

# 多传感器信息融合技术在变压器故障诊断中的应用

叶汉民,肖尊定

(桂林理工大学机械与控制工程学院,广西 桂林 541004)

**摘要:**本文中作者通过采集变压器运行时的状态参数,运用并行的两个 BP 神经网络对变压器进行局部诊断,再用 D-S 证据理论对局部诊断的结果进行全局融合,实现了对变压器故障的准确诊断。

**关键词:**变压器;BP 神经网络;信息融合;D-S 证据理论;故障诊断

中图分类号:TM406 文献标识码:B 文章编号:1001-8425(2014)09-0060-05

## Application of Multi-Sensor Information Fusion Technology to Transformer Fault Diagnosis

Ye Han-min, Xiao Zun-ding

(Guilin University of Technology, Guilin 541004, China)

**Abstract:** The parameters of transformer in operation are integrated by authors in this papers. Using two parallel BP neural networks to transformer in partial diagnosis, garnish with D-S evidence theory to partial diagnostic results,the authors realize accurate diagnosis of transformer fault.

**Key words:**Transformer;BP neural network;Information fusion;D-S evidence theory;Fault diagnosis

### 1 引言

变压器是电力系统中重要的设备之一,其运行状态直接关系到整个电力系统的安全与稳定。由于变压器结构复杂,影响运行因素众多,变压器在实际诊断过程中存在许多不确定因素,仅靠单一的故障特征和诊断方法得出的结论并不可靠,诊断结果可信度低。本文中笔者将 BP 神经网络和 D-S 证据理论方法结合起来,通过吸收神经网络很强的自学习、自适应和容错能力,使得对不确定信息的融合具有较强的鲁棒性,在此基础上笔者构造了两个 BP 神经网络分别处理色谱数据和电气试验数据,避免了单一诊断神经网络复杂的结构形式、训练时间长,以及某一传感器故障或数据源错误对整个诊断系统带来的不良影响,并且采用试凑法对隐含层节点数进行确定,减少了训练时间和训练次数,提高了诊断的准确率。另外,当变压器同时出现多种故障时,通过对 BP 网络输出结果的拆分,再分别与对应数组融合,有效的实现了对变压器同时出现多种故障的诊断。

### 2 信息融合技术与变压器故障

#### 2.1 多传感器信息融合技术

多传感器信息融合技术是指利用计算机技术对获得的若干传感器的信息在一定准则下加以分析、综合以完成所需决策和估计任务而进行的信息处理过程。

多传感器信息融合中采用的融合算法有贝叶斯估计法、最小二乘法、D-S 证据理论、神经网络、聚类分析和遗传算法等。多传感器信息融合技术具有以下特点:(1)容错性。在单个传感器出错时,仍能正常工作。(2)互补性。利用信息融合能实现不同传感器信息的互补,提高信息利用率。(3)实时性。以较少的时间获得更多的信息,大大提高系统的识别效率。

#### 2.2 变压器故障

变压器故障诊断的方法有:电气试验法、油中溶解气体分析法(DGA)和绝缘检测法等,由油中溶解气体分析法诊断变压器故障时,其内部产生的特征气体的组分和含量与故障严重程度有很密切的关

系,我国现行的《变压器油中溶解气体分析和判断导则》(GB/T7252-2001),将不同的故障类型产生的特征气体归纳成表1。

表1 变压器不同故障类型产生的特征气体成分

故障类型	主要气体成分	次要气体成分
油过热	CH <sub>4</sub> 、C <sub>2</sub> H <sub>2</sub>	H <sub>2</sub> 、C <sub>2</sub> H <sub>6</sub>
油和纸过热	CH <sub>4</sub> 、C <sub>2</sub> H <sub>4</sub> 、CO、CO <sub>2</sub>	H <sub>2</sub> 、C <sub>2</sub> H <sub>6</sub>
油纸绝缘中局部放电	H <sub>2</sub> 、CH <sub>4</sub> 、CO	C <sub>2</sub> H <sub>2</sub> 、C <sub>2</sub> H <sub>6</sub> 、CO <sub>2</sub>
油中火花放电	H <sub>2</sub> 、C <sub>2</sub> H <sub>2</sub>	—
油中电弧	H <sub>2</sub> 、C <sub>2</sub> H <sub>2</sub>	CH <sub>4</sub> 、C <sub>2</sub> H <sub>4</sub> 、C <sub>2</sub> H <sub>6</sub>
油和纸中电弧	H <sub>2</sub> 、C <sub>2</sub> H <sub>2</sub> 、CO、CO <sub>2</sub>	CH <sub>4</sub> 、C <sub>2</sub> H <sub>4</sub> 、C <sub>2</sub> H <sub>6</sub>

