

# 基于变量压缩-BP 神经网络模型的火电厂 SO<sub>2</sub> 排放预测 研究报告

作者：李东峰 李至瑞  
指导老师：石磊  
学校：山东省实验中学

声明：本人郑重声明：所提交的数学应用论文是本人在指导老师的指导下独立进行研究的成果，除文中已经注明引用的内容外，本文不含其它个人或集体已发表或撰写过的作品成果。对本人的研究做出重要贡献的个人和集体已在文中以明确方式标明。

## 一、摘要

论文题目: 基于变量压缩-BP 神经网络模型的火电厂 SO<sub>2</sub> 排放预测研究报告

针对济南黄台电厂湿式石灰石-石膏法脱硫系统能耗偏高且难以建立二氧化硫排放动态预测模型的问题, 提出一种基于的变量压缩-BP 神经网络的动态预测方法。现场测取了 2018 年 7 月间 2676 组现场数据, 针对脱硫过程中建模对象具有的强非线性及多个输入变量间的相关和耦合特性, 采用非负熵率 (nonnegative garrote, NNG) 方法进行了变量选择, 将选取的变量集作为 BP 神经网络 (back-propagation neural network) 的输入, 建立了 NNG-BP 神经网络模型。将该模型与原始的 BP 神经网络模型进行对比, 结果表明 NNG-BP 神经网络模型显著提高了模型的预测精度。本研究所建立的模型及研究结果, 可以为石灰石-石膏法脱硫系统的优化提供参考。

关键词: 石灰石-石膏法脱硫; 变量选择; BP 网络; NNG-BP 控制系统

To solve the problem of high energy consumption in existing in wet limestone-gypsum desulfurization process and dynamic prediction model, a dynamic prediction method for SO<sub>2</sub> concentration was proposed based on nonnegative garrote and back-propagation neural network. 2676 sets of field data were obtained on the spot in July 2018. In view of the industrial process modeling objects generally have strong nonlinear and the variables are relevant and coupled, this paper adopted nonnegative garrote (NNG) method for variable selection. NNG-BP model was established based on the selected optimal variable set as model input. In addition, combining NNG method with back-propagation neural network, the NNG-BP model was constituted. Comparing NNG-BP model with the back-propagation neural network model without the variable selection. The results show that the NNG-BP neural network model improves the prediction accuracy of the model significantly. The model and results can be used as reference for the optimization of limestone-gypsum desulfurization system.

Keywords: wet limestone-gypsum desulfurization, variable selection, BP networks, NNG-BP control system

## 目录

1 研究背景.....	4
2 湿法烟气脱硫技术.....	5
3 黄台电厂脱硫工艺.....	6
3.1 系统组成.....	6
3.2 石灰石-石膏湿法烟气脱硫工艺.....	8
4 基于神经网络的 SO <sub>2</sub> 软测量技术.....	11
4.1 神经网络结构.....	12
4.1.1 BP 神经网络结构.....	12
4.1.2 非负阈值算法.....	13
4.1.3 整合 NNG 和 ANN.....	14
4.1.4 评估标准.....	14
4.1.5 s 的选择.....	14
4.2 MATLAB 程序的实现.....	15
4.2.1 数据预处理的编程实现.....	15
4.2.2 NNG-BP 的编程实现.....	20
4.3 基于神经网络预测模型仿真实例.....	22
4.3.1 BP 神经网络算法对烟气.....	23
4.3.2 NNG-BP 神经网络算法对烟气 SO <sub>2</sub> 浓度的预测.....	24
5 小结.....	25
参考文献.....	27

# 1 研究背景

从改革开放至今的四十年里, 中国的电力工业迅猛发展, 中国的煤电电力装机容量提高了约 29 倍, 在发电总量上提升了约 23 倍<sup>[1]</sup>。发电行业是优质电能的创造者, 但同时也是主要的一次能源消耗者和多种形态污染物的排放大户。我的父亲是黄台电厂的职工, 我从小在电厂宿舍长大, 小时候宿舍区地面总是黑乎乎的, 我就想等我长大了一定要把电厂治理得干干净净, 现在的电厂治理得非常干净, 绿化的像花园了。2015 年国家规定所有燃煤机组必须进行超低排放改造, 在基准氧含量 6% 条件下, 二氧化硫排放浓度由 200 毫克/标立方米降低到不高于 35 毫克/标立方米。目前黄台电厂的脱硫厂里的能耗大户, 脱硫系统存在能耗偏高的不足, 因此脱硫设备的节能降耗及经济运行成为目前研究的热点。在国家“十二五”规划中, 对火电厂 SO<sub>2</sub> 排放提出了更严格的标准, 脱硫系统的节能优化运行对实现企业效益和社会效益的双赢具有重要意义。

国家出台的相关的火电排放标准, 且已经取得了比较理想的成绩, 据相关统计, 截止到 2017 年底, 全国已经有 71% 的燃煤机组已经完成了改造, 实现了超低排放。排放烟气中的二氧化硫、灰尘等完全达到国际最严格环保标准的要求<sup>[2]</sup>。在煤燃烧过程中释放出的二氧化硫是造成大气污染的主要污染源, 因此如何去除煤燃烧产生的二氧化硫至关重要。脱硫的方法按阶段可以分为燃烧前脱硫、燃烧中脱硫和燃烧后脱硫, 燃烧后脱硫又称烟气脱硫 (FGD), 主要是针对燃料燃烧后生成的烟气进行处理, 由于其操作方便, 是目前唯一大规模应用的脱硫方法。按照烟气脱硫过程中的干湿形态, 可以分为干法、半干法和湿法脱硫<sup>[3]</sup>, 其中湿法烟气脱硫技术脱硫效率高, 一般可以达到 95% 以上, 单机烟气处理量大, 可与大型锅炉单元匹配, 在世界范围内被广泛应用到各大电厂中。

石灰石-石膏湿法烟气脱硫技术是将石灰石粉加水制成浆液作为吸收剂泵入吸收塔与烟气充分接触混合, 烟气中的二氧化硫与浆液中的碳酸钙以及从塔下部鼓入的空气进行氧化反应生成硫酸钙, 硫酸钙达到一定饱和度后, 结晶形成二水石膏。经吸收塔排出的石膏浆液经浓缩、脱水, 使其含水量小于 10%, 然后用输送机送至石膏贮仓堆放, 脱硫后的烟气经过除雾器除去雾滴, 再经过换热器加热升温后, 由烟囱排入大气。由于吸收塔内吸收剂浆液通过循环泵反复循环与烟气接触, 吸收剂利用率很高, 钙硫比较低, 脱硫效率可大于 95%。

大多数公司采用湿式脱硫技术, 其反应原理基本类似, 主要区别在于吸收塔的结构上, 有的采用单回路喷淋空塔, 有的采用双回路喷淋塔和鼓泡塔等<sup>[4]</sup>。不同的脱硫工艺, 脱硫效率也会有所不同。本文所描述的电厂, 采用了石灰石-石膏湿法烟气脱硫工艺, 采用双塔结构, 在很大程度上提高了脱硫的效率。

近年来, 计算机技术和数据存储能力快速发展, 大量数据源源不断地在各行各业中累积, 如何从海量数据中获取价值信息并提升企业价值已经成为当下研究热点。当下的互联网行业中, 物联网 (IOT) 技术, 云计算技术不断地发展以及成熟, 大数据 (bigdata) 已经成为这个时代当之无愧的热点<sup>[5]</sup>。大数据技术为现代企业发展提供了机遇和挑战。对于一些传统产业, 大数据的出现带来了巨大的挑战, 倒闭产业进行重新整合, 产业内的一些企业也不得不开始变革。而对于现代化企业, 其数据资产的规模、活性以及其解释、运用数据的能力开始直接与企业价值挂钩。

人工智能是计算机科学的一个分支, 它企图了解智能的实质, 并生产出一种新的能以人类智能相似的方式做出反应的智能机器, 该领域的研究包括机器人、语言识别、图像识别、自然语言处理和专家系统等。人工智能从诞生以来, 理论和技术日益成熟, 应用领域也不断扩大。人工智能不是人的智能, 但能像人那样思考、也可能超过人的智能。人工智能是包括十分广泛的科学, 它由不同的领域组成, 如机器学习, 计算机视觉等等, 总的说来, 人工智

能研究的一个主要目标是使机器能够胜任一些通常需要人类智能才能完成的复杂工作。

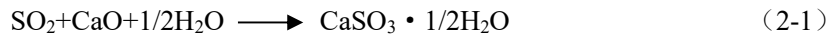
随着生产规模的日益扩大,工业过程的系统复杂度也不断提高,如何建立描述这些系统行为的模型,成为研究人员关注的一个焦点。建立工业过程模型一般有三种途径:<sup>[6]</sup> (1) 从系统内在机理出发,推导出机理模型,机理建模方法通常要针对具体问题仔细分析系统的动力学特性,建立物理化学平衡方程;(2) 根据输入输出信号,用系统辨识或参数估计求取一个微积分或差分数学模型,系统辨识或状态估计主要是一些信号处理方法;(3) 根据系统运行状态下的输入输出数据,通过智能方法,建立一个基于知识的模型。利用知识发现和人工智能的建模方法是近年以来兴起的一个方向。

