第 49 卷 第 6 期：2297-2307 高电压技术 Vol.49, No.6: 2297-2307
2023 年 6 月 30 日 High Voltage Engineering June 30, 2023
DOI: 10.13336/j.1003-6520.hve.20230052

多源异构数据分布下的配电网风险状态多维评估方法

陈 春[1]，黄俊贤[1]，孙辰昊[1]，曹一家[1]，李 勇[2]
（1. 长沙理工大学电气与信息工程学院，长沙 410114；2. 湖南大学电气与信息工程学院，长沙 410082）

摘 要：配电网运行风险状态的全面评估对电力系统的安全可靠运行具有重要意义。为此，综合考虑外部环境条件和内部电气参数，定义配电线路健康度和重要度，提出一种多源异构数据分布下的配电网风险状态多维评估方法。首先，健康度由外部环境信息分析，离散环境特征通过改进的关联规则挖掘(association rule mining，ARM)方法处理，从而能深入探索和分析稀有因子，并通过部件关键度度量(component importance measure，CIM)优化元素相对权重计算方法，进一步提高诊断准确度，连续特征由分层概率模糊推理系统(hierarchical probabilistic fuzzy inference system，HPFIS)处理，分层结构的应用减少了运算复杂性，概率模糊能具体诊断连续特征中的罕见元素。其次，重要度由电气参数分析，利用二阶锥规划和启发式规则(second order cone programming and heuristic rules，SOCPHR)的配电网故障恢复模型求解故障失电后果，继而反映内部用户对供电可靠性需求的差异。基于以上，构建了 ARM-CIM-HPFIS-SOCPHR(ACHS)评估模型，最后，通过对 39 节点和 84 节点实际配电系统进行仿真验证，描述了风险状态二维映射图。研究表明所提方法能有效评估配电网运行风险状态，便于检查和维护。

关键词：配电网风险状态；多维度评估；健康度；重要度；多源异构数据

Multidimensional Assessment Method for the Risk States of Distribution Networks Under Multi-source Heterogeneous Data Distributions

CHEN Chun[1], HUANG Junxian[1], SUN Chenhao[1], CAO Yijia[1], LI Yong[2]
(1. School of Electrical and Information Engineering, Changsha University of Science and Technology, Changsha 410114, China;
2. School of Electrical and Information Engineering, Hunan University, Changsha 410082, China)

Abstract: The comprehensive assessment of distribution network operation risk state is of great significance to the safe and reliable operation of a power system. Therefore, by comprehensively considering the external environmental conditions and internal electrical parameters, we define the healthiness index and importance index of distribution line, and propose a multidimensional assessment model for the risk states of distribution networks under multi-source heterogeneous data distributions. First, the healthiness index is analyzed by external environmental information, discrete environmental features are processed by improved association rule mining (ARM) method so that rare factors can be explored and analyzed in depth, and element relative weights are improved by component importance measure (CIM) to further improve diagnostic accuracy. The continuous features are processed by hierarchical probabilistic fuzzy inference system (HPFIS), the application of hierarchical structure reduces the computational complexity, and the probabilistic fuzzy can specifically diagnose the rare elements in the continuous features. Secondly, the importance index is analyzed by electrical parameters, and the consequences of fault loss are solved using a distribution network fault recovery model combining the second-order cone programming and heuristic rules (SOCPHR), which reflects the difference in internal users' demand for power supply reliability. Based on the above, the ARM-CIM-HPFIS-SOCPHR (ACHS) assessment model is constructed. Finally, through the simulation verification of the 39-node and 84-node actual distribution system, the risk state two-dimensional map is described. The results show that the proposed method can be adopted to effectively assess the operating risk state of the distribution network and facilitate inspection and maintenance.

Key words: distribution network risk status; multidimensional assessment; healthiness indices; importance indices; multi-source heterogeneous data

基金资助项目：国家自然科学基金(52007009)；深圳市科技计划项目(CJGJZD20200617102405015)。
Project supported by National Natural Science Foundation of China (52007009), Shenzhen Science and Technology Planning Project (CJGJZD20200617102405015).