### 3 诊断系统结构及原理

笔者综合诊断系统分为三大模块:数据预处理模块、BP神经网络的局部诊断模块和D-S证据理论的全局诊断模块。诊断系统结构如图1所示。

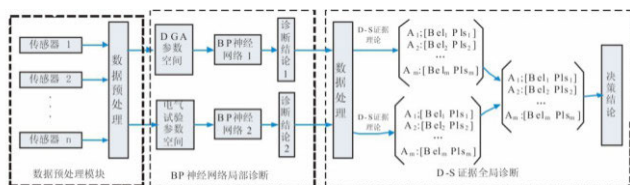


图1 变压器故障综合诊断结构

#### 3.1 数据预处理

神经网络的输入数据具有不同的物理意义,而且各输入数据的量值通常也有很大的差别。若输入数据的绝对值过大会使神经元输出饱和,继而使权值调整进入误差曲面的平坦区,此时就会使误差对权值的变化不太敏感。所以,在训练网络之前,需要对输入的数据作归一化处理,使其范围在[0,1]之间。在归一化处理时,考虑到各气体浓度值相差较大,分别对各个气体浓度值进行处理,本文中笔者采用如下归一化公式:

$$y_{ij} = (x_{ij} - x_{j\min}) / (x_{j\max} - x_{j\min}) \quad (1)$$

其中: $y_{ij}$ 为归一化后的输出数据, $x_{ij}$ 为第*i*个样本的第*j*个特征分量, $x_{j\min}$ 和 $x_{j\max}$ 分别为第*j*个特征分量的最小值和最大值。

#### 3.2 BP神经网络

(1) BP神经网络结构。BP网络是一种采用了误差反向传播(error back propagation)算法的单向传播的多层前馈网络。BP网络有明显的输入、隐含和输出层,隐含层可以有多层,相邻层采用全互连方式连接,同层节点之间没有连接,输出层与输入层

之间也没有直接的联系。图2是BP神经网络的拓扑结构。

(2) BP神经网络的算法原理。BP算法的原理:输入样本从输入层传入,经隐含层逐层处理后,传向输出层。若输出层的实际输出与期望的输出不符,则计算输出误差,将输出误差以某种形式通过隐含层向输入层反传,并将误差分摊给各层的所有单元,从而获得各层单元的误差信号,将此误差信号作为修正各单元权值的依据,周而复始,权值得到不断调整,直到网络输出的误差减少到可以接受的程度。图3为标准BP算法的流程图。

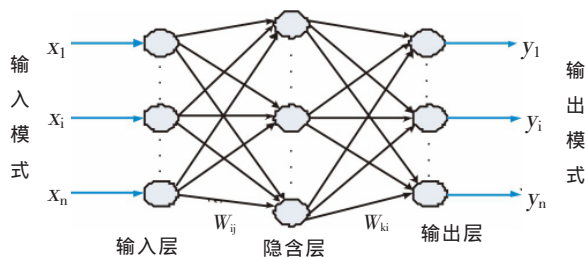


图2 BP神经网络拓扑结构

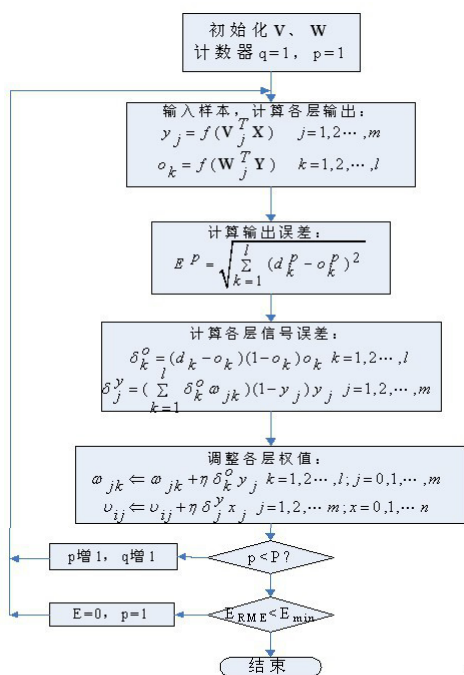


图3 标准BP算法流程图

其中: $V$ 、 $W$ 为初始化的权值随机矩阵; $P$ 为训练样本对总数; $q$ 为训练样本次数计数器; $p$ 为样本模式计数器; $\eta$ 为学习率,取(0,1); $E$ 为误差变量,初值为0; $E_{\min}$ 为网络训练精度,取正小数。