在这些多变量、非线性时变的工业过程系统中,其物理或化学动力学过程具有以下特点:状态变量维数高、各变量强相互关联、检测信号粗糙而不完整、干扰随机、系统参数时变等等。这就要求被控对象的动力学模型不仅有学习和识别的能力,而且对环境和干扰有较强的鲁棒性,所以建立数学解析模型具有一定的困难。因此,作为数学解析建模的一种补充方法,过程辨识和人工智能是研究者们关注的一个研究方向。由于过程辨识和人工智能能够在信息量不足、不完备的情况下,对那些用数学模型难以描述的复杂系统进行建模,在工业过程、计划决策、商品制造中得到了广泛应用<sup>[7]</sup>。

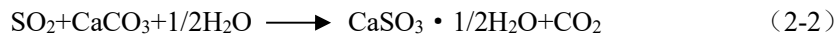
## 2 湿法烟气脱硫技术

目前广泛使用的脱硫方法为湿法烟气脱硫技术,主要是利用石灰或者石灰石与二氧化硫会产生化学反应,将烟气中的二氧化硫吸收。其化学反应原理为:

石灰法:

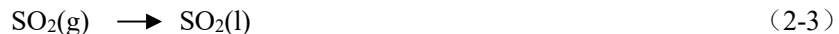


石灰石法:



其具体反应可以细分如下:

(1) 二氧化硫的吸收过程:



(2) 溶解过程:



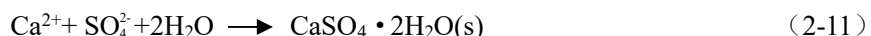
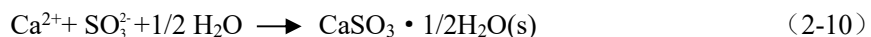
(3) 中和过程:



(4) 氧化过程:



(5) 结晶过程:



含有二氧化硫的烟气从一级吸收塔底部由下往上运动,遇到从喷淋层发出的液态悬浮液,根据双膜理论,二氧化硫以分子扩散的方式通过气膜和液膜,与碱性悬浮液发生化学反应,去除烟气中的二氧化硫,悬浮液下落汇集,在浆液池中完成  $\text{CaSO}_4$  的氧化与  $\text{CaSO}_4 \cdot 2\text{H}_2\text{O}$  的结晶<sup>[8]</sup>,最终在吸收塔底部形成石膏。

### 3 黄台电厂脱硫工艺

某电厂运行 2 台 33 万千瓦机组, 供热抽气能力 670 吨/小时, 是某市电压支撑点和最大的热源厂。其采用的就是湿法烟气脱硫技术, 本文描述的脱硫工艺流程和相关参数, 主要是来自某电厂 9 号机组, 该机组的主要特点是采用了双吸收塔结构, 这样的构造相对于单塔设

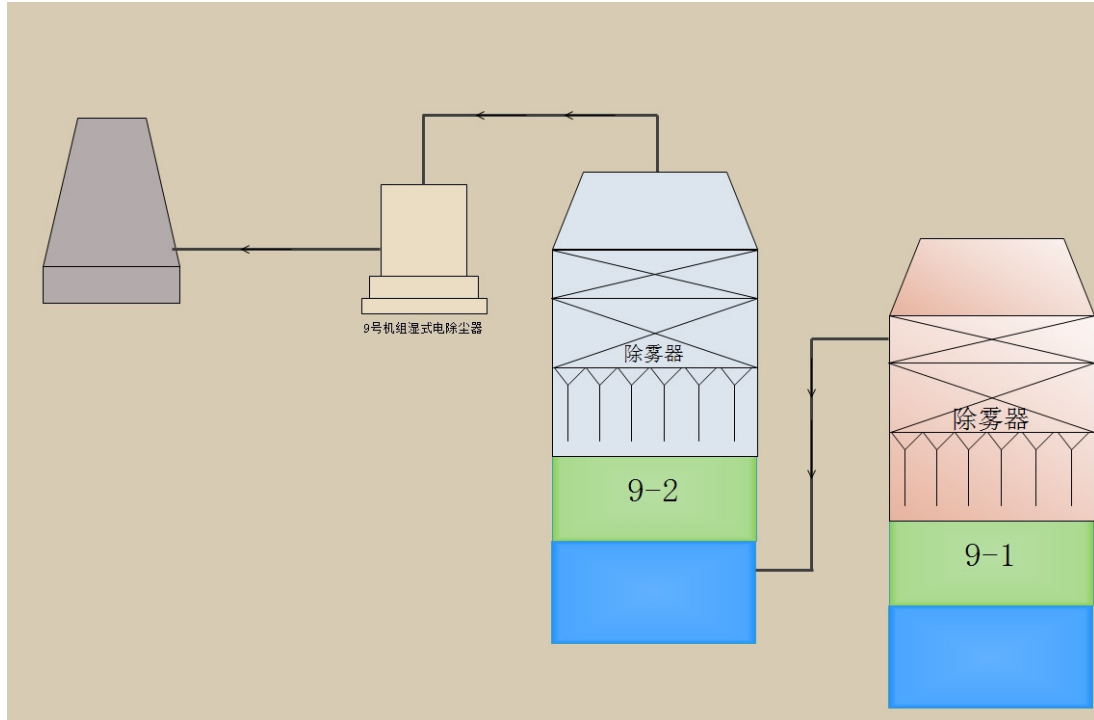


图 3-1 系统整体框架

计来说, 可以对排出的烟气进行二次处理, 能够更加有效的去除烟气中的二氧化硫, 整个系统框架如图 3-1 所示。某电厂的烟气脱硫系统主要包括二氧化硫吸收系统; 烟气系统; 吸收剂制备系统; 石膏处理系统; 工艺水系统; 事故浆液系统等。

#### 1.1 系统组成

##### (1) 烟气系统

本系统的作用是将锅炉原烟气引入一二级吸收塔, 将脱硫后的净烟气经湿式除尘器引入烟囱排放。升压风机将烟气送入 1 号吸收塔底部, 在进入吸收塔之前进行降温处理, 在一定的范围内, 温度越低, 越有利于二氧化硫的吸收。烟气由底部向上升, 经过处理后从一级吸收塔顶部进入二级吸收塔底部进行烟气的第二处理。

本系统锅炉、脱硫、脱硝风机三合一, 从锅炉引风机后的总烟道上引出的烟气直接进入吸收塔入口。烟气在吸收塔内洗涤净化, 经除雾器除去水雾后接入水平主烟道经烟囱排入大气。原设计在水平主烟道上设置的旁路挡板门进行封堵, 设置的挡板密封风系统和辅助电加热器退出运行。当机组启动时, FGD 系统也随机启动。

脱硫取消旁路后, 为防止多台循环浆泵同时故障跳闸, 造成吸收塔除雾器和防腐衬胶得不到及时冷却, 超温损坏。在吸收塔入口原烟道处增加事故喷淋系统, 以降低烟气温度。即: 从工艺水母管敷设到原烟道吸收塔入口的事故减温水管道, 工艺水母管接口位置安装 DN150



手动隔离门, 在烟道附件便于操作位置安装 1 个 DN150 气动阀门, 阀板材质 1.4529, 门后采用 316L 管道。手动隔离门正常运行时处于常开位置, 气动阀门在检修时手动隔离门关闭进行隔离。在事故减温水管道下方安装 8 根喷淋支管, 材质 316L, 喷淋支管底部焊接 316L 堵板后焊接固定在烟道底部, 每根支管上安装 3 个喷嘴, 24 个喷嘴在 9.6m \* 3.6m 原烟道截面上均布。喷淋母管采取内缠电伴热带, 外包硅酸铝棉和保温铝板进行保温。二级吸收塔不设喷淋降温装置。

## (2) 除雾器系统

湿法吸收塔在运行过程中, 易产生粒径为 10~60 微米的“雾”。“雾”不仅含有水分, 它还溶有硫酸、硫酸盐、SO<sub>2</sub>等, 如不妥善解决, 任何进入烟囱的“雾”, 实际就是把 SO<sub>2</sub> 排放到大气中。因此, 工艺上对吸收设备提出除雾的要求。被净化的气体在离开吸收塔之前要进行除雾。通常除雾器多设在吸收塔的顶部。

一级吸收塔上部设二级屋脊式除雾器除雾, 使出口烟气中液滴浓度小于 75mg/Nm<sup>3</sup>。除雾器配有带喷嘴的冲洗管道, 除雾器清洗水经喷嘴依次冲洗除雾器上沉积的固体颗粒, 除雾器冲洗水还可以起到调节吸收塔浆池液位的作用。

二级塔除雾器采用一层管式再加两层屋脊布置, 在除雾器上下部均设冲洗水系统。

运行时根据设定程序进行自动冲洗, 也可进行人工冲洗。冲洗用水为 FGD 工艺水。除雾器材料采用带加强的阻燃聚丙烯, 能承受高速水流冲刷, 特别是人工冲洗造成的高速水流冲刷。

除雾器包括分离器、冲洗管、喷嘴及构件等, 安装在吸收塔上部, 用以分离净烟气夹带的雾滴。除雾器出口烟气湿度不大于 75mg/Nm<sup>3</sup>(干基)。烟气经过除雾器处理后应当垂直离开除雾器。

除雾器的最大表面烟气流速按吸收塔内烟气流速+25%进行设计, 并保证在吸收塔内烟气流速偏差±25%情况下的除雾效率。除雾器应确保烟气温度 85℃除雾器不会损坏, 烟气温度 90℃条件下连续运行时间不少于 20 分钟不会损坏。