0 引言

配电网在电力系统中有着应对环境变化以及保障能源稳定供应的重要作用，然而由于配电网位置的特殊性，也使得其容易被外界环境所影响。据统计，超过 60%的用户停电是由于配电系统故障所引起的，而这些故障往往来自于诸如台风、冰灾、地震、洪水等极端环境的发生，可见，极端环境是影响配电系统可靠供电的主要原因[1]。

在这种情况下，如果能提前对配电线路的运行状态进行准确评估并作出规划，那么将大大减少潜在的经济损失。因此，通过对配电系统的运行状态进行全面评估，探索其潜在的风险情况，可以有效提高城市配电系统的供电可靠性水平。

目前，国内外已有不少学者通过电信号对电力系统的故障风险进行研究，主要是基于暂态模型[2-3]、电压与电流的提取[4]、频率波动的监测[5]等手段以达到风险评估的目的。上述应用电气属性的分析方法可以产生较为精准和及时的输出，更适合于在线监测和瞬时预测，以助于系统的自动应对。但与此同时，它们不能够提供人工应对和运维检修所需的充足准备时间，因此，为满足这一需要，应延长应对事故的提前准备时间。

近年来，由于环境特征在更长的时间尺度上对故障事件进行预测具有优势性，以及电网数据管理系统和数据挖掘技术的完善性，使得基于数据驱动模型和环境特征的故障风险分析方法得到广泛的应用[6-7]。文献[8-10]分别对电力系统在地震、台风以及极端冰雪影响下的故障风险进行了分析和评估。尽管上述这些研究取得了比较可观的进展，但有些问题仍有待改进。一是部分研究通常利用固定的天气模型分析特定的故障场景情况，缺少对环境场景的全面性考虑；二是部分方法对所有数据进行统一分析，未考虑数据的不平衡分布。

关联规则挖掘(association rules mining，ARM)方法常被用于探索个体变量之间的潜在关系，能更直观灵活的分析故障数据信息和电网的深层联系[11-12]，并且许多学者也致力于将关联规则应用于环境特征和配电系统的相关性研究[13-16]，其都表明关联规则在分析两者关系时，在适应性和全面性方面更具优势。

然而，这其中仍存在一些不足。首先，传统 ARM 算法仍采用相同且固定的评分方法分析所有变量，对于稀有变量以及高风险低概率(high risk-low probability，HRLP)变量，将会直接忽略它们不作任何分析。其次，目前的元素相对权重计算方法通常简单地基于统计频率，整个分析过程没有对元素权重进行调整和优化。此外，对于具有多类型数据的研究对象，现有预测方法一般采用相同模型处理不同类型的输入变量，缺乏针对性分析。最后，部分研究在分析环境影响下的故障风险时，便不再考虑电气属性方面的作用，忽略了内部效应。

为了解决这些问题，本文提出一种多源异构数据分布下的配电网风险状态多维评估方法。此方法引入健康度和重要度，分别代表来自外部环境特征和内部电气特征的线路潜在风险状态。首先通过求解环境特征的健康度贡献值进而求解线路健康度，离散环境特征由改进的 ARM 评分计算方法进行处理，并基于 CIM 优化元素相对权重计算，连续环境特征通过具有分层结构和概率模糊的 HPFIS 方法进行处理；接着通过 SOCPHR 模型求解线路故障时的负荷损失量和失电用户数以得到线路重要度；最后以实际的配电系统作为算例分析对象，构建其运行风险状态二维映射图，并验证本文所提方法的准确性和有效性。

1 数据库的建立和处理

1.1 数据的采集与汇总

本实证研究所需的故障事件记录来自中国南方某配电系统，通过对配电系统的监控和信息收集系统的调度，获得了 8558 条可用的故障记录，每条记录所包含的特征数据及所含元素如表 1 所示。

1.2 数据库的建立

关联规则挖掘经常被用来反映一个事务与其它事务之间的相互依存性和关联性。

表 1 所研究的环境特征数据汇总

Table 1 Summary of environmental characteristics data

离散/连续特征	所含元素
月份	1—12
日期	1—31
时刻	00:00—23:00
地形	平原、湿地、丘陵、森林、山顶
风向	东、南、西、北、东南、东北、西南、西北
天气	多云、晴、雨、雪、大风、阴
气象特征	平均气温(℃)、湿度(%)、平均气压(hPa)、平均风速(m/s)、平均降水量(mm/h)
故障结果	故障、未发生

