



(12)发明专利申请

(10)申请公布号 CN 109443419 A  
(43)申请公布日 2019.03.08

(21)申请号 201811011319.9

(22)申请日 2018.08.31

(71)申请人 广州市世科高新技术有限公司  
地址 510540 广东省广州市白云区太和镇  
北太路1633号广州民营科技园内白云  
电气科技大厦第一层105室  
申请人 广州市扬新技术研究有限责任公司

(72)发明人 银星茜 盛健 彭曼 王亚东  
张俊强 廖权保 黄伟峰

(74)专利代理机构 广州知友专利商标代理有限  
公司 44104  
代理人 周克佑

(51)Int.Cl.  
G01D 21/02(2006.01)  
G06N 20/00(2019.01)

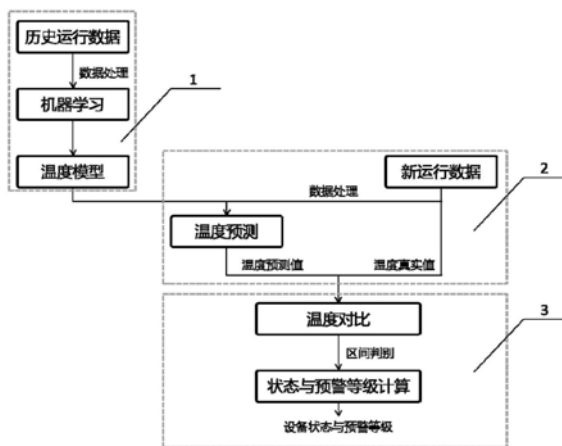
权利要求书2页 说明书6页 附图2页

(54)发明名称

一种基于机器学习的整流器在线监测方法

(57)摘要

一种基于机器学习的整流器在线监测方法，包括步骤：采用整流器历史数据进行建模，构建整流器正常运行状态下的温度模型，含“数据处理”与“机器学习”环节；整流器状态监测与预测：先使用“数据处理”对新数据进行处理并构建相同的特征变量，再将处理后的数据传递给整流器温度模型，由温度模型计算出整流器温度的预测值；整流器状态/预警等级计算：实时将整流器温度预测值与真实值进行对比，采用区间判断与动态监测判断整流器运行状态，实时给出整流器预警信息和状态信息，预警信息分为三个区间：第一个区间预警级别为0，表示设备正常运行；第二个区间预警级别为1-11，表示设备状态需要“注意”；第二个区间预警级别为12，表示设备状态“异常”。



1. 一种基于机器学习的整流器在线监测方法,其特征是包括以下步骤:

S1,数据采集:

使用电流传感器、温度传感器和湿度传感器采集整流器运行时的电流、设备温度、环境温度以及环境湿度的数据;

S2,数据处理:对电流数据进行动态平滑处理,方法包括移动平滑法、非对称局部加权回归散点平滑法、移动窗口拟或多项式平滑方法;

S3,机器学习:采用整流器历史运行数据以岭回归法、Lasso regression、随机森林、决策树、梯度提升决策树、神经网络或RNN算法进行建模;

S4,模型预测:

针对整流器运行产生的新数据,在进行数据处理后将数据传到整流器温度模型,预测整流器在当前电流、环境温度与环境湿度下的温度,通过对比整流器的真实温度与真实温度,实时甄别整流器温度是否按历史正常规律变化。

2. 根据权利要求1所述的基于机器学习的整流器在线监测方法,其特征是:所述的S2数据处理中,采用移动平滑法对电流数据进行动态平滑处理包括以下子步骤:

S2-1,公式(1):

$$\bar{I}_t = \frac{1}{N}(I_{t-1} + I_{t-2} + \dots + I_{t-N}) \dots\dots\dots (1);$$

其中, $\bar{I}_t$ 为当前时刻平滑后的电流值, $I_t$ 为当前时刻的实际电流值,(1,2,...,N)为当前时刻向后平滑多少个时间点,如 $I_{t-1}$ 为t-1时刻的电流值, $I_{t-2}$ 为t-2时刻的电流值, $I_{t-N}$ 为t-N时刻的电流值;

S2-2,使用电流数据构造新衍生变量:

针对电流数据构造其衍生变量,包含电流平方值、电流立方值、平滑电流平方值、平滑电流立方值、公式(2)的一段时间内的电流平方累计值、公式(3)的一段时间内的电流立方累计值、公式(4)的电流最大值、公式(5)的电流最小值和公式(6)的标准差;

$$\text{sum}(I_t^2) = \sum I_{t-1}^2 + I_{t-2}^2 + \dots + I_{t-N}^2 \dots\dots\dots (2);$$

$$\text{sum}(I_t^3) = \sum I_{t-1}^3 + I_{t-2}^3 + \dots + I_{t-N}^3 \dots\dots\dots (3);$$

$$I_{t\_max} = \max(I_{t-1}, I_{t-2}, \dots, I_{t-N}) \dots\dots\dots (4);$$

$$I_{t\_min} = \min(I_{t-1}, I_{t-2}, \dots, I_{t-N}) \dots\dots\dots (5);$$

$$I_{t\_std} = \text{std}(I_{t-1}, I_{t-2}, \dots, I_{t-N}) \dots\dots\dots (6);$$

