**表4-1 集油系统评价指标汇总表**

**Table4-1 Summary of evaluation index of oil gathering system**

| 系统 | 评价单元 | | 评价指标 | 单位 | 备注 |
| --- | --- | --- | --- | --- | --- |
| 集油系统 | 综合 | | 单位液量集油气耗 | m3/t |  |
| 单位液量集油电耗 | kW▪h/t |  |
| 单位液量集油综合能耗 | kgce/t |  |
| 站场 | | 热能利用率 | % |  |
| 电能利用率 | % |  |
| 能量利用率 | % |  |
| 输油单耗 | kW▪h/m3 |  |
| 设备 | 加热炉 | 加热炉负荷率 | % |  |
| 加热炉热效率 | % | 正平衡计算 |
| 泵 | 泵负荷率 | % | 按排量计算 |
| 泵机组效率 | % |  |
| 管线 | | 集油管线百米压降 | MPa/100m |  |
| 集油管线百米温降 | ℃/100m |  |
| 集油管线能量损失率 | % |  |
| 系统 | | 输油系统效率 | % |  |

1.综合指标

（1）单位液量集油气耗

单位液量集油气耗：集油系统中的总耗气量与测试期内集油系统中所有单井产液量之和的比值。

计算边界：集中处理站前包括所有的单井、计量间、转油站及管线。



（2）单位液量集油电耗

单位液量集油电耗：集油系统中的总耗电量与测试期内集油系统中所有单井产液量之和的比值。

计算边界：集中处理站前包括所有的单井、计量间、转油站及管线。



注：对于双管掺水流程，应包括转油站掺水泵耗电。不应包括热洗泵耗电。

（3）单位液量集油综合能耗

单位液量集油综合能耗：集油系统中总能耗与测试期内集油系统包含所有单井的产液量之和的比值。

计算边界：集中处理站前包括所有的单井、计量间、转油站及管线。

 式中：标煤的单位为kgce，测试期内集油系统所含所有单井产液量之和，单位为t。

2.站场

（1）热能利用率

计算边界：以某一个站场为单位，出站流体热能与进站流体热能的差值，比上站内供给的热能。

 （4-1）

其中，





式中，——热能利用率，%；

——介质带入、带出的热能，kJ/h；

——站消耗总热能，kJ/h；

，——物流数；

——进站第种介质的比热容，kJ/（kg·℃）；

——出站第种介质的比热容，kJ/（kg·℃）；

——进站第种介质流量，kg/h；

——出站第种介质流量，kg/h；

——出站第种介质的温度，℃；

——进站第种介质的温度，℃；

——参考温度，℃；

——燃料消耗量，Nm³（kg）/h；

——燃料热值，kJ/ kg（m³）。

（2）电能利用率

计算边界以某一个站场为单位，出站流体压能与进站流体压能的差值，比上站内供给的电能。

 （4-2）

式中，——电能利用率，%；

——进站、出站的压能，MJ/h；

——站总耗电量，(kW·h)/h。

（3）能量利用率

计算边界：以某一站为计算单位。

 （4-3）

式中，——站的能源效率，%；

——介质进出站的热能，kJ/h；

——介质进出站的压力能，kJ/h；

——站供给介质的热能，kJ/h；

——站耗电量，(kW·h)/h。

能量利用率也可表示为进出口物流焓差。热能与电能不能直接相加，把电能乘以折合系数折合为热能。

3.设备

加热炉（主要评价加热炉热效率、散热损失、排烟损失、不完全燃烧损失、空气系数、泵（泵效率、泵机组效率、节流损失率《GB/T 31453-2015 油田生产系统节能监测规范》）。