除雾器覆盖除设置支撑结构处外的整个吸收塔横截面, 除雾器应保证避免出现在支撑梁与吸收塔壁之间或在除雾器任何部分烟气未经除雾处理而“泄漏”的情况。除雾器系统的设计特别考虑 FGD 装置入口的飞灰浓度的影响。该系统还包括去除除雾器沉积物的冲洗和排水系统。

除雾器采用独立的单元结构, 冲洗系统运行时根据给定或可编程序自动冲洗, 既可进行自动冲洗, 也可进行人工冲洗。冲洗喷嘴为全锥喷射形, 喷角不大于 120°。为防止堵塞, 一级除雾器上下部及二级除雾器下部配备性能良好的冲洗系统。冲洗系统能够对除雾器进行全面冲洗。喷嘴与除雾器冲洗水母管连接, 邻近喷嘴的喷淋范围部分重叠, 以确保 100%的冲洗效果, 喷淋范围有 150%的叠加(平均值)。

## (3) 吸收塔溢流装置

一级吸收塔浆液池上部设溢流口, 保证吸收塔内浆液不会倒灌入入口烟道内。吸收塔顶部设有放空阀, 正常运行时阀门关闭, FGD 装置停运时阀门开启, 方便检修维护。

二级吸收塔浆液池上部设溢流口, 液位高时溢流至一级吸收塔。吸收塔顶部通过烟道与湿除连接, 不设放空阀。

氧化空气系统是存在于吸收氧化系统内的一个重要部分, 氧化空气的功能是促使吸收塔浆液池内石膏的形成。氧化空气注入不充分将会引起分离效率的降低, 还可能导致吸收塔中的结垢, 因此, 该部分的优化设置对提高整个设备的脱硫效率和石膏产品的质量显得尤为重要。

## (4) 吸收塔浆液搅拌系统

由于吸收塔浆液池容积较大, 容易造成吸收塔内固体物质的沉淀。同时新鲜石灰石浆液

是从吸收塔液相区某高度进入塔内的, 石灰石浆液量较循环浆液量很小, 仅靠自身的渗透作用很难实现与浆液池中的浆液充分混合。为防止吸收塔浆液中固体物的沉积,

一级塔布置两台流量 1650m<sup>3</sup>/h, 扬程 22m 的脉冲悬浮泵, 一用一备。

二级塔布置两台流量 1060m<sup>3</sup>/h, 扬程 22m 的脉冲悬浮泵, 一用一备。

#### (5) 氧化风机

一级吸收塔的氧化空气系统由氧化风机、氧化曝气装置及相应的管道、阀门组成。氧化风机供给的空气通过布置在浆池内的曝气装置进入塔内。氧化风机提供足够的氧化空气, 使吸收塔内的亚硫酸钙充分转化成硫酸钙。一级塔氧化风机更换为流量 12000Nm<sup>3</sup>/h, 压头 85kPa 的离心式风机。布置方式改造为管网式布置, 在塔底 4m 高处布置 5 条氧化空气管道。氧化风机出口设置减温装置, 进塔温度不得高于 50℃。二级塔生成的石膏较少, 不单独设置氧化风机。通过石膏旋流泵将生成浆液打回一级吸收塔氧化并排出。

在本系统每座一级吸收塔设置 2 台氧化风机 (1 用 1 备), 用以向吸收塔中输入氧化空气。

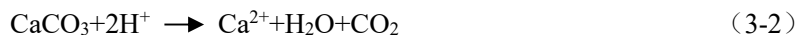
#### (6) 吸收塔排空装置

在脱硫系统解列或出现事故停机需要检修时, 一级吸收塔内的浆液通过石膏浆液排出泵排入事故浆液箱中, 二级吸收塔内的浆液通过脉冲悬浮泵排入事故浆液箱中, 以便对吸收塔进行维修。

### 3.2 石灰石-石膏湿法烟气脱硫工艺

某电厂 9 号机组拥有两座吸收塔, 吸收区域直径为 12 米, 高度为 31.4 米。机组采用石灰石-石膏湿法脱硫, 增容改造采用双塔串联方式, 一级吸收塔入口设计浓度为 4834 mg/Nm<sup>3</sup>, 二级吸收塔入口浓度为 600mg/Nm<sup>3</sup> (标态、干基、6%O<sub>2</sub>), 二级吸收塔出口排放浓度小于 32mg/m<sup>3</sup> (标态、干基、6%O<sub>2</sub>), 在二级吸收塔后增设湿式除尘器, 满足重点地区 SO<sub>2</sub> 排放浓度为 50mg/Nm<sup>3</sup>, 烟尘排放浓度 20mg/Nm<sup>3</sup> 的标准。运营中执行超低排放标准: 即 SO<sub>2</sub> 排放浓度小于 35mg/Nm<sup>3</sup>, 烟尘排放浓度小于 5mg/Nm<sup>3</sup>。原烟从锅炉引出后, 先进入一级吸收塔进行脱硫处理, 再进入二级吸收塔进行进一步的脱硫处理, 整个过程处理完成后的烟气, 将完全符合国家排放标准。

一级吸收塔包括四层由循环泵控制的喷淋层, 喷淋区直径为 12 米, 这是整个脱硫过程的关键; 喷淋层所需要的浆液有两部分构成, 一部分是来自供浆系统, 另一部分是循环使用的浆液, 两部分混合形成的浆液由吸收塔顶部向下成雾状喷射, 多层喷雾相互叠加, 形成雾滴下落, 与从塔底上升的烟气形成了逆流流, 浆液中的水吸收烟气中的 SO<sub>2</sub>, 反应式为:



雾滴也会将烟气中的灰尘包裹起来, 落入塔底的浆液中, 形成石膏。

根据电厂的工作人员长时间的工作经验, 在该吸收塔中浆液的最佳 PH 值为 5.5~6.5 之间, 超过 6.5 之后, 容易使吸收塔壁结垢, 低于 5.5 时, 会大大降低脱硫的效率。



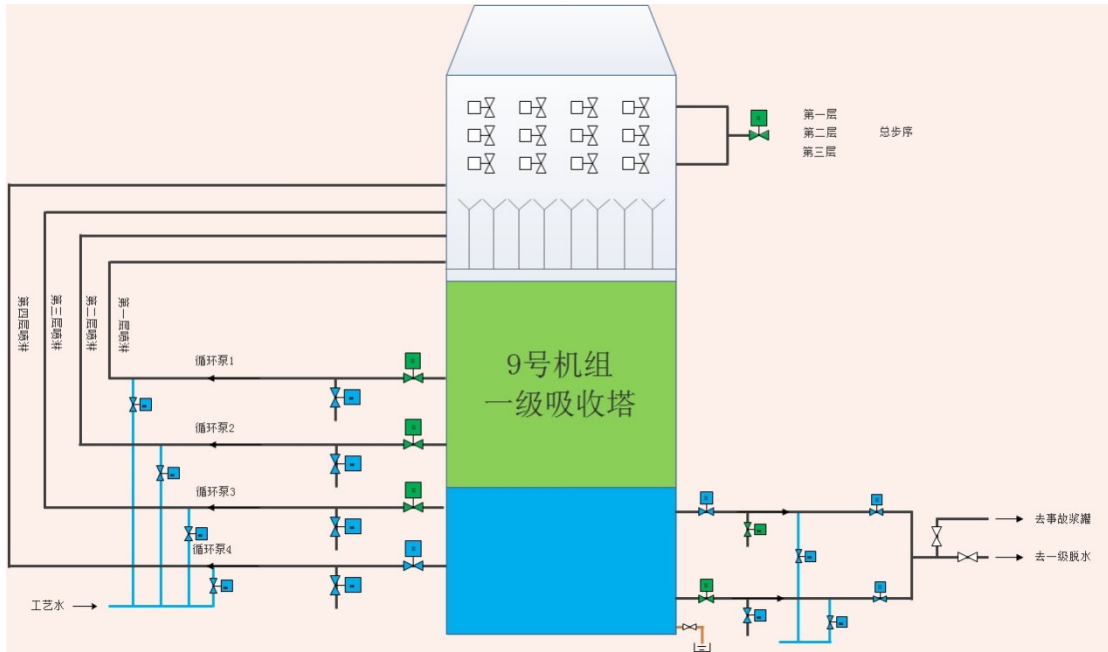


图 3-2 一级吸收塔

二氧化硫吸收系统是烟气脱硫系统的核心,其作用是将引入的锅炉原烟气在吸收塔内通过吸收塔浆液的喷淋洗涤去除大量的  $\text{SO}_2$ , 脱硫反应生成的脱硫产物在吸收塔浆池中被通入的氧化空气强制反应生成硫酸钙并在浆池中结晶生成二水石膏。

塔体的设计采用脉冲悬浮射流搅拌措施来避免浆池中浆液沉淀。吸收塔底面设计能完全排空浆液。氧化空气管采用管网式布置。

进入一级吸收塔的石灰石浆液在吸收塔浆池中溶解,通过调节进入吸收塔的石灰石浆液量或吸收塔排出浆液浓度,使吸收塔浆池 PH 值维持在 5.5~6.5 之间以保证石灰石的溶解及  $\text{SO}_2$  的吸收。