在公式中, $I_t^2$ 为t时刻的电流平方值, $I_{t-1}^2$ 为t-1时刻的电流平方值, $I_{t-2}^2$ 为t-2时刻的电流平方值, $I_{t-N}^2$ 为t-N时刻的电流平方值; $I_t^3$ 为t时刻的电流立方值, $I_{t-1}^3$ 为t-1时刻的电流立方值, $I_{t-2}^3$ 为t-2时刻的电流立方值, $I_{t-N}^3$ 为t-N时刻的电流立方值; $\text{sum}(I_t^2)$ 为[t-N,t-1]时间段内电流平方的累计值, $\text{sum}(I_t^3)$ 为[t-N,t-1]时间段内电流立方的累计值, $I_{t\_max}$ 为[t-N,t-1]时间段内电流的最大值, $I_{t\_min}$ 为[t-N,t-1]时间段内电流的最小值, $I_{t\_std}$ 为[t-N,t-1]时间段内电流值的标准差; $I_{t-1}$ 为t时刻的电流值, $I_{t-2}$ 为t-2时刻的电流值, $I_{t-N}$ 为t-N时刻的电流值;

S2-3,数据标准化:

采用(公式9)的z-score法对每一变量进行数据标准化:

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i \dots\dots\dots (7);$$

$$\sigma_x = \text{std}(x_1, x_2, \dots, x_N) \dots\dots\dots (8);$$

$$z_i = \frac{(x_i - \bar{x})}{\sigma_x} \dots\dots\dots (9)。$$

3. 根据权利要求1所述的基于机器学习的整流器在线监测方法,其特征是:所述的S3机器学习,采用整流器历史运行数据以及岭回归法进行建模,构建整流器运行时的电流、环境温度、环境湿度与设备温度间的对应关系,即整流器温度模型具体为;

对于线性回归问题,最小二乘法的目标函数为(公式10):

$$(\alpha, \beta) = \text{argmin} \left\{ \sum_{i=1}^n (y_i - \alpha - \sum_{j=1}^p \beta_j x_{ij})^2 \right\} \dots\dots\dots (10);$$

式中:

$x_{ij}$ —— $x^i = (x_{i1}, \dots, x_{ip})^T$ 为第i个观测样本的自变量值,该样本共有p个特征变量;

$y_i$ ——第i个样本的因变量值;

$\alpha$ ——截距项;

$\beta_j$ ——第j个特征变量的系数;

$\beta$ ——特征变量系数构成的向量;

令 $\theta = (\alpha, \beta)$ ,则目标函数变为最小化 $\|X\theta - y\|^2$ ;对目标函数的参数进行求导,可获得该目标函数的求解公式(公式11):

$$\theta = (X^T X)^{-1} X^T y \dots\dots\dots (11);$$

在公式中, $\theta$ 为待求参数向量, $X$ 为样本特征矩阵, $y$ 为因变量值向量;

当 $X$ 不是列满秩、或列之间存在多重共线性时,在目标函数加上一个正则化项,使得目标函数变为最小化 $\|X\theta - y\|^2 + \|\lambda I\|^2$ ;

对正则化后的目标函数进行求导后,可获得目标函数求解公式(公式12),该方法称为岭回归法;

$$\theta = (X^T X + \lambda I)^{-1} X^T y \dots\dots\dots (12);$$

在公式中, $\lambda$ 为岭参数, $I$ 为单位矩阵, $\lambda I$ 为增加的正则项。

4. 根据权利要求1至3任意一项所述的基于机器学习的整流器在线监测方法,其特征是:还包括步骤S5,整流器状态与预警等级计算:

在模型预测的基础上使用区间判断与动态监测的方式进行判别,区间判断与动态监测流程的步骤具体如下:

1) 判断当前设备预测温度与真实温度的残差绝对值是否大于设定阈值,当残差绝对值大于阈值时,监测程序自动启动,对设备状态进行监测;

2) 对后续N个数据点的残差绝对值进行统计,当其中n个数据点的残差绝对值大于阈值时,设备状态信息 $T_p$ 自动加1;

3) 重复步骤2),当出现连续p个区间均出现n个数据点的残差绝对值大于等于阈值,即 $T_p = p$ 时,则备预警信息 $T_w$ 自动加1;若其中某一区间内残差绝对值不满足条件时,则 $T_p$ 与 $T_w$ 数值清零,返回步骤(1),重新开始监测模型预测结果。