(3) BP神经网络隐含层节点数。BP网络中隐含层节点的作用是从样本中提取并存储其内在规律,每个节点有若干个权值,而每个权值都是增强网

络映射能力的一个参数。节点数量太少,网络从样本中获取的信息能力就差,不足以概括和体现训练集中的样本规律;节点数过多,又可能把样本中非规律性的内容掌握,出现“过度吻合”,降低了泛化能力,增加了训练时间。

隐含层节点数取决于训练样本数的多少、样本噪声的大小以及样本中蕴涵规律的复杂性。节点数的确定,一个常用的方法称为试凑法,通过把隐含层节点数的经验公式作为试凑法的初始值来确定最终的隐含层节点数。常用的隐含层节点数经验公式有:

$$m=\sqrt{n+l}+c \quad (2)$$

$$m=\log_2 n \quad (3)$$

$$m=\sqrt{nl} \quad (4)$$

式中, $m$  为隐含层节点数; $n$  为输入层节点数; $l$  为输出层节点数。 $c$  为 1~10 之间的常数。

### 3.3 D-S 证据理论

D-S 证据理论是由 Dempster 于 1967 年提出的,后由 Shafer 加以扩充和发展,所以证据理论又称为 D-S 证据理论。D-S 证据理论在多传感器信息融合中得到了广泛应用。证据理论中,为了描述和处理不确定性,引入了概率分配函数、信任函数及似然函数等概率。

#### (1) 概率分配函数。

设  $D$  为样本空间,领域内的命题都用  $D$  的子集表示,则概率分配函数定义如下。

设函数  $M:2^D \rightarrow [0,1]$ ,且满足: $M(\emptyset)=0, \sum_{A \subseteq D} M(A)=1$ ,则称  $M$  是  $2^D$  上的概率分配函数, $M(A)$  称为  $A$  的基本概率函数。

#### (2) 信任函数和似然函数。

信任函数用  $Bel$  表示, $Bel$  函数又称为下限函数,以  $Bel(A)$  表示对命题  $A$  为真的信任程度。似然函数用  $Pls$  表示, $Pls(A)$  表示不否认  $A$  的信任程度。 $[Bel(A), Pls(A)]$  称为  $A$  的信任区间。

#### (3) 概率分配函数的正交和。

当对同样的证据得到两个或多个不同的概率分配函数时,需要对它们进行组合,即概率分配函数的正交和。设  $M_1, M_2, \dots, M_n$  是  $n$  个概率分配函数,则其正交和  $M=M_1 \oplus M_2 \oplus \dots \oplus M_n$  为:

$$M(A)=K^{-1} \sum_{\bigcap_{i=1}^n A_i=A} \prod_{i=1}^n M_i(A_i) \quad (5)$$

$$\text{其中, } K=1-\sum_{\bigcap_{i=1}^n A_i=\emptyset} \prod_{i=1}^n M_i(A_i)。$$

如果  $K \neq 0$ ,则正交和  $M$  是一个概率分配函数;若  $K=0$ ,则不存在正交和,称  $M_1$  和  $M_2$  矛盾。

## 4 诊断实例

### 4.1 特征向量的确定和数据样本的收集

(1) 输入、输出特征向量的确定。变压器的故障主要与  $\text{CH}_4$ 、 $\text{H}_2$ 、总烃和  $\text{C}_2\text{H}_2$  4 种气体的浓度有关,所以在 BP 神经网络 1 中把这 4 种气体作为输入层的特征向量。其中各气体浓度的单位为  $\mu\text{L/L}$ 。在 BP 神经网络 2 中,通过对搜集资料的统计和对各项电气试验项目有效性的分析,选取绕组直流电阻不平衡系数、绕组变比、空载电流和空载损耗、局部放电、铁心绝缘电阻及油中糠醛含量作为网络 2 的输入特征向量。根据变压器常见的故障类型选取:一般过热、严重过热、局部放电、火花放电和电弧放电 5 种故障类型作为输出层的 5 个节点输出。