原烟气先进入一级吸收塔后,逆流通过喷淋区,与浆液充分接触,二氧化硫、氯化氢、氟化氢这些有害成份被吸收,烟气中的部分粉尘也被除去。再进入二级吸收塔,烟气中剩余的二氧化硫等有害成份再次,逆流通过喷淋区被吸收,最后通过湿式除尘器去除粉尘后排至烟囱。

一级吸收塔有四个循环浆泵,电机功率逐层增加,第四层喷淋层对应的电机功率最大。表 3-1 是一级吸收塔循环浆泵的规范

表 3-1 一级吸收塔循环浆泵规范

项目	单位	1 泵	2 泵	3 泵	4 泵	
台数	台	1*2	1*2	1*2	1*2	
形式		卧式离心泵	卧式离心泵	卧式离心泵	卧式离心泵	
转速	rpm	596	614	627	645	
轴功率	kw	630	710	800	800	
电机	型号	YKK450-4	YKK450-4	YKK450-4	YKK450-4TH	
	同步转速	rpm	1490	1490	1490	1490
	额定功率	kw	500	560	630	710
	额定电流	A	63	63	70.5	79.3
	额定电压	V	6000	6000	6000	6000

根据 BMCR 工况下烟气量以及烟气中  $\text{SO}_2$  含量,一级吸收塔设置 4 台(每台流量为

7600m<sup>3</sup>/h) 浆液循环泵 (编号为 1、2、3、4 循环浆泵), 采用 4 层浆液雾化喷淋方式, 在系统低负荷运行时, 可以通过停用部分浆液循环泵而保证一定的液气比。一级吸收塔为变径塔, 底部浆池区域直径为 15m, 正常液位为 10.0m, 容积为 2186m<sup>3</sup>, 上部二氧化硫吸收区域直径为 12m。二级吸收塔为圆柱塔, 直径为 12.5m, 高度 28 米。设三层喷淋层, 浆液分别由三台流量为 5000m<sup>3</sup>/h 循环浆泵提供 (编号为 5、6、7 循环浆泵)。正常液位为 10.8m, 容积为 1301m<sup>3</sup>

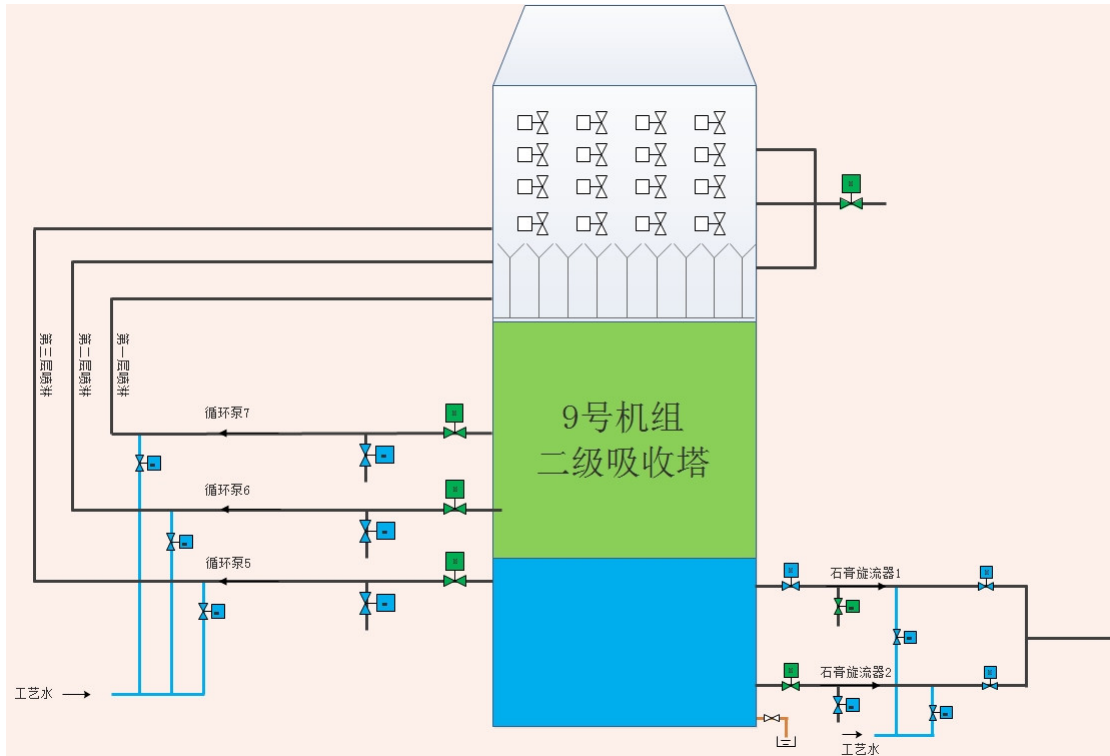


图 3-3 二级吸收塔

二级吸收塔与一级吸收塔采用相同的流程, 从一级吸收塔顶端排出的烟气从二级吸收塔的底部进入, 进行第二次脱硫处理, 如图 3-2 由于一级吸收塔吸收了烟尘中的绝大多数的二氧化硫, 二级吸收塔减少了喷淋的数量, 增加了雾化层的数量。烟气经过二级吸收塔的再次处理之后, 已经完全达到了排放标准。二级吸收塔的循环浆泵与一级吸收塔循环浆泵略有差别, 具体规范如表 3-2 所示。

表 3-2 二级吸收塔循环浆泵规范

项目	单位	5 泵	6 泵	7 泵	
台数	台	1	1	1	
形式		卧式离心泵	卧式离心泵	卧式离心泵	
转速	rpm	743	743	744	
电机	型号	YXKK560-8	YXKK560-8	YXKK560-8	
	同步转速	rpm	743	743	744
	额定功率	kw	500	560	630
	额定电流	A	63.1	70.5	77
	额定电压	V	6000	6000	6000

在实际生产过程中, 正常情况下, 两座吸收塔的循环水泵的开启方式采用了“3+2”的运行模式, 即一级吸收塔开启三个循环水泵, 二级吸收塔开启两个循环水泵, 其余水泵留作备用。在深夜电厂负荷较低的情况下, 为了节能, 会采取“2+1”的运行模式, 在达到脱硫

的目的的同时降低了消耗。

由于大部分的反应是在一级吸收塔内完成, 大部分的石膏也是在一级吸收塔内产生, 为了方便对石膏的处理, 二级吸收塔最初产生的石膏密度达到一定程度后通过旋流器旋流之后输送到一级吸收塔内进行统一处理。

## 4 基于神经网络的 SO<sub>2</sub> 软测量技术

软测量自 20 世纪 80 年代中后期作为间接测量技术的一个发展方向被提出以来, 研究异常活跃, 发展十分迅速, 应用日益广泛, 几乎渗透到工业领域的各个方面, 已成为检测技术的主要研究方向之一。

当前的生产过程极其复杂, 具有高度的非线性, 采用线性软测量方法建模会造成模型的解释性不足和预测精度低的问题, 难以满足测量与控制的需要。因此出现了很多非线性建模方法, 如神经网络 (ANN)、支持向量机(SVM)、多模型理论等。近年来, 对于非线性建模中的变量选择问题引起控制领域的广泛关注。

人工神经网络 (Artificial Neural Networks, ANN) 系统是 20 世纪 40 年代后出现的。它是由众多的神经元可调的连接权值连接而成, 具有大规模并行处理、分布式信息存储、良好的自组织自学习能力等特点。

(1) 人工神经网络也具有初步的自适应与自组织能力。

人类大脑有很强的自适应与自组织特性, 后天的学习与训练可以开发许多各具特色的活动功能。如盲人的听觉和触觉非常灵敏; 聋哑人善于运用手势; 训练有素的运动员可以表现出非凡的运动技巧等等。普通计算机的功能取决于程序中给出的知识和能力。显然, 对于智能活动要通过总结编制程序将十分困难。在学习或训练过程中改变突触权重值, 以适应周围环境的要求。同一网络因学习方式及内容不同可具有不同的功能。人工神经网络是一个具有学习能力的系统, 可以发展知识, 以致超过设计者原有的知识水平。通常, 它的学习训练方式可分为两种, 一种是有监督或称有导师的学习, 这时利用给定的样本标准进行分类或模仿; 另一种是无监督学习或称无为导师学习, 这时, 只规定学习方式或某些规则, 则具体的学习内容随系统所处环境 (即输入信号情况) 而异, 系统可以自动发现环境特征和规律性, 具有更近似人脑的功能。

(2) 泛化能力

泛化能力指对没有训练过的样本, 有很好的预测能力和控制能力。特别是, 当存在一些有噪声的样本, 网络具备很好的预测能力。

(3) 非线性映射能力

当对系统对于设计人员来说, 很透彻或者很清楚时, 则一般利用数值分析, 偏微分方程等数学工具建立精确的数学模型, 但当对系统很复杂, 或者系统未知, 系统信息量很少时, 建立精确的数学模型很困难时, 神经网络的非线性映射能力则表现出优势, 因为它不需要对系统进行透彻的了解, 但是同时能达到输入与输出的映射关系, 这就大大简化设计的难度。

(4) 高度并行性

并行性具有一定的争议性。承认具有并行性理由: 神经网络是根据人的大脑而抽象出来的数学模型, 由于人可以同时做一些事, 所以从功能的模拟角度上看, 神经网络也应具备很强的并行性。