## 一种基于机器学习的整流器在线监测方法

### 技术领域

[0001] 本发明涉及一种基于机器学习的整流器在线监测方法。

### 背景技术

[0002] 整流器(Rectifier)是一个整流装置,可将交流(AC)转化为直流(DC),该设备主要有两个功能,一是将交流电(AC)变成直流电(DC),经滤波后供给负载;二是给储能系统供电。作为供电系统的一部分,整流器的运行状态与供电系统的稳定性直接相关,通过对整流器设备的运行状态进行监测、判断与预警,可进一步提升供电系统运行时的安全性。

[0003] 到目前为止,国内外对整流器运行状态的监测方式主要是采集设备运行时的电流、环境温度、环境湿度、设备温度数据并通过数值进行展示。然而,现有的监测方式一方面无法描述出变量间的相互影响关系,另一方面在进行整流器状态评判时必须人为的设定指标阈值(如设备温度阈值),且当指标值在阈值范围内时则无法准确、有效的评估整流器的整体状态及其变化趋势。

### 发明内容

[0004] 本发明所要解决的技术问题,就是提供一种基于机器学习的整流器在线监测方法,该方法以整流器真实温度值为状态评判指标,构建整流器电流、功率、环境温湿度与整流器真实温度之间的温度模型,在对整流器状态进行监测时,使用温度模型实时计算整流器温度的预测值,并进行设备预测温度值与真实温度值的对比,通过对结果进行判别,实时甄别出整流器非正常的运行趋势,对整流器的运行状态进行监测。

[0005] 解决上述技术问题,本发明采用如下技术方案:

[0006] 一种基于机器学习的整流器在线监测方法,其特征是包括以下步骤:

[0007] S1,数据采集:

[0008] 使用电流传感器、温度传感器和湿度传感器采集整流器运行时的电流、设备温度、环境温度以及环境湿度的数据;

[0009] S2,数据处理:对电流数据进行动态平滑处理,方法包括移动平滑法、非对称局部加权回归散点平滑法、移动窗口拟或多项式平滑方法;

[0010] S2-1,采用移动平滑法即公式(1)对电流数据进行动态平滑处理;由于电流变化波动较大,直接使用电流进行预测会导致结果波动较大;因此,在进行建模及温度预测前需要对电流数据进行平滑处理,确保预测温度不出现较大的波动;

[0011] 在本发明中,采用移动平滑法对电流数据进行移动平滑处理:

$$[0012] \quad \bar{I}_t = \frac{1}{N}(I_{t-1} + I_{t-2} + \dots + I_{t-N}) \dots\dots\dots (1);$$

[0013] 其中, $\bar{I}_t$ 为当前时刻平滑后的电流值, $I_t$ 为当前时刻的实际电流值,(1,2,⋯, N)为当前时刻向后平滑多少个时间点,如 $I_{t-1}$ 为t-1时刻的电流值, $I_{t-2}$ 为t-2时刻的电流值, $I_{t-N}$ 为t-N时刻的电流值;

[0014] S2-2,使用电流数据构造新衍生变量:

[0015] 针对电流数据构造其衍生变量,衍生变量类型包含电流平方值、电流立方值、平滑电流平方值、平滑电流立方值、一段时间内的电流平方累计值(公式(2))、一段时间内的电流立方累计值(公式(3))、电流最大(公式(4))、最小值(公式(5))和标准差(公式(6));

[0016]  $\text{sum}(I_t^2) = \sum I_{t-1}^2 + I_{t-2}^2 + \dots + I_{t-N}^2 \dots\dots\dots (2);$

[0017]  $\text{sum}(I_t^3) = \sum I_{t-1}^3 + I_{t-2}^3 + \dots + I_{t-N}^3 \dots\dots\dots (3);$

[0018]  $I_{t\_max} = \max(I_{t-1}, I_{t-2}, \dots, I_{t-N}) \dots\dots\dots (4);$

[0019]  $I_{t\_min} = \min(I_{t-1}, I_{t-2}, \dots, I_{t-N}) \dots\dots\dots (5);$

[0020]  $I_{t\_std} = \text{std}(I_{t-1}, I_{t-2}, \dots, I_{t-N}) \dots\dots\dots (6);$