(2) 样本数据的收集。针对变压器故障时可能是一种故障也可能是几种故障的结合,笔者一共收集并模拟产生了 50 组变压器故障的样本数据,部分输入样本数据和期望输出值如表 2 所示。

### 4.2 基于 Matlab 的网络仿真

(1) 隐含层节点数的选择。在 BP1 网络中,采用公式(2)计算隐含层节点数  $m$  取[4,13],训练函数为  $\text{trainscg}$ ,采用试凑法并进行网络仿真,得到图 4 所示的不同隐含层节点数的训练误差曲线,通过对图 4 的分析为了保证较少的训练次数以及较小的均方误差性能函数(MSE)值。本文中 BP 神经网络 1 中隐含层节点数选为 12,网络结构为  $4 \times 12 \times 5$ 。同样的方法可以得到在 BP 神经网络 2 中隐含层节点选为 13,所以 BP2 网络的结构为  $6 \times 13 \times 5$ 。

由图 4(b)可以看出,误差减小到  $10^{-5}$  数量级,满足要求。

(2) 输出结果。选取变压器严重过热、变压器严重过热兼电弧放电两种故障类型进行模拟试验。分别将两组样本测试数据分别输入到 2 个已经训练好的 BP 网络中,经过网络学习后得到 2 个网络的输出  $Y, Z$  分别为:

$$Y_1 = [-0.102 \ 0 \ 0.798 \ 5 \ -0.201 \ 3 \ 0.200 \ 9 \ 0.304 \ 5]'$$

$$Y_2 = [-0.227 \ 8 \ 1.178 \ 9 \ -0.175 \ 8 \ 0.325 \ 7 \ -0.201 \ 2]'$$

$$Z_1 = [0.127 \ 8 \ 0.852 \ 3 \ -0.105 \ 6 \ 1.104 \ 9 \ 0.025 \ 0]'$$

$$Z_2 = [-0.145 \ 6 \ 1.132 \ 4 \ 0.257 \ 8 \ 0.827 \ 3 \ 0.071 \ 8]'$$

### 4.3 D-S 证据理论全局融合

针对 2 个 BP 网络的输出结果,利用 D-S 证据理论对两个网络的输出结果进行处理。神经网络的训练误差作为不确定因素  $m(\theta)$ ,计算公式为:

表 2 输入训练的样本数据

归一化后的数据					期望输出值				
组号	CH <sub>4</sub>	H <sub>2</sub>	总烃	C <sub>2</sub> H <sub>2</sub>	一般过热	严重过热	局部放电	火花放电	电弧放电
1	0.118 9	0.280 0	0.028 9	0.035 7	0	0	1	0	0
2	0.029 0	0.100 0	0.080 2	0.031 3	0	1	0	0	0
3	0.051 5	0.566 7	0.024 8	0.213 8	0	0	0	1	0
4	0.102 8	0.120 0	0.044 3	0	1	0	0	0	0
5	0.135 4	0	0.058 9	0.021 4	0	1	0	0	0
6	0.559 5	0.430 0	0.153 9	0.033 5	0	1	1	0	0
7	0.110 9	0.200 0	0.036 6	0.017 9	1	0	1	0	0
8	0.170 2	0.420 0	0.094 4	0.504 0	0	1	0	0	1
9	0.118 0	0.696 7	0.036 2	0.186 4	1	0	0	0	1
10	0.119 1	0.435 5	0.064 9	0.398 8	1	0	0	1	1

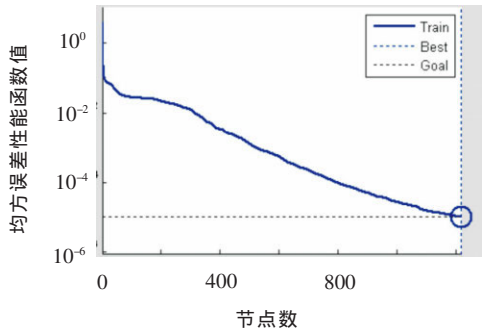
$$E_n=\frac{1}{2}\sum(t_{nj}-y_{nj}) \tag{6}$$