多少年以来, 人们从医学、生物学、生理学、哲学、信息学、计算机科学、认知学、组织协同学等各个角度企图认识并解答上述问题。在寻找上述问题答案的研究过程中, 这些年来逐渐形成了一个新兴的多学科交叉技术领域, 称之为“神经网络”。神经网络的研究涉及

众多学科领域, 这些领域互相结合、相互渗透并相互推动。不同领域的科学家又从各自学科的兴趣与特色出发, 提出不同的问题, 从不同的角度进行研究。

当对系统对于设计人员来说, 很透彻或者很清楚时, 则一般利用数值分析, 偏微分方程等数学工具建立精确的数学模型, 但当对系统很复杂, 或者系统未知, 系统信息量很少时, 建立精确的数学模型很困难时, 神经网络的非线性映射能力则表现出优势, 因为它不需要对系统进行透彻的了解, 但是同时能达到输入与输出的映射关系, 这就大大简化设计的难度。

下面将人工神经网络与通用的计算机工作特点来对比一下:

若从速度的角度出发, 人脑神经元之间传递信息的速度要远低于计算机, 前者为毫秒量级, 而后者的频率往往可达几百兆赫。但是, 由于人脑是一个大规模并行与串行组合处理系统, 因而, 在许多问题上可以做出快速判断、决策和处理, 其速度则远高于串行结构的普通计算机。人工神经网络的基本结构模仿人脑, 具有并行处理特征, 可以大大提高工作速度。

人脑存贮信息的特点为利用突触效能的变化来调整存贮内容, 也即信息存贮在神经元之间连接强度的分布上, 存贮区与计算机区合为一体。虽然人脑每日有大量神经细胞死亡 (平均每小时约一千个), 但不影响大脑的正常思维活动。

## 4.1 神经网络结构

人们熟悉的神经网络有几十种, 由于 BP 神经网络实质上实现了一个从输入到输出的映射功能, 而数学理论已证明它具有实现任何复杂非线性映射的功能。这使得它特别适合于求解内部机制复杂的问题; BP 网络能通过学习带有正确答案的实例集自动提取“合理的”求解规则, 即具有自学习能力; BP 网络具有一定的推广、概括能力。因此 BP 神经网络成为一种广泛应用的神经网络。

### 4.1.1 BP 神经网络结构

图 4-1 显示了多层感知器 (MLP) 神经网络<sup>[9]</sup>的架构, 该神经网络由三层组成: 输入层, 输出层和隐藏层。每层由多个神经元组成, 这些神经元与相邻层中的神经元相连。让  $x = \{x_1, x_2, \dots, x_p\}$  表示网络的候选输入变量, 隐藏层具有表示为的  $q$  个节点  $h = \{h_1, h_2, \dots, h_q\}$ , 权重  $w_{ij} (i = 1, 2, \dots, p; j = 1, 2, \dots, q)$  表示输入变量  $x_i$  和第  $j$  个隐藏神经元  $j$ th 之间的输入权重。隐藏层的第  $j$  个神经元的输出信号  $O_j^h$  由下式给出:

$$O_j^h(x) = f\left(\left(\sum_{i=1}^p w_{ij}x_i\right) + b_j\right) \quad j = 1, 2, \dots, q \quad (4-1)$$

其中  $b_j^h$  是隐藏层的第  $j$  个神经元的偏差,  $f$  表示隐藏层的激活功能。

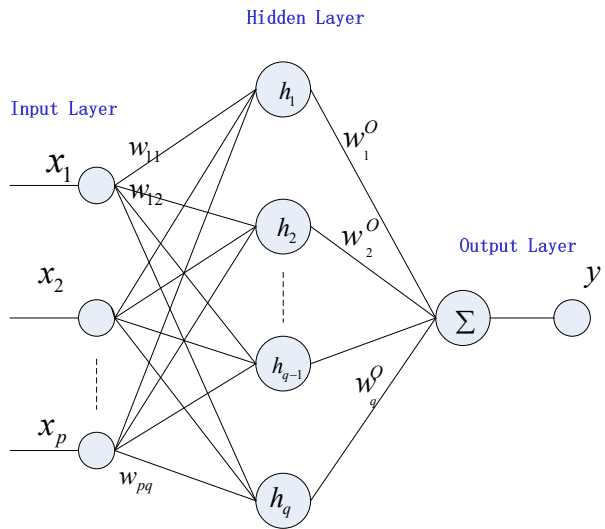


图 4-1. 人工神经网络的体系结构

此外, 网络的输出  $y$  代表响应变量, 由下式给出:

$$y = g\left(\sum_{j=1}^q w_j^o O_j^h(x)\right) + \theta \quad (4-2)$$

其中  $\theta$  是输出层的偏差,  $g$  表示输出层的激活功能, 并且  $w_j^o (j = 1, 2, \dots, q)$  表示隐藏层和输出层之间的第  $j$  个输出权重。将等式 (4-1) 代入等式 (4-2),  $y$  可以表述为:

$$y = g\left(\sum_{j=1}^q w_j^o f\left(\sum_{i=1}^p w_{ij} x_i\right) + \theta\right) + \theta \quad (4-3)$$

### 4.1.2 非负阈值算法

Breiman<sup>[10]</sup>提出了原始 NNG (nonnegative garrote) 来解决线性问题的子集回归, 其中真实模型具有以下形式:

$$y = \sum_{i=1}^p \beta_i x_i + \varepsilon \quad (4-4)$$

$x = \{x_1, x_2, \dots, x_p\}$  和  $y$  分别是输入和输出变量,  $\beta = \{\beta_1, \beta_2, \dots, \beta_p\}$  是变量系数,  $\varepsilon$  是随机误差。假设  $X \in \mathbb{R}^n \times \mathbb{R}^p$  是输入数据矩阵, 其中每列代表候选解释变量,  $Y \in \mathbb{R}^n$  是响应变量的向量。让  $\hat{\beta}_i$  表示普通最小二乘 (OLS) 估计; NNG 可以表述为:

$$c^*(s) = \operatorname{argmin}\left\{\sum_{k=1}^n \left(y_k - \sum_{i=1}^p c_i \hat{\beta}_i x_{ik}\right)^2\right\} \quad (4-5)$$

subject to:

$$c_i \geq 0, \quad \sum_{i=1}^p c_i \leq s$$

其中  $s$  是 garrote 参数。一组新的预测器系数,  $\tilde{\beta}_i$ , 可以通过以下方式获得:

$$\tilde{\beta}_i = c_i \hat{\beta}_i, \quad i = 1, 2, \dots, p \quad (4-6)$$

然后  $y$  的预测方程可以表示为:

$$\tilde{y} = \sum_{i=1}^p \tilde{\beta}_i x_i \quad (4-7)$$

如果 garrote 参数  $s \geq p$ , 约束  $\sum_{i=1}^p c_i < s$  无效, 这意味着  $\tilde{\beta}_i = \hat{\beta}_i, \forall i$ 。当 garrote 参数  $s$  减少时, 更多的幅度系数  $c_i$  将倾向于为零。当  $s$  从 “ $p$ ” 减小到零时, 出现解决方案路径。预测误差可以定义为在模型处理期间未使用的未来实例的  $y$  预测中的均方误差。



### 4.1.3 整合 NNG 和 ANN

本文提出了一种输入变量选择方法, 将 NNG 引入神经网络来解决非线性问题。与原始 NNG 类似, 新方法可分为两个阶段。在第一阶段, 它获得了一个训练有素的 MLP 网络, 其中激活函数  $f, g$ , 偏差  $b_j^h, b^o$ , 并且等式 (4-3) 的输出权重  $w_j^o$  是已知的并且是固定的<sup>[11]</sup>。在第二阶段, 可以通过在输入变量前面添加新的幅度系数  $c = \{c_1, c_2, \dots, c_p\}$  来重新计算等式 (4-3):

$$y = g\left(\sum_{j=1}^q w_j^o f\left(\sum_{i=1}^p c_i w_{ij} x_j\right) + b\right) + B \quad (4-8)$$

假设  $\widehat{w}_{ij} (i = 1, 2, \dots, p; j = 1, 2, \dots, q)$  是训练有素的神经网络的初始输入权重; NNG-ANN 的表述为:

$$c^*(s) = \operatorname{argmin}\left\{\sum_{k=1}^n (y_k - g\left(\sum_{j=1}^q w_j^o f\left(\sum_{i=1}^p c_i w_{ij} x_j\right) + b\right) + B)\right\}^2 \quad (4-9)$$

subject to:

$$c_i \geq 0, \sum_{i=1}^p c_i \leq s$$

显然, 等式 (4-9) 是非线性约束二次问题。本文使用信赖域反射优化算法, 这是提出的内反射牛顿法的子空间信赖域方法。没有稳定性问题, 并且随着输入变量的数量增加, 计算时间仅适度增加。

一组新的输入权重  $\widetilde{w}_{ij}$  可以通过以下方式获得:

$$\widetilde{w}_{ij} = c_i \widehat{w}_{ij}, i = 1, 2, \dots, p; j = 1, 2, \dots, q \quad (4-10)$$

如果  $c_i = 0$ , 则在新神经网络中消除输入变量  $i$ 。因此,  $y$  的预测模型通过以下方式获得:

$$\tilde{y} = g\left(\sum_{j=1}^q w_j^o f\left(\sum_{i=1}^p \widetilde{w}_{ij} x_j\right) + b\right) + B \quad (4-11)$$