[0021] 在公式中,  $I_t^2$ 为t时刻的电流平方值,  $I_{t-1}^2$ 为t-1时刻的电流平方值,  $I_{t-2}^2$ 为t-2时刻的电流平方值,  $I_{t-N}^2$ 为t-N时刻的电流平方值;  $I_t^3$ 为t时刻的电流立方值,  $I_{t-1}^3$ 为t-1时刻的电流立方值,  $I_{t-2}^3$ 为t-2时刻的电流立方值,  $I_{t-N}^3$ 为t-N时刻的电流立方值;  $\text{sum}(I_t^2)$ 为[t-N, t-1]时间段内电流平方的累计值,  $\text{sum}(I_t^3)$ 为[t-N, t-1]时间段内电流立方的累计值,  $I_{t\_max}$ 为[t-N, t-1]时间段内电流的最大值,  $I_{t\_min}$ 为[t-N, t-1]时间段内电流的最小值,  $I_{t\_std}$ 为[t-N, t-1]时间段内电流值的标准差;  $I_{t-1}$ 为t时刻的电流值,  $I_{t-2}$ 为t-2时刻的电流值,  $I_{t-N}$ 为t-N时刻的电流值;

[0022] S2-3,数据标准化:

[0023] 将数据按比例缩放,使之落入一个小区间内,消除数据的单位限制,便于不同单位或量级的指标能够进行比较和加权;

[0024] 在本发明中,采用z-score法(正规化方法)(公式9)对每一变量进行数据标准化:

[0025]  $\bar{x} = \frac{1}{N} \sum_1^N x_i \dots\dots\dots (7);$

[0026]  $\sigma_x = \text{std}(x_1, x_2, \dots, x_N) \dots\dots\dots (8);$

[0027]  $z_i = \frac{(x_i - \bar{x})}{\sigma_x} \dots\dots\dots (9);$

[0028] 除了z-score法外还可以是:min-max标准化、log函数转换和atan函数转换方法;

[0029] S3,机器学习:采用整流器历史运行数据以岭回归法、Lasso regression、随机森林、决策树、梯度提升决策树、神经网络或RNN算法进行建模;

[0030] 本发明采用整流器历史运行数据以及岭回归法进行建模,构建整流器运行时的电流、环境温度、环境湿度与设备温度间的对应关系,即整流器温度模型;

[0031] 对于线性回归问题,最小二乘法的目标函数为(公式10):

[0032]  $(\alpha, \beta) = \text{argmin} \left\{ \sum_{i=1}^n (y_i - \alpha - \sum_{j=1}^p \beta_j x_{ij})^2 \right\} \dots\dots\dots (10);$

[0033] 式中:

[0034]  $x_{ij}$ —— $x^i = (x_{i1}, \dots, x_{ip})^T$ 为第i个观测样本的自变量值,该样本共有p个特征变量;

[0035]  $y_i$ ——第i个样本的因变量值(即需要设备实际温度值);

[0036]  $\alpha$ ——截距项；

[0037]  $\beta_j$ ——第j个特征变量的系数；

[0038]  $\beta$ ——特征变量系数构成的向量；

[0039] 令  $\theta = (\alpha, \beta)$ ，则目标函数变为最小化  $\|X\theta - y\|^2$ ；对目标函数的参数进行求导，可获得该目标函数的求解公式(公式11)：

[0040]  $\theta = (X^T X)^{-1} X^T y \dots \dots \dots (11)$ ；

[0041] 在公式中， $\theta$ 为待求参数向量， $X$ 为样本特征矩阵， $y$ 为因变量值向量；当 $X$ 不是列满秩时，或列之间存在多重共线性时， $X^T X$ 的行列式接近于0(即 $X^T X$ 接近于奇异)，则计算  $(X^T X)^{-1}$  时误差会很大，导致传统最小二乘法缺乏稳定性与可靠性；

[0042] 因此，通过放弃最小二乘法的无偏性，以损失部分信息、降低精度为代价获得回归系数更为符合实际、更可靠的回归方法：

[0043] 在目标函数加上一个正则化项，使得目标函数变为最小化  $\|X\theta - y\|^2 + \|\lambda I\|^2$ ；

[0044] 对正则化后的目标函数进行求导后，可获得目标函数求解公式(公式12)，该方法称为岭回归法；

[0045]  $\theta = (X^T X + \lambda I)^{-1} X^T y \dots \dots \dots (12)$ ；

[0046] 在公式中， $\lambda$ 为岭参数， $I$ 为单位矩阵， $\lambda I$ 为增加的正则项；

[0047] S4，模型预测：

[0048] 针对整流器运行产生的新数据，在进行数据处理后将数据传到整流器温度模型，预测整流器在当前电流、环境温度与环境湿度下的温度，通过对比整流器的真实温度与真实温度，实时甄别整流器温度是否按历史正常规律变化；

[0049] S5，整流器状态与预警等级计算

[0050] 由于电流波动较大，模型在部分数据点的预测结果较差，为避免程序误报，在模型预测的基础上使用区间判断与动态监测的方式进行判别；

[0051] 区间判断与动态监测流程的步骤具体如下(详见图3)：

[0052] 4) 判断当前设备预测温度与真实温度的残差绝对值是否大于设定阈值，当残差绝对值大于阈值时，监测程序自动启动，对设备状态进行监测；

[0053] 5) 对后续N个数据点的残差绝对值进行统计，当其中n个数据点的残差绝对值大于阈值时，设备状态信息  $T_p$  自动加1；

[0054] 6) 重复步骤(2)，当出现连续p个区间均出现n个数据点的残差绝对值大于等于阈值(即  $T_p = p$ )，则备预警信息  $T_w$  自动加1；若其中某一区间内残差绝对值不满足条件时，则  $T_p$  与  $T_w$  数值清零，返回步骤(1)，重新开始监测模型预测结果。