（1）加热炉效率

加热炉效率需要通过测试及计算确定。计算对象为加热炉单体，指标测算包括正平衡、反平衡两种测试方法。

正平衡计算法：

加热炉有效利用的能量与总的对加热炉的能量供给的比值为加热炉的热效率：

 （4-4）

其中，；；为原油输量，kg /h；为原油比热容，kJ /(kg·C);*TO*为原油出炉温度，℃；为原油进站温度，℃；为燃气消耗量，m3/h；为燃气低发热值，kJ /m3。

（2）泵效率

泵效率的测算对象为单泵，包含泵效率指标和泵机组效率指标，同时，离心泵与柱塞泵的计算方法也有差别。

1）离心泵评价方法

A、泵效率

泵的运行评价按照实际泵效来评价，具体计算公式为：

 （4-5）

其中， 为泵效，%；为有效功率，kW；**为电机功率，kW；为传动效率。

 （4-6）

 （4-7）

其中，U为电机电压，V；I为电机电流，A；为电机功率因子。

B、泵机组效率

 （4-8）

其中， 为泵机组的有效功率，kW，；为即泵机组的输入功率，kW： 

1）柱塞泵评价方法

A、泵机组效率

 （4-9）

其中，为泵机组效率，%；为泵出口压力，MPa；为泵入口压力，MPa；Q 为流量，m3/h。为即泵机组的有功功率，kW。

可通过下式计算得出：

 （4-10）

其中，*U*为平均电压，V；*I*为平均电流，A；为电机功率因数。

B、泵机组单耗

 （4-11）

其中，泵机组单耗，kW·h/m3；*Q*为流量，m3/h；为泵机组的有功功率，kW。

4.管线

管线流速、单位管长压力损失、管线百米温降、管道能损率、管道单位能损、管线效率、管道效率（模型校核得出）。

（1）集油管线百米压降

计算边界：以每根管线为计算对象。

**** （4-12）

式中，为单位管长的压降损失， Pa/m；为单管出口压力，Pa；单管入口压力，Pa；单管长度，m。

（2）集油管线百米温降

计算边界：以每根管线为单位，计算起点到终点每百米的温降。

 （4-13）

式中，为每百米单位管长温降，℃；为管线出口温度，℃；为管线入口温度，℃；为管线长度，m。

（3）集油管线能量损失率

计算边界：以每根管道为单位，散失的热能、压能之和与管道入口的热能、压能之和的比值。

 （4-14）

式中：为管道散失的热能，kJ/h，为管道散失的压能，kJ/h，为管道入口热能，kJ/h，为管道入口压能，kJ/h。

通过以上公式，我们可以计算出联合站单位燃料消耗、联合站单位原油电能消耗以及联合站热能利用率，补充原始数据中缺失的部分，由此形成此次预测模型的建模数据。

4.3 联合站单位燃料消耗预测

4.3.1 Adaptive-Lasso变量选择模型

运用LARS算法来解决Adaptive-Lasso估计，对于每给一个*γ*，该算法会寻求一个最优的*λn*。用Python编制相应的程序，代码清单见附表5。

运行结果如下表4-1所示。

**表4-1 系数表**

**Table4-1 Coefficient table**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| x1 | x2 | x3 | x4 | x5 | x6 | x7 | x8 | x9 | x10 | x11 |
| 0 | 0 | 0 | 0 | 0 | 0.0001 | -0.0002 | 0 | 0 | 0 | 0 |
| x12 | x13 | x14 | x15 | x16 | x17 | x18 | x19 | x20 | x21 | x22 |
| 0.002 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.001 | 0 |
| x23 | x24 | x25 | x26 | x27 | x28 | x29 | x30 | x31 | x32 | x33 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0011 | 0 | 0 |
| x34 | x35 | x36 | x37 | x38 | x39 | x40 | x41 | x42 | x43 | x44 |
| 0 | 0.0006 | 0 | 0 | -0.0017 | 0.0009 | 0.002 | 0 | -0.0003 | 0 | 0 |