假设 $V=\{l_1, l_2, l_3, \cdots\}$为所研究对象的所有输入变量的集合，项集 O 是 V 的子集，那么一条关联规则可以被写为 $S\rightarrow O$，其中 O 是目标变量或目标变量集。为了能从罕见变量中挖掘出 HRLP 变量，需将常见变量与罕见变量分开进行各自分析。于是本文将一条关联规则进一步拓展为两部分：

$$S^{\mathrm{c}}+S^{\mathrm{r}}\rightarrow O \tag{1}$$

式中：S^{c} 和 S^{r} 分别代表常见变量集和罕见变量集。在本小节中，所有收集的原始输入数据都将进行预处理，形成一个统一的数据处理空间，用于后续评估。

假设 $Z=\{e_1, e_2, e_3, \cdots, e_m\}$是所研究对象的 m 个输入事务的数据库(故障事件的记录)。设 $F=\{f_1, f_2, f_3, \cdots, f_n\}$为一组包含所有特征的集合，其中 j=1, 2, …, n 表示所有 n 个特征中的一个。每一个环境特征 f_j 又由一组环境元素 $a_{j,1}, a_{j,2}, \cdots, a_{j,k}, \cdots, a_{j,w}$ 组成。设 $O=\{O_1, O_2, \cdots, O_i, \cdots, O_m\}$为一组包含所有目标变量的集合，在本文中所预测的目标对象即为实际的故障处理综合后果(故障或未发生)。

基于以上假设，为建立评估数据库矩阵，将各个集合写成矩阵形式。评估数据库可以写成矩阵 $\boldsymbol{G}$，$\boldsymbol{G}$ 由子矩阵 $\boldsymbol{Z}=[e_1, e_2, e_3, \cdots, e_m]^{\mathrm{T}}$、$\boldsymbol{F}=[f_1, f_2, f_3, \cdots, f_n]$、$\boldsymbol{S}$、$\boldsymbol{O}=[O_1, O_2, \cdots, O_i, \cdots, O_m]^{\mathrm{T}}$ 所构成，上标 T 表示矩阵的转置。为了对离散和连续数据进行分别处理，将子矩阵 $\boldsymbol{S}$ 分为两部分：属于离散特征的元素用 Y_{d} 表示，属于连续特征的元素用 T_{g} 表示。因此，评估数据库 $\boldsymbol{G}$ 可以表示为

$$\boldsymbol{G}=\begin{bmatrix}\boldsymbol{Z} & \boldsymbol{Y}_{\mathrm{d}} & \boldsymbol{T}_{\mathrm{g}} & \boldsymbol{O}\end{bmatrix}=\begin{bmatrix} Z & f_1 & \cdots & f_d & f_{d+1} & \cdots & f_n & O \\ e_1 & a_{11} & \cdots & a_{1d} & a_{1(d+1)} & \cdots & a_{1n} & O_1 \\ e_2 & a_{21} & \cdots & a_{2d} & a_{2(d+1)} & \cdots & a_{2n} & O_2 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots \\ e_m & a_{m1} & \cdots & a_{md} & a_{m(d+1)} & \cdots & a_{mn} & O_m \end{bmatrix} \tag{2}$$

式中：从第二行开始的每一行都代表一次故障事件的记录；$f_1, \cdots, f_d$ 代表离散特征，$f_{d+1}, \cdots, f_n$ 代表连续特征；a_{ij} 代表记录 e_i 中特征 f_j 的一个元素。因此，在一个故障事件 e_i 中，$a_{i1}, \cdots, a_{in}$ 属于每个特征中相匹配的元素；O_i 是多因素决定的故障处理后果；e_i 是此故障事件的编号标签。

2　基于改进关联规则和分层概率模糊系统的健康度求解

本文将输入环境特征划分为离散特征和连续特征，通过求解每个环境要素的健康贡献值，进而求解配电线路的健康度。首先对传统关联规则的诊断标准计算方法进行改进，从而能够提取罕见环境元素和 HRLP 变量。并且，CIM 用于建立健康度贡献值，以衡量离散特征中各个元素对系统稳定性的影响差异。接着，针对不同类型数据的特性，构建分层概率模糊推理系统以求解连续变量。因此，各要素对线路健康度的贡献值都能被求解，线路 L 的一条故障记录的健康度计算式如下所示。

$$H_{\mathrm{line}}=\sum_{a}^{N_1}\lambda_{a(\mathrm{CONT})}+\sum_{b}^{N_2}\lambda_{b(\mathrm{DISC})} \tag{3}$$

