**乐清电厂烟气脱硫运行数据分析报告**

* **数据基本情况**：共含有1061个样本，20个过程变量，1个质量变量（具体详情见表1），样本按时间顺序采集，采样间隔为5分钟，数据集时间跨度为一周。由于数据按时序排列，因此并未采用随机方式划分数据，而是按时间顺序将前80%样本作为训练集，将后20%样本作为测试集。

表1 数据集中具体变量展示

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **过程变量** | #1机组负荷 | #1机组送风量 | #1原烟气反吹保持O2浓度 | #1原烟气SO2基准浓度 |
| #1原烟气烟尘基准浓度 | #1机组选择后原烟气温度 | #1机组选择后原烟气压力 | #1脱硫入口原烟气流量 |
| #1吸收塔石膏排出泵出口母管#1PH计 | #1吸收塔石膏排出泵出口母管#2PH计 | #1吸收塔浆液密度仪 | #1吸收塔供浆电磁流量计 |
| #1吸收塔A再循环泵电流 | #1吸收塔B再循环泵电流 | #1吸收塔C再循环泵电流 | 1D吸收塔再循环泵电流 |
| 选择后#1吸收塔出口烟温 | 1吸收塔出口烟气压力 | #1FGD净烟气流速 | #1标准状态下干烟气流量 |
| **质量变量** | #1净烟气SO2基准浓度 | | | |

* **软测量性能分析**：分别建立了6种软测量模型，通过表2可以看出各模型还是具有一定的预测能力，但性能较差，即便是预测性能最好的LSTM在红圈部分也存在较大偏差，原因可能是数据量较小，导致测试部分出现了训练部分不存在的工况条件，如果能有更多数据支撑应该能够提升性能。同时，利用数据间的时序特性对性能提升比较有效，因此动态软测量模型更适合解决该问题。后续可以考虑在预测某一时刻质量变量时，同时利用过去时刻的过程变量和质量变量信息，以提升模型性能。

表2 各模型软测量预测性能对比

|  |  |  |
| --- | --- | --- |
| **模型** | **R2(%)** | **RMSE(浓度)** |
| Lasso | 23.74 | 5.15 |
| Linear SVR | 31.07 | 4.90 |
| RandomForest | 36.33 | 4.71 |
| XgBoost | 59.28 | 3.77 |
| Fully Connected Networks | 66.98 | 3.39 |
| **Long Short-Term Memory** | **72.06** | **3.12** |

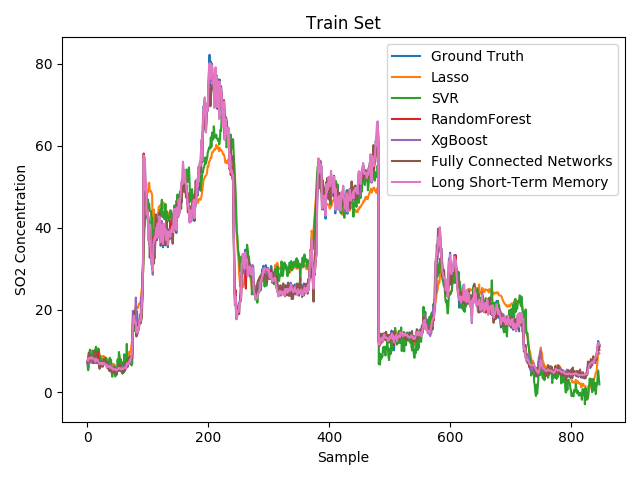
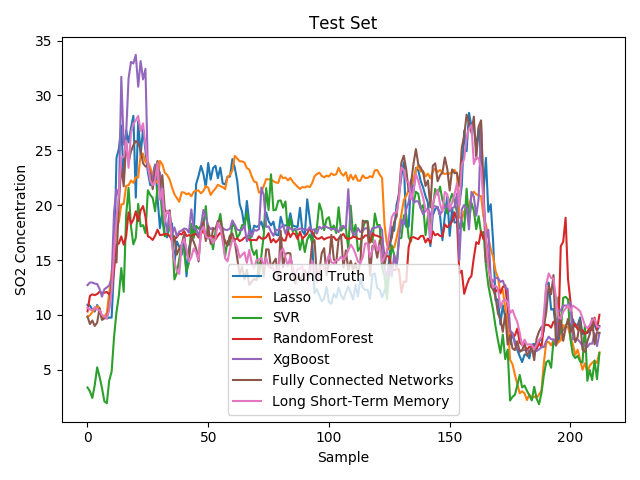
 