由表4-1可以看出，利用Adaptive-Lasso方法识别影响单位燃料消耗的关键影响因素是2#三相分离器的**处理液量**和**处理气量**、3#三相分离器的**处理气量**、加热炉的**燃油用量**、**燃气用量**和**通过油量**、**稳定塔的进油量**、**外输泵的排量、工作电流、工作压力和进口压力。**

4.3.2 灰色预测

利用第二章的灰色预测算法来编写灰色预测函数，具体步骤如下：

（1）利用cunsum函数对原始序列进行累加；

（2）紧邻均值生成序列；

（3）将序列变为数组；

（4）建立灰微分方程；

（5）求解参数；

（6）计算方差比和小残差概率。

代码清单如附表6所示。

载入编写的灰色预测函数，2#三相分离器的处理液量（x6）和处理气量（x7）、3#三相分离器的处理气量（x12）、加热炉的燃油用量（x21）和通过油量（x31）、稳定塔的进油量（x35）、外输泵的排量（x38）、工作电流（x39）、工作压力（x40）和进口压力（x42）这些生产参数的第1576、1577数值通过Python建立灰色预测模型得出。处理过程为：载入编写的灰色预测函数，输入并读取生产数据，利用灰色预测函数预测第1576、1577数值。单位燃料消耗灰色预测的数据处理得代码清单如附表7所示。

预测结果的精度等级见表4-2。

**表4-2 灰色预测模型单位燃料消耗相关因素精度表**

**Table4-2 The precision tables of grey forecasting model of unit energy consumption factors**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | x6 | x7 | x12 | x21 | x31 |
| 1576预测值 | 4149.21 | 5.89 | 3.4 | 225.93 | 276.98 |
| 预测精度等级 | 好 | 一般 | 好 | 好 | 好 |
|  | x35 | x38 | x39 | x40 | x42 |
| 1576预测值 | 1901.04 | 236.73 | 105.53 | 406.23 | 212.68 |
| 预测精度等级 | 一般 | 好 | 好 | 好 | 一般 |

4.3.3 神经网络预测

利用第二章的神经网络算法来预测联合站的单位燃料消耗，其参数设置为误差精度10-7，学习次数10000次，神经元个数为Lasso变量选择方法选择的变量个数10。将数据零均值标准化后，带入联合站单位燃料消耗预测所建立的3层神经网络预测模型（输入层10节点，隐藏层20节点，输出层1节点），得到联合站单位燃料消耗1576和1577的预测值。具体处理过程为：读取数据，建立模型，用relu函数作为激活函数，大幅度提高准确度，编译模型，训练模型，保存模型参数，预测并还原结果。代码清单如附表8所示。

运行代码后，相关的数据结果如表 4-3所示。

**表4-3 神经网络模型预测结果及相对误差**

**Table4-3 Neural network model prediction results and the relative error**

|  |  |  |  |
| --- | --- | --- | --- |
|  | y | y\_pred | 相对误差（%） |
| 1 | 16.587 | 17.741 | 6.95 |
| 2 | 19.328 | 17.923 | 7.26 |

**表4-3 神经网络模型预测结果及相对误差（续）**

**Table4-3 Neural network model prediction results and the relative error（continued）**

|  |  |  |  |
| --- | --- | --- | --- |
|  | y | y\_pred | 相对误差（%） |
| 3 | 18.486 | 18.020 | 2.52 |
| 4 | 19.09 | 17.783 | 6.84 |
| 5 | 14.166 | 15.320 | 8.14 |
| 6 | 17.73 | 18.129 | 2.25 |
| 7 | 15.304 | 15.569 | 1.73 |
| 8 | 15.298 | 15.281 | 0.11 |
| 9 | 16.728 | 17.500 | 4.61 |
| …… | …… | …… |  |
| 1573 | 16.878 | 15.338 | 9.12 |
| 1574 | 17.414 | 16.640 | 4.44 |
| 1575 | 15.784 | 15.799 | 0.09 |
| 1576 |  | 16.230 |  |
| 1577 |  | 16.227 |  |