### 4.1.4 评估标准

本文使用贝叶斯信息准则(BIC)作为有限模型集中模型选择的评价标准。BIC 由 Schwarz 和 Gideon 提出, 它被用来衡量模型的准确性和复杂性之间的权衡。它被给予:

$$\text{BIC} = n \cdot \log_e(\text{MSE}) + k \cdot \log(n) \quad (4-12)$$

其中  $n$  表示观察数,  $k$  表示所选变量的数量。RMSE 由下面给出:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum (y_j - \tilde{y}_j)^2} \quad (4-13)$$

其中  $y$  是实际数据,  $\tilde{y}_j$  是通过等式 (4-11) 计算的预测值。

### 4.1.5 s 的选择

对于给定的 garrote 参数  $s$ , 可以通过求解方程 (4-9) 的非线性约束二次问题来获得模型。因此, NNG-ANN 估计的准确度和模型大小在很大程度上取决于 garrote 参数  $s$  的选择。可以使用  $v$ -折叠交叉验证确定最佳  $s$ , 其程序如下所示。

数据集  $\mathcal{L} = \{X, Y\}$  被分成  $v$  个子集:  $\mathcal{L} = \{\mathcal{L}_1, \mathcal{L}_2, \dots, \mathcal{L}_v\}$ 。设  $\overline{\mathcal{L}}_v = \mathcal{L} - \mathcal{L}_v$ ; 训练集为  $\overline{\mathcal{L}}_v$ , 验证集为  $\mathcal{L}_v$ 。NNG-ANN 使用数据集  $\overline{\mathcal{L}}_v$  来建立适当的模型, 并使用验证数据集  $\mathcal{L}_v$  来计算验证值  $y^{(v)}$ 。 $v$  的预测模型误差可以通过以下方式获得:

$$\text{CV}^{(v)}(s) = n \cdot \log_e\left(\frac{1}{n} \sum_{j=1}^n (y_j - \tilde{y}_j^{(v)})^2\right) + k \cdot \log(n) \quad (4-14)$$



和  $V$  折叠交叉验证平均误差是:

$$\overline{CV}(s) = \frac{1}{V} \sum_{v=1}^V CV^{(v)}(s) \quad (4-15)$$

$s$  的最优解可以通过最小化整体  $v$  折交叉验证来确定:

$$s^* = \operatorname{argmin}_{s_{lb} < s < s_{ub}} \left( \frac{1}{V} \sum_{v=1}^V CV^{(v)}(s) \right) \quad (4-16)$$

其中  $[s_{lb}, s_{ub}]$  表示  $s$  的有界域。  $s$  的下限通常设置为接近零的值。 在论文中将其设置为 0.01, 并将  $s$  的上限设置为  $p$ , 这是输入变量的总数。

## 4.2 MATLAB 程序的实现

MATLAB 是美国 MathWorks 公司出品的商业数学软件, 用于算法开发、数据可视化、数据分析以及数值计算的高级技术计算语言和交互式环境。它将数值分析、矩阵计算、科学数据可视化以及非线性动态系统的建模和仿真等诸多强大功能集成在一个易于使用的视窗环境中, 为科学研究、工程设计以及必须进行有效数值计算的众多科学领域提供了一种全面的解决方案。

### 4.2.1 数据预处理的编程实现

```
>> clear all;
load('A.mat');
corrcoef(A);
if ~isempty(varargin) && isnumeric(varargin{1})
    y = varargin{1};
    varargin(1) = [];
    % Convert two inputs to equivalent single input.
    if numel(x)~=numel(y)
        error(message('MATLAB:corrcoef:XYmismatch'));
    end
    x = [x(:) y(:)];
elseif ~ismatrix(x)
    error(message('MATLAB:corrcoef:InputDim'));
end

% Quickly dispose of most common case.
if nargout<=1 && isempty(varargin)
    r = corral(x);
    return;
end

if ~isreal(x) && nargout>1
    error(message('MATLAB:corrcoef:ComplexInputs'));
end
```

```
% Deal with empty inputs
[n,m] = size(x);
if isempty(x)
    if n <= 1
        % Zero observations with m variables results in an m x m NaN,
        % unless m == 0 where we return a scalar NaN by convention.
        % Note, an empty row vector is treated as zero observations with
        % one variable, consistent with treatment of non-empty row vectors.
        r = NaN(max(m,1), class(x));
    else
        % n > 1 observations with no variables returns an empty matrix of
        % correlations.
        r = ones(0, class(x));
    end
    p = r; rlo = r; rup = r;
    return;
end
% Treat all vectors like column vectors.
if isvector(x)
    x = x(:);
    [~, m] = size(x);
end
% Process parameter name/value inputs.
[alpha,userrows,emsg] = getparams(varargin{:});
if ~isempty(emsg)
    error(message(emsg));
end
% Compute correlations.
t = isnan(x);
removemissing = any(t,:);
if isequal(userrows, 'all') || ~removemissing
    [r, n] = correl(x);
elseif isequal(userrows, 'complete')
    % Remove observations with missing values.
    x = x(~any(t,2),:);
    [r, n] = correl(x);
else
    % Compute correlation for each pair.
    r = zeros(m,class(x));
    n = zeros(m,class(x));
    jk = 1:2;
    for j = 1:m
        jk(1) = j;
```

```

    for k=1:j-1
        jk(2) = k;
        tjk = ~any(t(:,jk),2);
        n(j,k) = sum(tjk);
        if n(j,k)<=1
            r(j,k) = NaN;
        else
            rjk = correl(x(tjk,jk));
            r(j,k) = rjk(1,2);
        end
    end
    end
    tj = ~t(:,j);
    n(j,j) = sum(tj);
    if n(j,j)<=1
        r(j,j) = NaN;
    else
        r(j,j) = correl(x(tj,j));
    end
    end
    end
    r = r + tril(r,-1)';
    n = n + tril(n,-1)';
end
% Compute p-value if requested.
if nargout>=2
    % Operate on half of symmetric matrix.
    lowerhalf = (tril(ones(m),-1)>0);
    rv = r(lowerhalf);
    if length(n)>1
        nv = n(lowerhalf);
    else
        nv = n;
    end
    % Tstat = +/-Inf and p = 0 if abs(r) == 1, NaN if r == NaN.
    Tstat = rv .* sqrt((nv-2) ./ (1 - rv.^2));
    p = zeros(m,class(x));
    p(lowerhalf) = 2*tpvalue(-abs(Tstat),nv-2);
    p = p + p' + diag(diag(r)); % Preserve NaNs on diag.
    % Compute confidence bound if requested.
    if nargout>=3
        % Confidence bounds are degenerate if abs(r) = 1, NaN if r = NaN.
        z = 0.5 * log((1+rv)/(1-rv));
        zalpha = NaN(size(nv),class(x));
        if any(nv>3)
            zalpha(nv>3) = (-erfinv(alpha - 1)) .* sqrt(2) ./ sqrt(nv(nv>3)-3);
        end
    end
end

```

```

        end
        rlo = zeros(m,class(x));
        rlo(lowerhalf) = tanh(z-zalpha);
        rlo = rlo + rlo' + diag(diag(r)); % Preserve NaNs on diag.
        rup = zeros(m,class(x));
        rup(lowerhalf) = tanh(z+zalpha);
        rup = rup + rup' + diag(diag(r)); % Preserve NaNs on diag.
    end
end
function [r,n] = correl(x)
%CORREL Compute correlation matrix without error checking.
[n,m] = size(x);
r = cov(x);
d = sqrt(diag(r)); % sqrt first to avoid under/overflow
r = r ./ d ./ d'; % r = r ./ d*d';
% Fix up possible round-off problems, while preserving NaN: put exact 1 on the
% diagonal, and limit off-diag to [-1,1].
r = (r+r')/2;
t = find(abs(r) > 1); r(t) = r(t)./abs(r(t));
r(1:m+1:end) = sign(diag(r));
function p = tpvalue(x,v)
%TPVALUE Compute p-value for t statistic.
normcutoff = 1e7;
if length(x)~=1 && length(v)==1
    v = repmat(v,size(x));
end
% Initialize P.
p = NaN(size(x));
nans = (isnan(x) | ~(0<v)); % v == NaN ==> (0<v) == false
% First compute F(-|x|).
%
% Cauchy distribution. See Devroye pages 29 and 450.
cauchy = (v == 1);
p(cauchy) = .5 + atan(x(cauchy))/pi;
% Normal Approximation.
normal = (v > normcutoff);
p(normal) = 0.5 * erfc(-x(normal) ./ sqrt(2));
% See Abramowitz and Stegun, formulas 26.5.27 and 26.7.1.
gen = ~(cauchy | normal | nans);
p(gen) = betainc(v(gen) ./ (v(gen) + x(gen).^2), v(gen)/2, 0.5)/2;
% Adjust for x>0. Right now p<0.5, so this is numerically safe.
reflect = gen & (x > 0);
p(reflect) = 1 - p(reflect);
% Make the result exact for the median.