[0055] 有益效果：本技术方案从根本上解决了目前人为设定预警状态值、指标值在阈值范围内时无法有效的评估整流器状态的简单监测方式。通过采用历史运行数据构建整流器正常运行时的电流、环境温湿度与整流器温度间的温度模型，可甄别出设备在后续运行过程中的非正常变化趋势。同时，通过使用动态监测程序与区间判别法，使得该技术方案具有较强的抗干扰性，避免程序误报。

## 附图说明

[0056] 图1发明流程图；

- [0057] 图2整流器温度实时状态监测与预测图；
- [0058] 图3整流器状态/预警等级计算流程图；
- [0059] 图4整流器状态/预警等级显示图。

具体实施方式

[0060] 为详细说明技术方案的技术内容、构造特征、所实现的目的及效果,以下结合具体实例并配合附图详细说明。

[0061] 本发明的基于机器学习的整流器在线监测方法实施例,包括以下步骤:

[0062] S1,数据采集:

[0063] 使用电流传感器、温度传感器和湿度传感器采集整流器运行时的电流、设备温度、环境温度以及环境湿度的数据;

[0064] S2,数据处理:

[0065] S2-1,对电流进行动态平滑处理(公式1);由于电流变化波动较大,直接使用电流进行预测会导致结果波动较大;因此,在进行建模及温度预测前需要对电流数据进行平滑处理,确保预测温度不出现较大的波动;

[0066] 在本发明中,采用移动平滑法对电流数据进行移动平滑处理:

[0067] 
$$\bar{I}_t = \frac{1}{N}(I_{t-1} + I_{t-2} + \dots + I_{t-N}) \dots\dots\dots (1);$$

[0068] 其中, $\bar{I}_t$ 为当前时刻平滑后的电流值, $I_t$ 为当前时刻的实际电流值,(1,2,...,N)为当前时刻向后平滑多少个时间点,如 $I_{t-1}$ 为t-1时刻的电流值, $I_{t-2}$ 为t-2时刻的电流值, $I_{t-N}$ 为t-N时刻的电流值;

[0069] 此外,还可以采用非对称局部加权回归散点平滑法(LOWESS)、移动窗口拟和多项式平滑方法;

[0070] S2-2,使用电流数据构造新衍生变量:

[0071] 针对电流数据构造其衍生变量,衍生变量类型包含电流平方值、电流立方值、平滑电流平方值、平滑电流立方值、一段时间内的电流平方(公式(2))、立方累计值(公式(3))、电流最大值(公式(4))、最小值(公式(5))和标准差(公式(6));

[0072] 
$$\text{sum}(I_t^2) = \sum I_{t-1}^2 + I_{t-2}^2 + \dots + I_{t-N}^2 \dots\dots\dots (2);$$

[0073] 
$$\text{sum}(I_t^3) = \sum I_{t-1}^3 + I_{t-2}^3 + \dots + I_{t-N}^3 \dots\dots\dots (3);$$

[0074] 
$$I_{t\_max} = \max(I_{t-1}, I_{t-2}, \dots, I_{t-N}) \dots\dots\dots (4);$$

[0075] 
$$I_{t\_min} = \min(I_{t-1}, I_{t-2}, \dots, I_{t-N}) \dots\dots\dots (5);$$

[0076] 
$$I_{t\_std} = \text{std}(I_{t-1}, I_{t-2}, \dots, I_{t-N}) \dots\dots\dots (6);$$

[0077] 在公式中, $I_t^2$ 为t时刻的电流平方值, $I_{t-1}^2$ 为t-1时刻的电流平方值, $I_{t-2}^2$ 为t-2时刻的电流平方值, $I_{t-N}^2$ 为t-N时刻的电流平方值; $I_t^3$ 为t时刻的电流立方值, $I_{t-1}^3$ 为t-1时刻的电流立方值, $I_{t-2}^3$ 为t-2时刻的电流立方值, $I_{t-N}^3$ 为t-N时刻的电流立方值; $\text{sum}(I_t^2)$ 为[t-N,t-1]时间段内电流平方的累计值, $\text{sum}(I_t^3)$ 为[t-N,t-1]时间段内电流立方的累计值, $I_{t\_max}$ 为[t-N,t-1]时间段内电流的最大值, $I_{t\_min}$ 为[t-N,t-1]时间段内电流的最小值, $I_{t\_std}$ 为[t-N,t-1]时