根据表中数据，作出单位原油电耗的真实值和预测值的对比图，如图4-3所示。



**图4-3 预测值与真实值对比图**

**Fig4-3 The comparison of predicted values and real values**

由表4-3可知，预测值与真实值的相对误差均小于9.2%，具有较好的可信度。由图4-3可知，预测曲线很好的温和了实际值曲线的变化趋势，两条曲线基本重合。第1576和1577个单位原油耗电量的预测值分别为16.230和16.227，可以起到能耗预警的作用，对数字化油田的发展具有建设性的意义。随着油田不断进入数字化时代，利用系统积累的海量数据进行分析、挖掘出影响联合站能耗的关键因素，利用往年的数据实现对未来能耗的预测。

4.4 联合站单位原油电能消耗预测

4.4.1 Adaptive-Lasso变量选择模型

利用Adaptive-Lasso方法识别影响单位原油电能消耗的关键影响因素，运行结果如表4-4所示。

**表4-4 系数表**

**Table4-4 Coefficient table**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| x1 | x2 | x3 | x4 | x5 | x6 | x7 | x8 | x9 | x10 | x11 |
| -0.0001 | 0 | 0 | 0 | 0 | -0.0001 | 0 | 0 | 0 | 0 | 0.0009 |
| x12 | x13 | x14 | x15 | x16 | x17 | x18 | x19 | x20 | x21 | x22 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| x23 | x24 | x25 | x26 | x27 | x28 | x29 | x30 | x31 | x32 | x33 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| x34 | x35 | x36 | x37 | x38 | x39 | x40 | x41 | x42 | x43 | x44 |
| -0.0002 | 0 | -0.0002 | -0.0002 | 0 | 0.0003 | 0 | 0 | 0 | 0 | 0 |

由表4-4可以看出，单位原油电能消耗的关键影响因素是1#三相分离器的处理液量、2#三相分离器的处理液量、3#三相分离器的处理液量、稳定塔的塔顶压力、外输泵的泵压力、干压、工作电流。

4.4.2 灰色预测

载入编写的灰色预测函数，1#三相分离器的处理液量（x1）、2#三相分离器的处理液量（x6）、3#三相分离器的处理液量（x11）、稳定塔的塔顶压力（x34）、外输泵的泵压力（x36）、干压（x37）、工作电流（x39）这些生产参数的第1576、1577数值通过Python建立灰色预测模型得出。

预测结果的精度等级见表4-5。

**表4-5 灰色预测模型单位燃料消耗相关因素精度表**

**Table4-5 The precision tables of grey forecasting model of unit energy consumption factors**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | x1 | x6 | x11 | x34 | x36 | x37 | x39 |
| 1576预测值 | 6515.89 | 4149.21 | 3374.36 | -0.04 | 0.85 | 0.78 | 105.53 |
| 1577预测值 | 6520.75 | 4148.77 | 3372.74 | -0.04 | 0.85 | 0.78 | 105.54 |
| 预测精度等级 | 好 | 好 | 好 | 好 | 一般 | 好 | 一般 |

4.4.3 神经网络预测

利用第二章的神经网络算法来预测联合站的单位原油电能消耗，其参数设置为误差精度10-7，学习次数10000次，神经元个数为Lasso变量选择方法选择的变量个数7。将数据零均值标准化后，带入联合站吨液能耗预测所建立的3层神经网络预测模型（输入层7节点，隐藏层14节点，输出层1节点），得到联合站吨液能耗1576和1577的预测值。