式中：H_{line} 为线路的健康度，H_{line} 越小，表明配电线路的工作状态越理想；$\lambda_{a(\mathrm{CONT})}$、$\lambda_{b(\mathrm{DISC})}$分别为一条故障记录中第 a 个连续变量和第 b 个离散变量的健康度贡献值；N_1、N_2 分别为一条故障记录中连续变量和离散变量的总数。

2.1　考虑罕见变量的关联规则挖掘模型

2.1.1　诊断标准计算方法的改进

稀有变量中的 HRLP 变量具备发生率低但后果严重性大的特点。因此，需要对 HRLP 变量作进一步诊断。

由式(1)可知，在传统 ARM 中，由于常见变量集 S^{c} 在数据库中的占比大，重要性诊断标准评分方法主要受常见变量集 S^{c} 的影响。此外，在传统 ARM 中，显著性测量的方式一般都是固定的。所以罕见变量往往被忽略。为此，本文根据特征元素改进 ARM 诊断标准计算方法，可以根据不同环境特征下罕见变量的分布情况，计算出不同形式的罕见变量的得分。如果关联规则 $S^{\mathrm{c}}+S^{\mathrm{r}}\rightarrow O$ 包含一个特征 f_j 的任何罕见元素，则改进后用以计算该特征中稀有变量的 ARM 支持度和置信度可以用式(4)和式(5)表示。

$$Q_{j(\mathrm{supp})}=\frac{J_{\mathrm{num}}\left(e_i\in\boldsymbol{G}(i,1);S^{\mathrm{c}}\subseteq\boldsymbol{G}(i,M_{\mathrm{c}})\neq\varnothing;\boldsymbol{G}(i,j)\in S^{\mathrm{r}}\neq\varnothing\right)}{J_{\mathrm{num}}\left(e_i\in\boldsymbol{G}(i,1);\boldsymbol{G}(i,j)\in S^{\mathrm{r}}\neq\varnothing\right)} \tag{4}$$

$$Q_{j(\mathrm{conf})}=\frac{J_{\mathrm{num}}\left(e_i\in\boldsymbol{G}(i,1);S^{\mathrm{c}}\subseteq\boldsymbol{G}(i,M_{\mathrm{c}})\neq\varnothing;\boldsymbol{G}(i,j)\in S^{\mathrm{r}}\neq\varnothing;\boldsymbol{G}(i,n+2)=O(\mathrm{c})\right)}{J_{\mathrm{num}}\left(e_i\in\boldsymbol{G}(i,1);\boldsymbol{G}(i,j)\in S^{\mathrm{r}}\neq\varnothing\right)} \tag{5}$$

式中：i=2, 3, …, m+1 代表评估数据库 $\boldsymbol{G}$ 中的一行；j=2, 3, …, n+1 代表一种特征属性；$J_{\mathrm{num}}()$代表 G 中同时满足所包含全部条件的故障记录的基数；O(c)代表故障事件的综合后果；M_{c} 代表从 2 到(n+1)的数值区间。

2.1.2　基于 CIM 的元素权重改进

显然，配电系统外部不同要素对系统整体可靠性的影响程度是不同的。因此，应设计更为合理准确的输入要素相对权重计算方法，从而规定每个要素的健康度贡献值。为了能单独评估罕见变量，建

立 $\boldsymbol{G}$ 的子矩阵，G_{jr} 代表 $\boldsymbol{G}$ 中包含特征 f_j 中任意罕见元素的故障记录。此外，本节基于单个特征 f_j 所包含元素的集合构建了两个子集合，即常见元素子集合 f_j^{c} 和罕见元素子集合 f_j^{r}。

本节令单个元素 $a_{j,u}\in f_j$ 的健康度为 $\varphi_{a_{j,u}}$，且由两部分组成，可写成

$$\varphi_{a_{j,u}}=\varphi_{j,u}^{\mathrm{c}}+\varphi_{j,u}^{\mathrm{r}} \tag{6}$$

式中：$\varphi_{j,u}^{\mathrm{c}}$ 表示来自常见元素的影响；$\varphi_{j,u}^{\mathrm{r}}$ 表示来自罕见元素的影响。其中，$\varphi_{j,u}^{\mathrm{c}}$ 的数学表达可写为

$$\varphi_{j,u}^{\mathrm{c}}=\begin{cases}0, & a_{j,u}\in f_j^{\mathrm{r}}\\ \sum\limits_{i=2}^{J_{\mathrm{num}}(e_i\in\boldsymbol{G})}J_{\mathrm{num}}\left(\boldsymbol{G}(i,j)=a_{j,u}\right)\Big/J_{\mathrm{num}}(m), & a_{j,u}\in f_j^{\mathrm{c}}\end{cases} \tag{7}$$