间段内电流值的标准差;  $I_{t-1}$ 为t时刻的电流值,  $I_{t-2}$ 为t-2时刻的电流值,  $I_{t-N}$ 为t-N时刻的电流值;

[0078] S2-3, 数据标准化:

[0079] 将数据按比例缩放, 使之落入一个小的特定区间内, 消除数据的单位限制, 便于不同单位或量级的指标能够进行比较和加权;

[0080] 在本发明中, 采用z-score法(正规化方法)(公式9)对每一变量进行数据标准化:

[0081]  $\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i$  ..... (7);

[0082]  $\sigma_x = \text{std}(x_1, x_2, \dots, x_N)$  ..... (8);

[0083]  $z_i = \frac{(x_i - \bar{x})}{\sigma_x}$  ..... (9);

[0084] S3, 机器学习:

[0085] 采用整流器历史运行数据以及岭回归法进行建模, 构建整流器运行时的电流、环境温度、环境湿度与设备温度间的对应关系, 即整流器温度模型;

[0086] 对于线性回归问题, 最小二乘法的目标函数为(公式10):

[0087]  $(\alpha, \beta) = \text{argmin} \left\{ \sum_{i=1}^n (y_i - \alpha - \sum_{j=1}^p \beta_j x_{ij})^2 \right\}$  ..... (10);

[0088] 式中:

[0089]  $x_{ij}$ —— $x^i = (x_{i1}, \dots, x_{ip})^T$ 为第i个观测样本的自变量值, 该样本共有p个特征变量;

[0090]  $y_i$ ——第i个样本的因变量值(即需要设备实际温度值);

[0091]  $\alpha$ ——截距项;

[0092]  $\beta_j$ ——第j个特征变量的系数;

[0093]  $\beta$ ——特征变量系数构成的向量;

[0094] 令 $\theta = (\alpha, \beta)$ , 则目标函数变为最小化  $\|X\theta - y\|^2$ ; 对目标函数的参数进行求导, 可获得该目标函数的求解公式(公式11):

[0095]  $\theta = (X^T X)^{-1} X^T y$  ..... (11);

[0096] 在公式中,  $\theta$ 为待求参数向量,  $X$ 为样本特征矩阵,  $y$ 为因变量值向量。当 $X$ 不是列满秩时, 或列之间存在多重共线性时,  $X^T X$ 的行列式接近于0(即 $X^T X$ 接近于奇异), 则计算  $(X^T X)^{-1}$ 时误差会很大, 导致传统最小二乘法缺乏稳定性与可靠性;

[0097] 因此, 通过放弃最小二乘法的无偏性, 以损失部分信息、降低精度为代价获得回归系数更为符合实际、更可靠的回归方法:

[0098] 在目标函数加上一个正则化项, 使得目标函数变为最小化  $\|X\theta - y\|^2 + \|\lambda I\|^2$ ;

[0099] 对正则化后的目标函数进行求导后, 可获得目标函数求解公式(公式12), 该方法称为岭回归法;

[0100]  $\theta = (X^T X + \lambda I)^{-1} X^T y$  ..... (12);

[0101] 在公式中,  $\lambda$ 为岭参数,  $I$ 为单位矩阵,  $\lambda I$ 为增加的正则项。

[0102] S4, 模型预测

[0103] 针对整流器运行产生的新数据, 在进行数据处理后将数据传到整流器温度模型,



预测整流器在当前电流、环境温度与环境湿度下的温度,通过对比整流器的真实温度与真实温度,实时甄别整流器温度是否按历史正常规律变化;

[0104] S5,整流器状态与预警等级计算

[0105] 由于电流波动较大,模型在部分数据点的预测结果较差,为避免程序误报,在模型预测的基础上使用区间判断与动态监测的方式进行判别;

[0106] 区间判断与动态监测流程的步骤具体如下(详见图3):

[0107] 7)判断当前设备预测温度与真实温度的残差绝对值是否大于设定阈值,当残差绝对值大于阈值时,监测程序自动启动,对设备状态进行监测;

[0108] 8)对后续N个数据点的残差绝对值进行统计,当其中n个数据点的残差绝对值大于阈值时,设备状态信息 $T_p$ 自动加1;

[0109] 9)重复步骤(2),当出现连续p个区间均出现n个数据点的残差绝对值大于等于阈值(即 $T_p=p$ ),则备预警信息 $T_w$ 自动加1;若其中某一区间内残差绝对值不满足条件时,则 $T_p$ 与 $T_w$ 数值清零,返回步骤(1),重新开始监测模型预测结果。

[0110] 具体的流程如图1所示,共分为三个步骤。

[0111] 1)步骤1为整流器历史数据建模。采用整流器历史数据进行建模,构建整流器正常运行状态下的温度模型,该步骤采用“数据处理”与“机器学习”的技术环节,先使用“数据处理”技术对整流器的原始数据进行处理与构建特征工程,再使用“机器学习”技术环节中的岭回归法构建整流器温度模型。