运行代码后，相关的数据结果如表4-6所示。

**表4-6 神经网络模型预测结果及相对误差**

**Table4-6 Neural network model prediction results and the relative error**

|  |  |  |  |
| --- | --- | --- | --- |
|  | y | y\_pred | 相对误差（%） |
| 1 | 9.184 | 8.774 | 4.46 |
| 2 | 8.533 | 8.825 | 3.42 |
| 3 | 9.108 | 8.988 | 1.31 |
| 4 | 9.609 | 9.169 | 4.57 |
| 5 | 8.804 | 8.645 | 1.80 |
| 6 | 8.948 | 8.866 | 0.91 |
| 7 | 7.458 | 7.880 | 5.65 |
| 8 | 7.899 | 8.271 | 4.70 |
| 9 | 8.676 | 8.860 | 2.12 |
| …… | …… | …… |  |
| 1573 | 7.328 | 7.652 | 4.42 |
| 1574 | 7.467 | 7.634 | 2.23 |
| 1575 | 6.782 | 7.748 | 4.36 |
| 1576 |  | 7.254 |  |
| 1577 |  | 7.252 |  |

根据表中数据，作出单位原油电耗的真实值和预测值的对比图，如图4-4所示。



**图4-4 预测值与真实值对比图**

**Fig4-4 The comparison of predicted values and real values**

由表4-6可知，预测值与真实值的相对误差均小于5.7%，具有较好的可信度。由图4-4可知，预测曲线很好的温和了实际值曲线的变化趋势，两条曲线基本重合。第1576和1577个单位原油耗电量的预测值分别为7.254和7.252，可以起到能耗预警的作用，对数字化油田的发展具有建设性的意义。随着油田不断进入数字化时代，利用系统积累的海量数据进行分析、挖掘出影响联合站能耗的关键因素，利用往年的数据实现对未来能耗的预测。

4.5 联合站热能利用率预测

4.5.1 Adaptive-Lasso变量选择模型

利用Adaptive-Lasso方法识别影响联合站热能利用率的关键影响因素，运行结果如表4-7所示。

**表4-7 系数表**

**Table4-7 Coefficient table**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| x1 | x2 | x3 | x4 | x5 | x6 | x7 | x8 | x9 | x10 | x11 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| x12 | x13 | x14 | x15 | x16 | x17 | x18 | x19 | x20 | x21 | x22 |
| -0.0006 | 0 | 0 | 0 | 0 | -0.0007 | 0 | 0 | 0 | 0 | 0 |
| x23 | x24 | x25 | x26 | x27 | x28 | x29 | x30 | x31 | x32 | x33 |
| 0 | 0 | 0 | -0.0007 | 0 | 0 | -0.0003 | 0.0031 | 0 | 0 | 0 |
| x34 | x35 | x36 | x37 | x38 | x39 | x40 | x41 | x42 | x43 | x44 |
| 0 | 0.0006 | 0 | 0 | -0.0017 | 0 | 0 | 0 | 0 | 0 | 0 |

由表4-7可以看出，单位原油电能消耗的关键影响因素是3#三相分离器的处理气量、1#加热炉和2#加热炉的出口温度，2#加热炉的燃油用量和燃气用量、稳定塔的进油量、外输泵的排量。

4.5.2 灰色预测

载入编写的灰色预测函数，3#三相分离器的处理气量（x12）、1#加热炉出口温度（x17）和2#加热炉的出口温度（x26）、燃油用量（x29）和燃气用量（x30）、稳定塔的进油量（x35）、外输泵的排量（x38）。这些生产参数的第1576、1577数值通过Python建立灰色预测模型得出。

预测结果的精度等级见表4-8。

**表4-8 灰色预测模型单位燃料消耗相关因素精度表**

**Table4-8 The precision tables of grey forecasting model of unit energy consumption factors**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | x12 | x17 | x26 | x29 | x30 | x35 | x38 |
| 1576预测值 | 6515.89 | 4149.21 | 3374.36 | -0.04 | 0.85 | 0.78 | 105.53 |
| 1577预测值 | 6520.75 | 4148.77 | 3372.74 | -0.04 | 0.85 | 0.78 | 105.54 |
| 预测精度等级 | 好 | 好 | 好 | 一般 | 好 | 好 | 一般 |