图1 训练集拟合效果 图2 测试集预测效果

* **质量变量影响因素分析**：分别利用基于相关系数和互信息的统计量、Lasso和Linear SVR的回归系数、RandomForest和XgBoost中基于不纯度的变量重要性得分，评价不同变量的重要程度，图3~8为各方法中的变量得分情况（得分越高表示越重要）。图9为综合考虑这些方法的变量重要性排名情况（排名越小表示越重要），用绿色框框出了排名前5的变量，在表3中用蓝色对应标出。其中可以看到原烟气SO2浓度对出口烟气SO2浓度影响最大，较符合基本的认知；两个PH计本质上反映了对于SO2的吸收作用强弱，因此也有较大影响。实际中需要在可操作的变量中寻找影响较大的变量，进而能够根据预测结果及时调整操作，使得出口烟气中SO2浓度符合排放标准。因此通过上述分析，供浆流量为影响较强且可操纵的变量，可通过调整供浆流量有效调节出口烟气的SO2浓度。后续应当在明确哪些变量是可操纵的基础上，再进行变量重要性的评价。

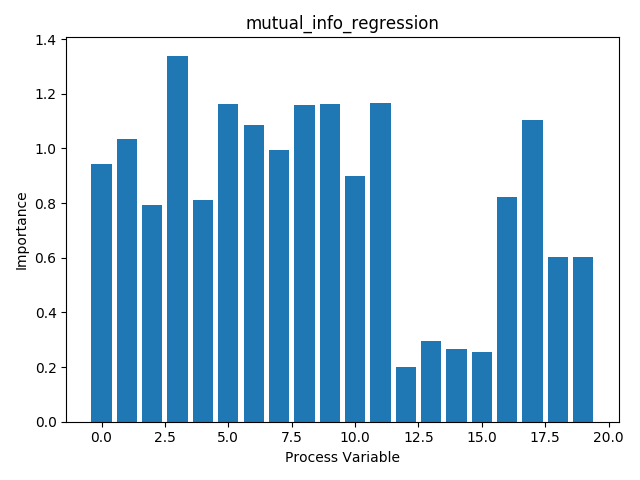
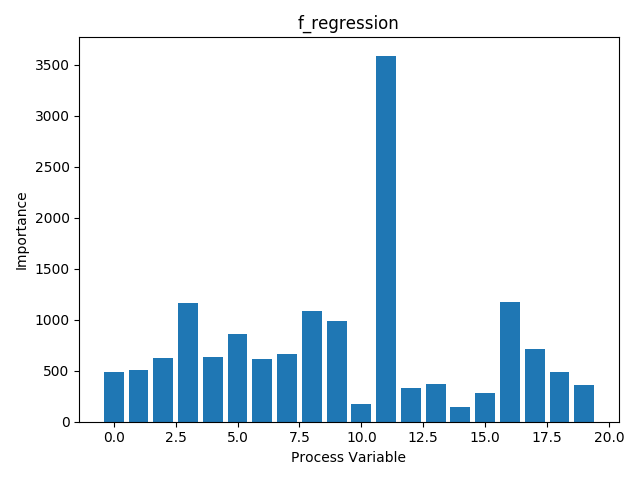


图3 f\_regression 图4 mutual\_info\_regression

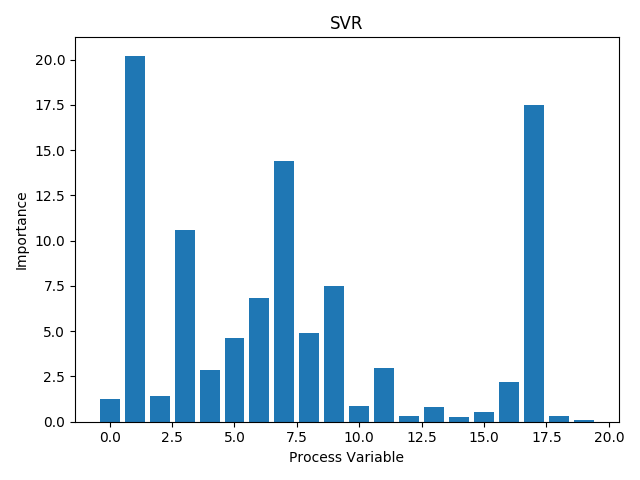
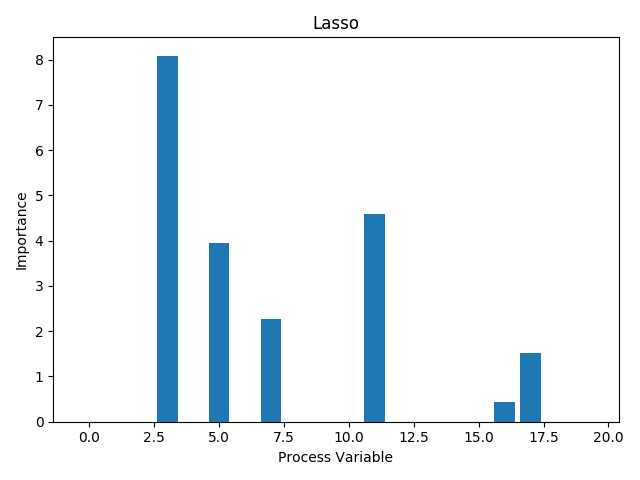