式中：$J_{\mathrm{num}}(m)$代表 $\boldsymbol{G}$ 中所有故障记录的基数。

为了针对性的描述每种罕见元素对配电系统产生的影响，可以基于部件关键度度量(component importance measure, CIM)来计算罕见元素的权重。CIM 主要用于衡量整个系统中各部件的灵敏度和影响程度。本文利用 CIM 工具中的风险减少度(risk reduction level, RRL)[17]来衡量部件的权重影响。

当确认元素 $a_{j,u}$ 在事件 e_i 中未出现即未产生任何影响时，RRL 衡量系统整体风险的相对下降程度，可以由式(8)描述。

$$Y^{\mathrm{RRL}}\left(a_{j,u}|\ e_i\right)=\left(1-o\left(p\left(e_i\right)\right)\right)\Big/\left(1-o\left(0,p\left(e_i\right)\right)\right) \tag{8}$$

式中：1–o(0, p(e_i))表示当元素 $a_{j,u}$ 确定未出现时，事件 e_i 出现在系统中的风险；1–o(p(e_i))表示故障 e_i 未发生时系统中的总体风险。

考虑到在一条真实故障记录中，故障发生与否由所有影响它的各个元素所决定，若其中任何一个元素发生变化，这个故障情况就会发生改变甚至不再发生。基于此，若系统中所包含的组成部件都独立，则系统的整体故障风险可以通过所有特征中相应元素的风险乘积来求解，其数学表达式为

$$\begin{aligned}1-o\left(p\left(e_i\right)\right)&=1-\prod_{j=2}^{n+1}o\left(1_{\mathrm{u}},p\left(e_i\right)\right)=1-\\ &\prod_{j=2}^{n+1}\left(\sum_{i=2}^{J_{\mathrm{num}}\left(e_i\in\boldsymbol{G}_{jr}\right)}\frac{J_{\mathrm{num}}\left(e_i\in\boldsymbol{G}(i,1);\boldsymbol{G}(i,j)=a_{j,u};\boldsymbol{G}(i,j)\in f_j^{\mathrm{r}}\right)}{J_{\mathrm{num}}\left(e_i\in\boldsymbol{G}_{\mathrm{t}}(i,1);\boldsymbol{G}(i,j)\in f_j\right)}\right)\end{aligned} \tag{9}$$

因此，结合式(8)—式(9)，式(6)中 $\varphi_{j,u}^{\mathrm{r}}$ 的数学表达式为

$$\varphi_{j,u}^{\mathrm{r}}=\begin{cases}\dfrac{1-\prod\limits_{j=2}^{n+1}\left(\sum\limits_{i=2}^{J_{\mathrm{num}}\left(e_i\in\boldsymbol{G}_{jr}\right)}\dfrac{J_{\mathrm{num}}\left(e_i\in\boldsymbol{G}(i,1);\boldsymbol{G}(i,j)=a_{j,u};\boldsymbol{G}(i,j)\in f_j^{\mathrm{r}}\right)}{J_{\mathrm{num}}\left(e_i\in\boldsymbol{G}(i,1);\boldsymbol{G}(i,j)\in f_j\right)}\right)}{1-\prod\limits_{u=1}^{w}\left(\sum\limits_{i=2}^{J_{\mathrm{num}}\left(e_i\in\boldsymbol{G}_{jr}\right)}\dfrac{J_{\mathrm{num}}\left(e_i\in\boldsymbol{G}(i,1);\boldsymbol{G}(i,j)\neq a_{j,u};\boldsymbol{G}(i,j)\in f_j^{\mathrm{r}}\right)}{J_{\mathrm{num}}\left(e_i\in\boldsymbol{G}(i,1);\boldsymbol{G}(i,j)\in f_j\right)}\right)}, & a_{j,u}\in f_j^{\mathrm{r}}\\ 0, & a_{j,u}\in f_j^{\mathrm{c}}\end{cases} \tag{10}$$