4.5.3 神经网络预测

利用第二章的神经网络算法来预测联合站的单位原油电能消耗，其参数设置为误差精度10-7，学习次数10000次，神经元个数为Lasso变量选择方法选择的变量个数7。将数据零均值标准化后，带入联合站吨液能耗预测所建立的3层神经网络预测模型（输入层7节点，隐藏层14节点，输出层1节点），得到联合站吨液能耗1576和1577的预测值。

运行代码后，相关的数据结果如表4-9所示。

**表4-9 神经网络模型预测结果及相对误差**

**Table4-9 Neural network model prediction results and the relative error**

|  |  |  |  |
| --- | --- | --- | --- |
|  | y | y\_pred | 相对误差（%） |
| 1 | 83.38 | 83.306 | 0.56 |
| 2 | 80.1 | 80.701 | 0.75 |
| 3 | 81.25 | 81.080 | 1.03 |
| 4 | 89.56 | 89.595 | 1.16 |
| 5 | 81.07 | 81.475 | 0.50 |
| 6 | 89.23 | 89.196 | 1.09 |
| 7 | 87.9 | 88.129 | 0.86 |
| 8 | 81.452 | 81.474 | 1.27 |
| 9 | 79.809 | 80.346 | 1.95 |
| …… | …… | …… |  |
| 1573 | 78.215 | 80.801 | 3.30 |
| 1574 | 81.812 | 80.907 | 1.10 |
| 1575 | 81.886 | 80.911 | 1.18 |
| 1576 |  | 82.395 |  |
| 1577 |  | 82.394 |  |

根据表中数据，作出单位原油电耗的真实值和预测值的对比图，如图4-5所示。



**图4-5 预测值与真实值对比图**

**Fig4-5 The comparison of predicted values and real values**

由表4-9可知，预测值与真实值的相对误差均小于3.3%，具有较好的可信度。由图4-5可知，预测曲线很好的温和了实际值曲线的变化趋势，两条曲线基本重合。第1576和1577个热能利用率的预测值分别为82.395%和82.394%，可以起到能耗预警的作用，对数字化油田的发展具有建设性的意义。随着油田不断进入数字化时代，利用系统积累的海量数据进行分析、挖掘出影响联合站能耗的关键因素，利用往年的数据实现对未来能耗的预测。

4.6 本章小结

本章利用第二章给出的灰色预测算法和神经网络算法对联合站单位燃料气（油）消耗、单位原油电能消耗、热能利用率进行了相关预测，预测结果相对真实可靠，能够起到能耗预警的作用，对数字化油田的发展具有建设性的意义。

**附表5 Adaptive-Lasso变量选择**

**Appendix5 Adaptive Lasso variable selection**

|  |
| --- |
| #-\*- coding: utf-8 -\*-  import pandas as pd  import numpy as np  inputfile = '../data/energy\_consumption.xlsx' #输入的数据文件  data = pd.read\_excel(inputfile) #读取数据  outputfile='../tmp/data2.xls'  #导入AdaptiveLasso算法  from sklearn.linear\_model import Lasso  model=Lasso()  model.fit(data.iloc[:,0:44],data['y'])  q=model.coef\_#各特征的系数  q=pd.DataFrame(q,index=['x1','x2','x3','x4','x5','x6','x7','x8','x9','x10','x11','x12','x13','x14','x15','x16','x17','x18','x19','x20','x21','x22','x23','x24','x25','x26','x27','x28','x29','x30','x31','x32','x33','x34','x35','x36','x37','x38','x39','x40','x41','x42','x43','x44',]).T  np.round(q,4).to\_excel(outputfile) |