式中  $t_{nj}$ 、 $y_{nj}$  分别为第  $j$  个神经元的期望输出和实际输出值。

将神经网络节点的输出归一化处理，作为各焦点元素的基本概率值，计算公式如下：

$$m(F_i)=y(F_i)/\sum_{i=1}^k y(F_i)+E_n \tag{7}$$

式中， $m(F_i)$  为各焦元的基本概率， $F_i$  为第  $i$  种故障模式， $y(F_i)$  为神经网络的实际输出。



(c) 训练误差曲线(隐含层节点数 13)

图 4 不同隐含层节点数训练误差曲线

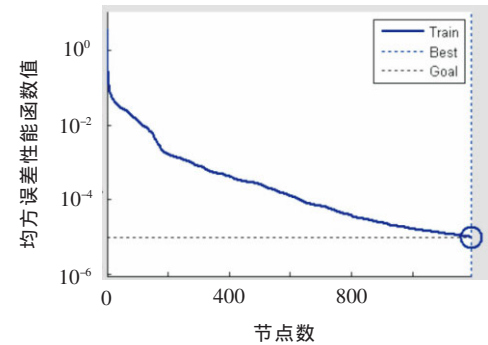
(1)对输出结果  $Y_1$  和  $Y_2$  按公式(1)、公式(6)和公式(7)处理后,得到诊断目标的信度分配,表 3 为两个 BP 网络的信度分配。

通过计算 BP 网络的  $Bel$  和  $Pls$ , 得到各个状态的置信区间;用公式对 BP 网络 1 和 BP 网络 2 进行融合,得到全局诊断后的结果如表 4 所示。

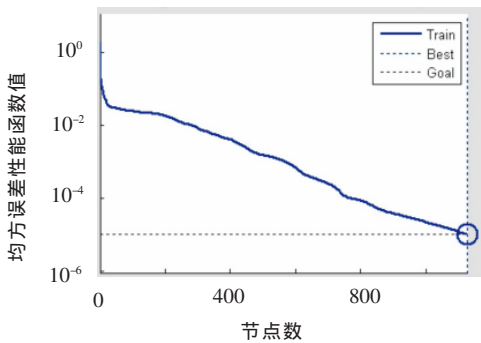
若选取  $Bel>0.6$  且  $m(\theta)<0.15$  作为融合判别准则,从表 4 可以得出,BP 网络融合前的置信区间和不确定性  $m(\theta)$  的信度分配都比较大,诊断结论也不一定确定,但经过融合处理后,故障状态的置信区间明显凸显,不确定性  $m(\theta)$  的信度分配也明显减小,变压器故障类型为严重过热,与实际的故障模拟相一致。

表 3 两个 BP 网络的信度分配

神经网络	$m(F_1)$	$m(F_2)$	$m(F_3)$	$m(F_4)$	$m(F_5)$	$m(\theta)$
BP1	0.046 9	0.471 9	0	0.164 8	0.203 7	0.112 0
BP2	0	0.623 1	0.023 4	0.229 2	0.012 0	0.130 1



(a) 训练误差曲线(隐含层节点数 11)



(b) 训练误差曲线(隐含层节点数 12)



表 4 两个 BP 网络以及融合后的各状态置信区间表

证据元素	$[Bel(F_1), Pls(F_1)]$	$[Bel(F_2), Pls(F_2)]$	$[Bel(F_3), Pls(F_3)]$	$[Bel(F_4), Pls(F_4)]$	$[Bel(F_5), Pls(F_5)]$	$m(\theta)$	诊断结论
BP1	[0.046 9, 0.158 9]	[0.471 9, 0.583 9]	[0.000 0, 0.112 0]	[0.164 8, 0.276 8]	[0.203 7, 0.315 7]	0.112 0	不确定
BP2	[0.000 0, 0.130 1]	[0.623 1, 0.753 2]	[0.023 4, 0.153 5]	[0.229 2, 0.359 3]	[0.012 0, 0.142 1]	0.130 1	严重过热
BP1&BP2	[0.010 8, 0.036 7]	[0.756 1, 0.782 0]	[0.004 6, 0.030 5]	[0.150 9, 0.176 8]	[0.053 8, 0.079 7]	0.025 9	严重过热