图5 Lasso 图6 SVR

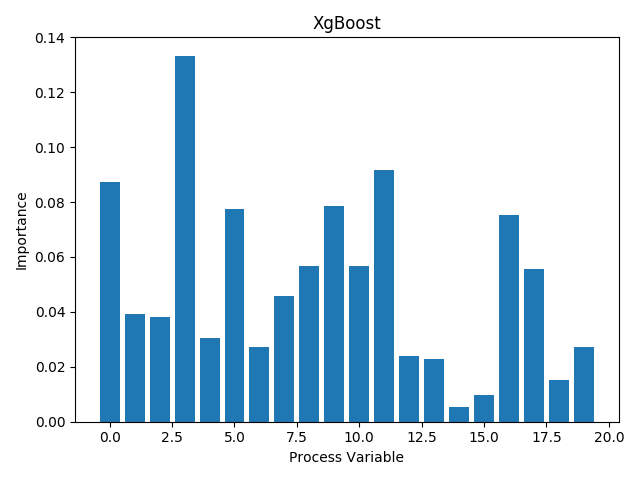
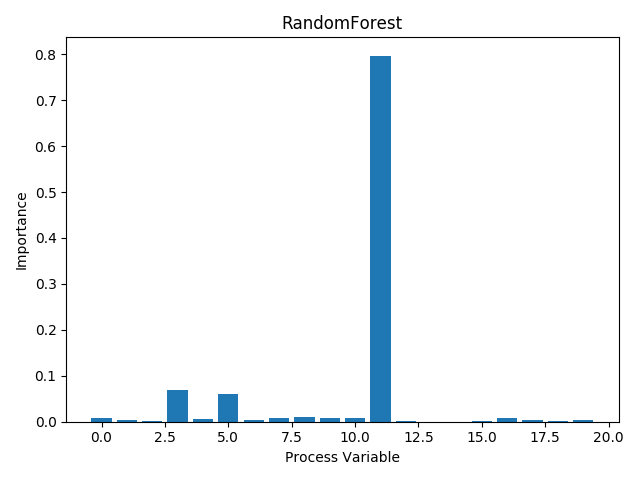


图7 RandomForest 图8 XgBoost

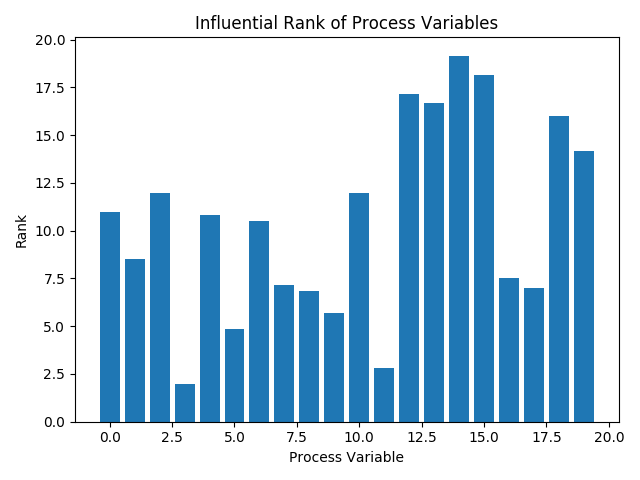


图9 各变量综合排名情况

表3 各过程变量具体综合排名情况

|  |  |  |  |
| --- | --- | --- | --- |
| 过程变量 | 平均排名 | 过程变量 | 平均排名 |
| #1机组负荷 | 11.0 | #1吸收塔浆液密度仪 | 12.0 |
| #1机组送风量 | 8.5 | **#1吸收塔供浆电磁流量计** | **2.8** |
| #1原烟气反吹保持O2浓度 | 12.0 | #1吸收塔A再循环泵电流 | 17.2 |
| **#1原烟气SO2基准浓度** | **2.0** | #1吸收塔B再循环泵电流 | 16.7 |
| #1原烟气烟尘基准浓度 | 10.8 | #1吸收塔C再循环泵电流 | 19.2 |
| **#1机组选择后原烟气温度** | **4.8** | 1D吸收塔再循环泵电流 | 18.2 |
| #1机组选择后原烟气压力 | 10.5 | 选择后#1吸收塔出口烟温 | 7.5 |
| #1脱硫入口原烟气流量 | 7.2 | 1吸收塔出口烟气压力 | 7.0 |
| **#1吸收塔石膏排出泵出口母管#1PH计** | **6.8** | #1FGD净烟气流速 | 16.0 |
| **#1吸收塔石膏排出泵出口母管#2PH计** | **5.7** | #1标准状态下干烟气流量 | 14.2 |