式中：w 代表所有元素的基数。

2.2 分层概率模糊推理系统

模糊集合与清晰集合的不同之处在于，清晰集合中的单个元素对于集合的隶属度只能为“0”或“1”，而模糊集合中的模糊变量的隶属度在 0 到 1 之间[18]。可以看出，采用清晰集合时往往会导致事态的准确性描述过分依赖于数值边界的设定，产生的不确定过高。

2.2.1 输入连续特征的预处理

本文所采用的模糊形式为概率模糊而非直接模糊，概率模糊的优势在于能够对不同特征以相同的标准来划分模糊集，提高 FIs 在不同场景下的适应能力。并且，概率模糊以各特征值的出现概率为基础进行分析，从而使相应的 FIs 能对罕见数据进行更直接和有效的处理。

因此，本小节所评估的目标为模糊概率健康值(fuzzy probability healthiness value, FPHV)。通过对所收集故障记录中的连续特征概率分布进行求解，再利用 FIS 将所计算的概率分布转换为 FPHV。表 1 中 5 个连续特征在输入数据库中的出现概率通过概率分布函数(probability distribution function, PDF)描述，并且通过预设的边界值，将这些 PDF 曲线划分为 4 个区域，通过这些边界值的设定，可以生成相应的 PDF 曲线，如图 1 所示。在图 1 中，每个连续特征的 PDF 曲线被划分成 4 段，原因是本文所设置的输入和输出隶属度函数考虑了 4 个划分等级，并且每个 PDF 曲线图中的 4 段与 4 个等级存在对应关系。例如，PDF 曲线中概率密度最小的分段对应输入隶属度函数中的“罕见”等级，在“湿度”的概率密度图中，第一个区间对应“罕见”，在“平均降水”的概率密度图中，第 4 个区间对应“罕见”。

基于每个 PDF，各个特征值出现在相应特征中的每一个数值区域的概率都能被求解。区域划分如下所述。

平均温度(℃)：≤12,12~21,21~28,>28；平均气压(kPa)：≤99.3,99.3~100,100~100.9,>100.9；平均降水量(mm/h)：≤1.8,1.8~3.8,3.8~5.6,>5.6；平均湿度(%)：≤0.6,0.6~0.75,0.75~0.9,>0.9；平均风速(m/s)：≤1.6,1.6~3.2,3.2~4.6,>4.6。