```

```

p(x == 0 & ~nans) = 0.5;
function [alpha,userows,estruc] = getparams(varargin)
%GETPARAMS Process input parameters for CORRCOEF.
alpha = 0.05;
userows = 'all';
estruc = [];
while ~isempty(varargin)
    if length(varargin)==1
        estruc = ('MATLAB:corrcoef:unmatchedPVPair');
        return
    end
    pname = varargin{1};
    if ~ischar(pname)
        estruc = ('MATLAB:corrcoef:invalidArgName');
        return
    end
    pval = varargin{2};
    j = find(strcmp(pname,{'alpha' 'rows'}));
    if isempty(j)
        estruc = ('MATLAB:corrcoef:invalidArgName');
        return
    end
    if j==1
        alpha = pval;
    else
        userows = pval;
    end
    varargin(1:2) = [];
end
% Check for valid inputs.
if ~isnumeric(alpha) || ~isscalar(alpha) || alpha<=0 || alpha>=1
    estruc = ('MATLAB:corrcoef:invalidAlpha');
    return
end
oktypes = {'all' 'complete' 'pairwise'};
if isempty(userows) || ~ischar(userows)
    i = [];
else
    i = strcmpi(userows, oktypes);
end
if ~any(i)
    estruc = ('MATLAB:corrcoef:invalidRowChoice');
    return
end
    
```

```

userrows = oktypes{i};
b=find(abs(ans(:,53))>0.3);
[m,n]=size(b);
newd=[];
c=[A(:,b)];
    
```

#### 4.2.2 NNG-BP 的编程实现

```

>> clear all;
load('x.mat')
load('y.mat')
function [s,shrcf,pmse]=NNGMain(x,y)
[m,p]=size(x);
m1=floor(m*0.5);
x1=x(1:m1,:);
y1=y(1:m1);
x2=x(m1:m,:);
y2=y(m1:m);
[net,rmse]=newtrain(x1,y1);
[n,p]=size(x1);
nfolds= 5;
s=linspace(1,10,20);
press=zeros(length(s),1);
b=[];
for i=1:length(s)
    press(i)=newnng_kcv(x1,y1,s(i),net);
end
[a,b]=min(press);
s=s(b);
[beta, shrcf ,nnet]=nng_ann(x1,y1,s,net)
yb=sim(nnet,x2');
e=y2-yb';
rmse=mse(e);
corrcoef(y2,yb);
plot (y2);
hold;
plot(yb,'r');

function [x1,y1,x2,y2]=datadiv(x,y)
[n,p]=size(x);
ind=[];
for i=1:n
    ind=[ind i];
end
m=floor(n/5);
    
```



```
testx=(4*m+1:5*m);  
x2=x(testx,:);  
y2=y(ind(testx),:);
```

```
newx=x;  
newx(4*m+1:5*m,:)=[];  
newy=y;  
newy(4*m+1:5*m,:)=[];
```

```
x1=newx;  
y1=newy;
```

```
function [net,rmse]=newtrain(x,y)  
[x1,y1,x2,y2]=datadiv(x,y);  
rmse=100;  
hiddennode=5;  
net = newff(x1',y1',hiddennode);  
net=train(net,x1',y1');  
yb=sim(net,x2');  
e=y2-yb';  
rmse=mse(e);  
for i=1:5
```

```
    nnet = newff(x1',y1',hiddennode);  
    nnet=train(nnet,x1',y1');
```

```
    yb=sim(nnet,x2');  
    e=y2-yb';
```

```
    newrmse=mse(e);  
    if newrmse<rmse  
        rmse=newrmse;  
        net=nnet;  
    end
```

```
end
```

```
function amse=newnng_kcv(x,y,s,net)  
[n,p]=size(x);  
ind=[];  
amse=0;  
nfolde=5;  
for i=1:n  
    ind=[ind i];
```

```

end
m=floor(n/nfolds);

for i=1:nfolds %for i=1:nfolds    nfolds =5

    newx=x;
    newy=y;
    vx=newx((i-1)*m+1:i*m,:);
    vy=newy((i-1)*m+1:i*m,:);
    newx((i-1)*m+1:i*m,:)=[];
    newy((i-1)*m+1:i*m,:)=[];
    [beta,shrcf ,nnet]=nng_ann(newx,newy,s,net);
    yb=sim(nnet,vx');
    perror(i)=mse(vy-yb');
end
amse= mean(perror);
a=find(abs(shrcf)>1e-5)
t=size(a,1);
BIC=2*t+n*log(amse/n);
amse=amse+BIC;
    
```

### 4.3 基于神经网络预测模型仿真实例

本文以黄台电厂 9 号机组的脱硫指标参数的数据为研究对象，数据集由 52 个候选输入变量和一个目标输出变量烟气 SO<sub>2</sub> 浓度组成。所有候选变量列于表 4-1 中。时间跨度 2018 年 7 月 1 日到 2018 年 7 月 10 日，时间间隔为 5min，共 2676 组。将样本的前八天作为训练集，最后两天的数据作为验证集，构建预测模型，并通过 matlab 工具实现基于神经网络算法对电厂脱硫控制系统净烟气 SO<sub>2</sub> 浓度的预测。

有 53 个输入变量，首先对数据进行预处理，分析数据的相关性，相关性求取方法： $\text{corrcoef}(A)$ ，即返回从矩阵 A 形成的一个相关系数矩阵，若 A 是一个  $m \times n$  的矩阵，那么得到的相关系数矩阵 ans 就是一个  $n \times n$  的对称矩阵，ans 中的第 i 行第 j 列的元素表示的就是 A 第 i 列和第 j 列的相关系数。所有输入变量数据相关性如表 4-1 所示。

$$\text{corrcoef}(ij) = \frac{\text{cov}(ij)}{\sqrt{\text{cov}(i) \cdot \text{cov}(j)}} \quad (4-17)$$

表 4-1 输入变量的相关性分析表

输入变量	相关性	输入变量	相关性
发电机功率	0.386033283	9号机组脱硫效率中间值	-0.183940649
#9炉原烟气SO2浓度	0.528331369	#9-2吸收塔出口温度1	0.167989947
#9-1吸收塔出口烟气SO2浓度	0.777421195	#9-2吸收塔出口温度2	0.135694197
#9吸收塔石膏浆液PH值	0.173663444	#9-2吸收塔出口压力	0.377781995
石灰石浆液去#9吸收塔流量	0.519010954	#9-1吸收塔出口烟气O2浓度	-0.276712916
#9-2吸收塔PH值1	-0.439259386	#9吸收塔石膏浆液密度	-0.01849067
#9-2吸收塔供浆流量	0.263436328	#9 吸收塔液位计算值	-0.214665548
#1机#1循环浆泵电流	-0.218303124	#9-2吸收塔液位1	-0.088706513
#1机#2循环浆泵电流	-0.063402028	#9-2吸收塔液位2	-0.124991662
#1机#3循环浆泵电流	-0.195058963	#9-2吸收塔液位3	-0.142561341
#1机#4循环浆泵电流	0.006734046	石灰石浆液密度	0.002305334
#9循环浆液泵E B相电流	0.001947611	吸收塔石膏排出泵1入口压力	-0.067476199
#9循环浆液泵G B相电流	-0.041596627	吸收塔石膏排出泵2电流	0.091325816
#9吸收塔入口烟气温度	0.271838357	吸收塔石膏排出泵1转速反馈	-0.044383181
总风量	0.371246588	吸收塔石膏排出泵1电流	-0.045030103
总煤量	0.480572605	吸收塔石膏排出泵2转速反馈	0.096818219
#9吸收塔出口烟气温度1	0.178149963	吸收塔石膏排出泵2入口压力	0.029678334
#9吸收塔出口烟气温度2	0.173817322	#9吸收塔石膏浆液流量	-0.120995441
#9吸收塔出口烟气温度3	0.165674696	石膏浆液排出泵A电流	0.145857469
9号机组脱硫效率	-0.127734248	石膏浆液泵A转速反馈	-0.077843131
#9炉烟囱入口净烟气流量	0.236413339	#9石膏浆液旋流器入口压力	0.252420174
#9炉原烟气流量	0.073357887	石膏浆液泵至#9吸收塔母管压力	-0.047091524
#9炉烟囱入口净烟气压力	-0.19771103	石膏浆液箱液位	-0.064845518
#9炉原烟气烟尘浓度	-4.53E-15	水环式真空泵A电流	-0.233548315
#9炉原烟气NOx浓度	-0.318632551	真空皮带脱水机A转速反馈	0.191006335
#9炉原烟气O2浓度	-0.277244127	皮带脱水机A石膏滤饼厚度	0.16985829

### 4.3.1 BP 神经网络算法对烟气 SO<sub>2</sub> 浓度预测

通过复杂系统建模, 利用 BP 神经网络算法对电厂脱硫控制系统净烟气 SO<sub>2</sub> 浓度进行预测。神经网络具有一个具有双曲正切激活函数的隐藏层和具有线性激活函数的输出层。该体系结构为 52-4-10-1, 意味着它有 52 个输入节点, 两个隐含层, 分别有 4 个隐藏节点和 10 个隐藏节点, 并有一个输出节点。预测结果如图 4-3 所示。图中横坐标表示时间 (以 5min 为间隔), 纵坐标表示烟气 SO<sub>2</sub> 浓度值。