**附表6 灰色预测函数代码**

**Appendix6 Grey prediction function code**

|  |
| --- |
| #-\*- coding: utf-8 -\*-  def GM11(x0): #自定义灰色预测函数  import numpy as np  x1 = x0.cumsum() #1-AGO序列  z1 = (x1[:len(x1)-1] + x1[1:])/2.0 #紧邻均值（MEAN）生成序列  z1 = z1.reshape((len(z1),1))  B = np.append(-z1, np.ones\_like(z1), axis = 1)  Yn = x0[1:].reshape((len(x0)-1, 1))  [[a],[b]] = np.dot(np.dot(np.linalg.inv(np.dot(B.T, B)), B.T), Yn) #计算参数  f = lambda k: (x0[0]-b/a)\*np.exp(-a\*(k-1))-(x0[0]-b/a)\*np.exp(-a\*(k-2)) #还原值  delta = np.abs(x0 - np.array([f(i) for i in range(1,len(x0)+1)]))  C = delta.std()/x0.std()  P = 1.0\*(np.abs(delta - delta.mean()) < 0.6745\*x0.std()).sum()/len(x0)  return f, a, b, x0[0], C, P #返回灰色预测函数、a、b、首项、方差比、小残差概率 |

**附表7 灰色预测代码**

**Appendix7 Grey prediction code**

|  |
| --- |
| #-\*- coding: utf-8 -\*-  import numpy as np  import pandas as pd  from GM11 import GM11  inputfile = '../data/energy\_consumption.xlsx'  outputfile = '../tmp/data1\_GM11energy\_consumption.xls'  modelfile = '../tmp/net.model'  data = pd.read\_excel(inputfile)  data.index = range(1,1576)  data.loc[1576] = None  data.loc[1577] = None  l = ['x6', 'x7', 'x12', 'x21', 'x31', 'x35','x38', 'x39', 'x40','x42']  for i in l:  f = GM11(data[i][range(1,1576)].as\_matrix())[0]  data[i][1576] = f(len(data)-1)  data[i][1577] = f(len(data))  data[i] = data[i].round(2)  data[l+['y']].to\_excel(outputfile) |

**附表8 神经网络预测代码**

**Appendix8 Neural network prediction code**

|  |
| --- |
| #-\*- coding: utf-8 -\*-  import pandas as pd  import numpy as np  inputfile = '../tmp/data1\_GM11energy\_consumption.xls' #灰色预测后保存的路径  outputfile = '../data/revenuepressure.xls' #神经网络预测后保存的结果  modelfile = '../tmp/1-net.model' #模型保存路径  data = pd.read\_excel(inputfile) #读取数据  feature = ['x6', 'x7', 'x12', 'x21', 'x31','x35', 'x38','x39','x40','x42'] #特征所在列  data\_train = data.loc[range(1,1576)].copy() #取2014年前的数据建模  data\_mean = data\_train.mean()  data\_std = data\_train.std()  data\_train = (data\_train - data\_mean)/data\_std #数据标准化  x\_train = data\_train[feature].as\_matrix() #特征数据  y\_train = data\_train['y'].as\_matrix() #标签数据  from keras.models import Sequential  from keras.layers.core import Dense, Activation  model = Sequential() #建立模型  model.add(Dense(input\_dim=10,output\_dim=20))  model.add(Activation('relu')) #用relu函数作为激活函数，能够大幅提供准确度  model.add(Dense(input\_dim=20,output\_dim=1))  model.compile(loss='mean\_squared\_error', optimizer='adam') #编译模型  model.fit(x\_train, y\_train, nb\_epoch = 10000, batch\_size = 16) #训练模型，学习一万次  model.save\_weights(modelfile) #保存模型参数  #预测，并还原结果。  x = ((data[feature] - data\_mean[feature])/data\_std[feature]).as\_matrix()  data[u'y\_pred'] = model.predict(x) \* data\_std['y'] + data\_mean['y']  data.to\_excel(outputfile)  import matplotlib.pyplot as plt #画出预测结果图  p = data[['y','y\_pred']].plot(subplots = True, style=['b-o','r-\*'])  plt.show() |