(2)当变压器同时发生 2 种或 2 种以上故障时,期望输出结果中有 2 项或 2 项以上的值为 1, 则可以把结果进行拆分。例如输出结果  $O=[0 \ 1 \ 0 \ 0 \ 1]'$  (可判断故障类型为严重过热和电弧放电), 把  $O'$  拆分为  $O_1=[0 \ 1 \ 0 \ 0 \ 0]'$  (严重过热) 与  $O_2=[0 \ 0 \ 0 \ 0 \ 1]'$  (电弧放电) 之和。规定: 对 BP 网络的输出  $O=[a_1 \ a_2 \ a_3 \ a_4 \ a_5]'$ , 若  $O$  中有 2 项或 2 项以上大于 0.6, 则认为变压器故障类型可能由多种组成, 对  $O$  进行拆分, 使得到的数组中只有一项大于 0.6, 小于 0.6 的保持不变, 其他大于 0.6 的项用 0 代替。

对 BP1 网络训练结果  $Z_1$  和 BP2 网络训练结果  $Z_2$  进行拆分, 得到:

$$Z_{11}=[0.127 \ 8 \ 0.852 \ 3 \ -0.105 \ 6 \ 0.000 \ 0 \ 0.025 \ 0]'$$

$$Z_{12}=[0.127 \ 8 \ 0.000 \ 0 \ -0.105 \ 6 \ 1.104 \ 9 \ 0.025 \ 0]'$$

$$Z_{21}=[-0.145 \ 6 \ 1.132 \ 4 \ 0.257 \ 8 \ 0.000 \ 0 \ 0.071 \ 8]'$$

$$Z_{22}=[-0.145 \ 6 \ 0.000 \ 0 \ 0.257 \ 8 \ 0.827 \ 3 \ 0.071 \ 8]'$$

把 BP1 中的  $Z_{11}$  和对应 BP2 中  $Z_{21}$ 、 $Z_{12}$  和  $Z_{22}$  进行证据理论融合。得到的结果如表 5 所示。

从表 5 可以得出, 通过对 BP 网络输出的数组进行拆分, 再对拆分的数组进行融合诊断, 诊断结论为严重过热和电弧放电, 与实际模拟结果相一致, 实现了对变压器同时出现多种故障的诊断。

## 5 结论

本文中笔者将双 BP 神经网络和 D-S 证据理论相结合的数据融合方法引入到故障诊断中, 有效的克服了单一 BP 网络诊断的缺陷。通过把 BP 网络的

输出进行融合, 减少了故障分类识别的不确定性, 提高了诊断系统的可靠性。同时, 通过对变压器的两种故障类型进行模拟实验, 验证了该方法应用于变压器故障诊断的可行性, 为变压器故障诊断的研究提供了有效的途径。

## 参考文献:

- [1] 陈新岗, 田晓霞. 一种结合 DS 证据信息融合的变压器故障诊断方法[J]. 变压器, 2010, 47(8): 59-63.
- [2] 胡文平, 尹项根. 变压器故障诊断与多传感器信息融合[J]. 高压电技术, 2002, 28(2): 32-34.
- [3] 殷跃. 基于 BP 神经网络的电力变压器故障诊断的研究[D]. 吉林: 吉林大学, 2007.
- [4] 韩力群. 人工神经网络理论、设计与应用[M]. 北京: 化学工业出版社, 2007.
- [5] 金聪, 戴上平. 人工智能教程[M]. 北京: 清华大学出版社, 2007.
- [6] 杨海马, 刘瑾. BP 神经网络在变压器故障诊断中的应用[J]. 变压器, 2009, 46(1): 67-70.

## 作者简介:



叶汉民(1965-), 男, 湖北蕲春人, 桂林理工大学教授, 研究方向为自动控制系统。

收稿日期: 2012-12-17

表 5 变压器多种故障融合后的诊断结果

拆分数组融合	$[Bel(F_1), Pls(F_1)]$	$[Bel(F_2), Pls(F_2)]$	$[Bel(F_3), Pls(F_3)]$	$[Bel(F_4), Pls(F_4)]$	$[Bel(F_5), Pls(F_5)]$	$m(\theta)$	诊断结果	最终决策结论
BP1 $Z_{11}'$ & BP2 $Z_{21}'$	[0.022 9, 0.026 7]	[0.853 7, 0.857 5]	[0.010 3, 0.014 1]	[0.052 4, 0.056 2]	[0.063 8, 0.067 6]	0.003 8	严重过热	严重过热兼 电弧放电
BP1 $Z_{12}'$ & BP2 $Z_{22}'$	[0.087 6, 0.100 1]	[0.008 2, 0.020 7]	[0.102 3, 0.114 8]	[0.793 6, 0.806 1]	[0.013 2, 0.025 7]	0.012 5	电弧放电	