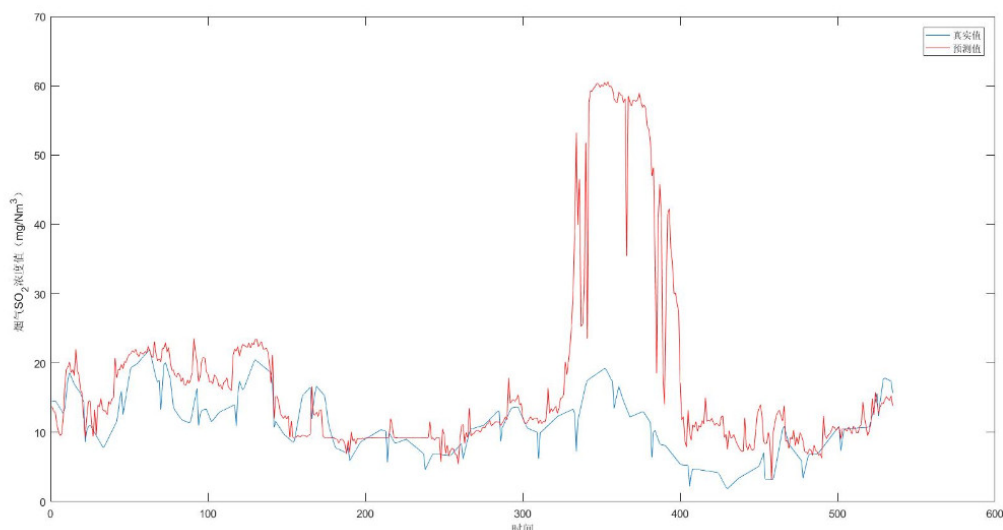


图 4-3 BP 神经网络算法对烟气 SO<sub>2</sub> 浓度的预测曲线图

### 4.3.2 NNG-BP 神经网络算法对烟气 SO<sub>2</sub> 浓度的预测

通过复杂系统建模, 利用 NNG-BP 神经网络算法对电厂脱硫控制系统净烟气 SO<sub>2</sub> 浓度的预测。变量压缩后剩余 9 个输入变量。剩余的 9 个标量如表 4-2 所示。神经网络具有一个具有双曲正切激活函数的隐藏层和具有线性激活函数的输出层。通过在 5 折交叉验证中训练神经网络和训练数据来确定神经网络的结构。该体系结构为 52-4-10-1, 意味着它有 52 个输入节点, 两个隐含层, 分别有 4 个隐藏节点和 10 个隐藏节点, 并有一个输出节点。预测结果如图 4-4 所示。图中横坐标表示时间 (以 5min 为间隔), 纵坐标表示烟气 SO<sub>2</sub> 浓度值。

变量选择的目的是要找到对输出变量影响较大的辅助变量, 还希望由辅助变量对输出变量可能出现的情况进行预测。当把预测作为我们的主要任务时, 常把预测误差平方和作为衡量预测优劣的指标, 其值越小则预测精度越高。最佳的 s 值就是使误差平方和取得最小的 s 值。然而, 在实践中由于样本数据的限制, 误差平方和的值并不是可直接观测到的, 我们必须对数据进行重用, 对数据进行多次切分来得到较好的估计。V 折交叉验证法由于其简洁性和普遍性被认为是一种行之有效的办法。

V 折交叉验证法首先是把数据集平均分为 V 份, 每次从 V 份数据集中拿出一份数据集作为验证集, 剩下的 V-1 份数据集作为训练集, 重复进行 V 次, 最后平均 V 次的结果作为最后泛化误差的估计。通常 V 的取值为 5 到 10 时能得到较好的结果, 当 V 取值太大时, 方差也会随之增大; 当 V 取值较小时由于参与训练的样本数据减少会导致预测误差的增大。

表 4-2 NNG-BP 算法变量选择结果

输入变量	相关系数
发电机功率	0.386033283
#9炉原烟气SO2浓度	0.528331369
#9-1吸收塔出口烟气SO2浓度	0.777421195
石灰石浆液去#9吸收塔流量	0.519010954
#9-2吸收塔PH值1	-0.439259386
总风量	0.371246588
总煤量	0.480572605
#9炉原烟气NO <sub>x</sub> 浓度	-0.318632551

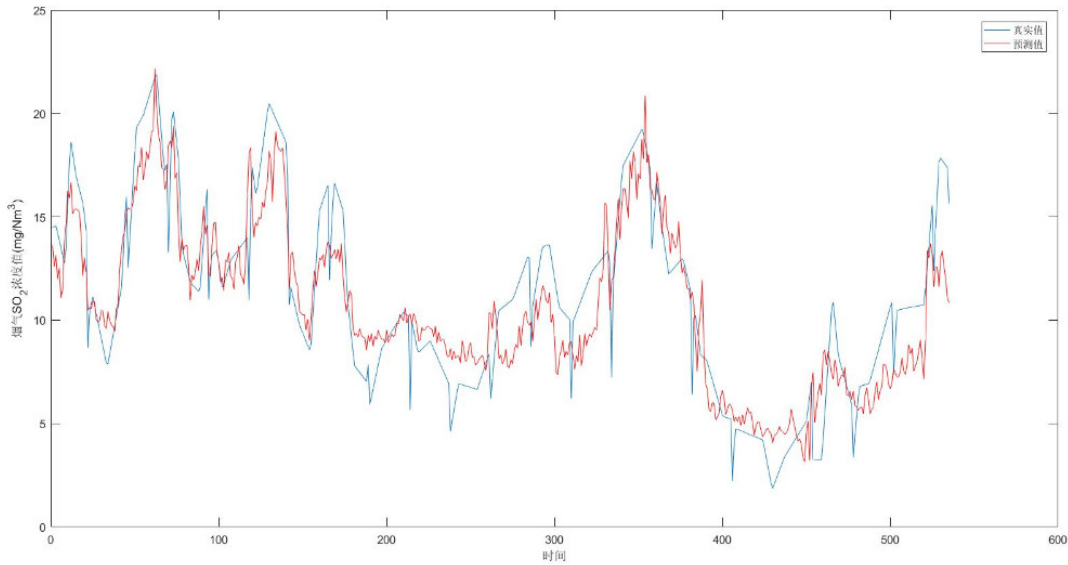


图 4-4 NNG-BP 神经网络算法对烟气 SO<sub>2</sub> 浓度的预测曲线图

表 4-3 为使用 BP 神经网络算法和 NNG-BP 算法对电厂脱硫控制系统净烟气 SO<sub>2</sub> 浓度的预测进行预测时的均方根误差 (Root Mean Squared Error, RMSE), 由表可知, 应用 NNG-BP 算法对电厂脱硫控制系统净烟气 SO<sub>2</sub> 浓度进行预测比 BP 神经网络算法的预测精度高。

表 4-3 预测结果对比

预测方法	NNG-BP 神经网络算法	BP 神经网络算法
RMSE	3.9645	187

## 5 小结

本文以 BP 神经网络和 NNG-BP 神经网络为基础, 建立了某电厂 9 号机组短时烟气 SO<sub>2</sub> 浓度预测模型。并利用济南黄台电厂 9 号机组脱硫指标参数数据为样本进行仿真预测。并将预测结果以均方根误差(RMSE)为标准进行了比较。结果表明: NNG-BP 神经网络能较准确地预测出烟气 SO<sub>2</sub> 浓度值。我们可以建立较准确地模型, 进行反馈控制, 具有较好的实用性和经济性。本研究所建立的模型及研究结果, 可以为石灰石-石膏法脱硫系统的优化提供参考。

## 致谢

本研究项目由李至瑞提出, 报告中数据的搜集由李东峰和李至瑞共同完成, 模型构建和程序设计调试由李东峰完成, 在数据采集过程中得到黄台电厂任主任的大力支持, 得到黄台电厂徐静工程师和陈明县工程师的大力协助, 建模过程得到山东省实验中学石磊老师的指导, 在此对给与大力帮助的各位专家和老师表示诚挚的谢意!



## 参考文献

- [1]. 王志轩, et al., *中国煤电清洁发展现状及展望*. 电力科技与环保, 2018(1): p. 1-8.
- [2]. 赵雪, 程茜, and 侯俊先, *脱硫脱硝行业 2017 年发展综述*. 中国环保产业, 2018(07): p. 10-24.
- [3]. 靳晓悦 and 刘雅举, *湿法烟气脱硫技术综述*. 当代化工研究, 2017(6): p. 5-6.
- [4]. 刘春辉, *火力发电厂烟气脱硫控制系统设计及应用*. 2012, 华北电力大学.
- [5]. 韩晶, *大数据服务若干关键技术研究*. 2013, 北京邮电大学.
- [6]. Frank, P.M.J.E.J.o.C., *Analytical and Qualitative Model-based Fault Diagnosis – A Survey and Some New Results*. 1996. 2(1): p. 6-28.
- [7]. 黄晓清, *基于核方法的复杂工业过程建模研究*. 2008, 大连理工大学.
- [8]. 张彦军, *石灰石-石膏湿法烟气脱硫吸收塔参数优化研究*. 2018, 河北科技大学.
- [9]. Haykin, S.J.N.N.A.C.F., *Neural Networks: A Comprehensive Foundation*. 1994: p. 71-80.
- [10]. Breiman, L., *Better Subset Regression Using the Nonnegative Garrote*. 1995: American Society for Quality Control and American Statistical Association. 373-384.
- [11]. Coleman, T.F. and Y.J.M.P. Li, *On the convergence of interior-reflective Newton methods for nonlinear minimization subject to bounds*. 1994. 67(1-3): p. 189-224.