[0112] 2)步骤2为整流器状态监测与预测(详见图2)。采集整流器状态数据,一方面进行状态监测。另一方面,结合使用步骤1构建的整流器温度模型,采用“数据处理”、“模型预测”的技术环节,预测整流器在正常状态下的温度值。先使用“数据处理”技术对新数据进行处理并构建相同的特征变量,再将处理后的数据传递给整流器温度模型,由温度模型计算出整流器温度的预测值。

[0113] 3)步骤3为整流器状态/预警等级计算(详见图3)。该步骤使用“整流器状态与预警等级计算”环节的技术。实时的将整流器温度预测值与真实值进行对比,采用区间判断与动态监测判断整流器的运行状态,实时给出整流器的预警信息和状态信息(详见图4)。该预警信息分为三个区间。第一个区间预警级别为0(绿色),表示设备正常运行。第二个区间预警级别为1-11(黄色),表示设备状态需要“注意”。第二个区间预警级别为12(红色),表示设备状态“异常”。

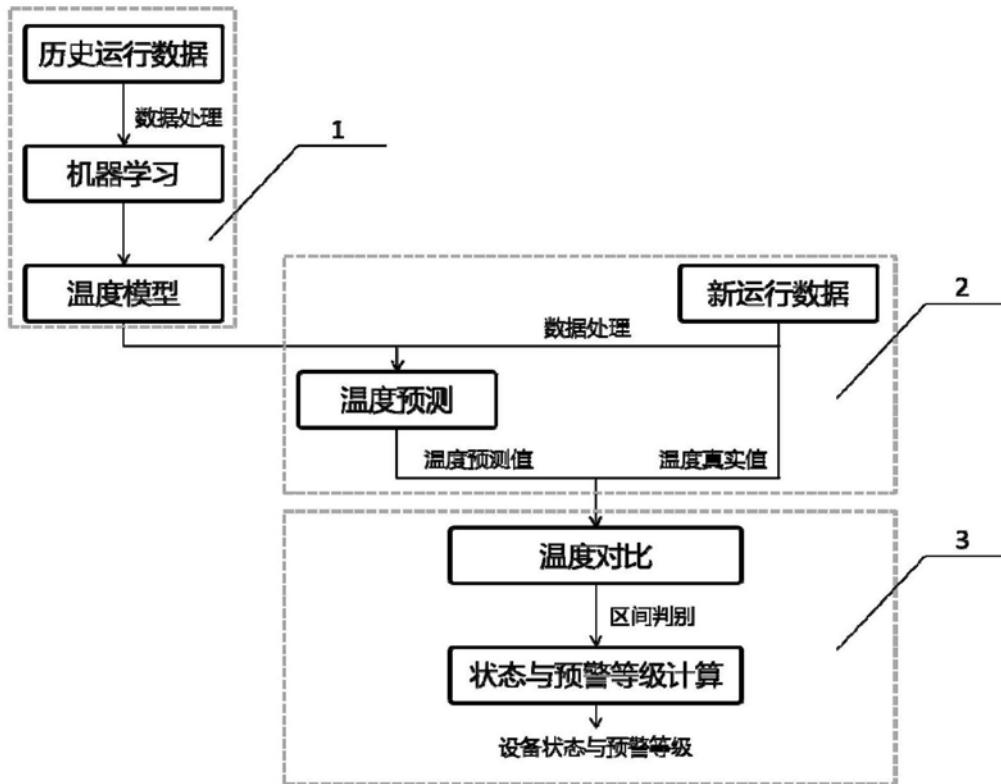


图1

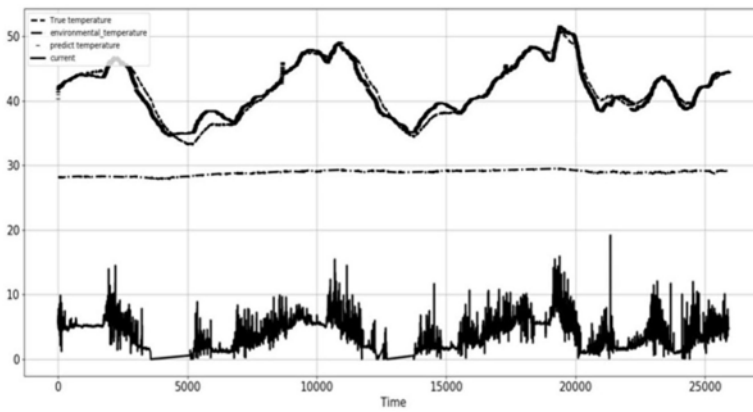


图2

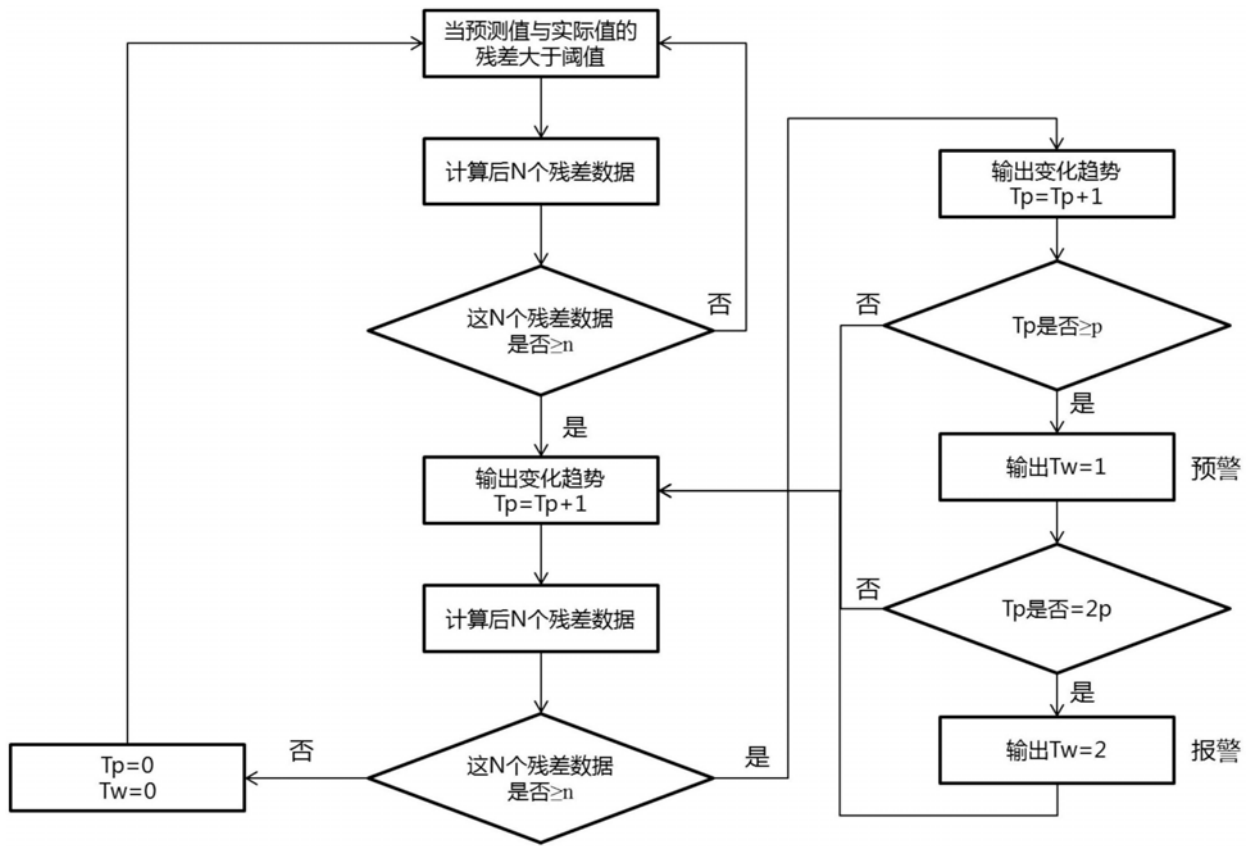


图3

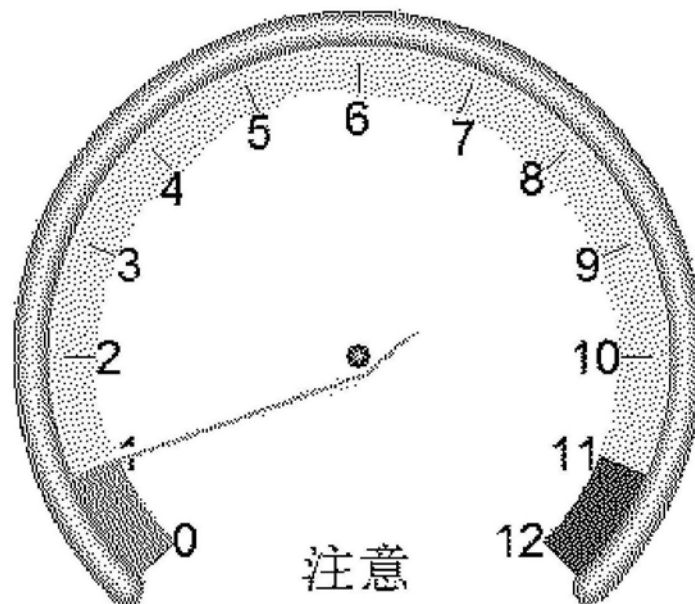


图4