2.2.2 连续特征健康度的求解

通过 PDF 曲线可以得到连续特征值在所划分区域中的概率，接下来，通过创建输入隶属度函数从而将概率模糊化。经综合考虑，创建了如图 2 所

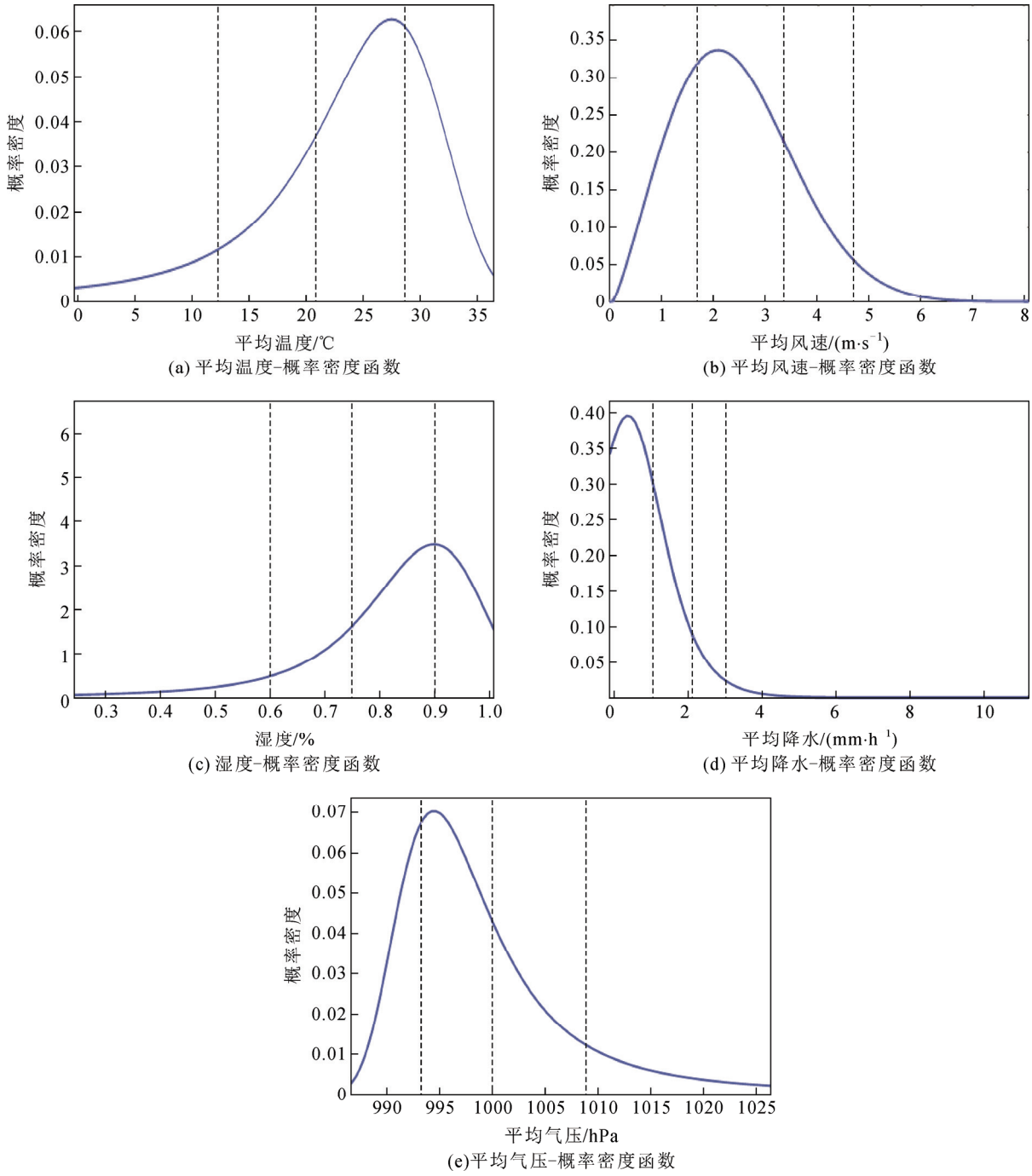


图 1　连续环境特征的 PDF 及各个区域的概率

Fig.1　PDF of continuous environmental features and probabilities for each area

示的三角形输入隶属度函数，4 个三角形所对应的 4 个区域被描述为：罕见(R)、一般(O)、可能(P)、广泛(W)。由于每个连续特征均采用相同的输入隶属度函数，所以在图 2 中仅展示针对“平均温度”的三角形函数形状。

本文采用分层模糊推理系统(hierarchical fuzzy inference system，HFIS)[18]求解连续特征的健康度。如果采用单层 FIs，则分析各包含 4 个模糊集的 5

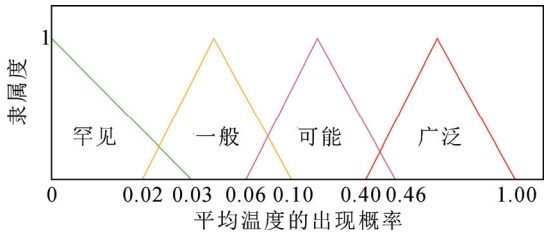


图 2　输入隶属度函数

Fig.2　Input membership functions

个连续特征时，共需要 4^5=1024 条模糊规则，采用多层次模糊系统仅需要考虑 4×16=64 条模糊规则，所构建的 HFIS 如图 3 所示。

基于 HFIS 结构，本文将输出模糊健康指数贡献值设计为 4 类：低(L)、中等(M)、高(H)、极高(E)，相应的模糊权重为 0.2、0.3、0.5、0.9。罕见元素的出现往往伴随着极端事件的发生，即对应着极高模糊健康指数贡献值(E)。综合考虑，可以构建如表 2 所示的模糊规则集。其中：(1)表 2 左半部分代表图 3 中第一个 FIs 所应用的模糊规则(平均温度和平均气压作为输入)，右半部分代表图 3 中其余 4 个 FIs 所使用的的模糊规则。(2)表 2 中右半部分的等级表示前一个 FIs 的产出。(3)∩代表的隶属度计算：$S_C=S_A\cap S_B=\min\{S_A,S_B\}$。

最后，输出隶属度函数由 4 组梯形函数所确定。4 个模糊集对应表 2 中的 4 个健康度级别：低(L)、中等(M)、高(H)、极高(E)。通过对输入数据的统计，可以获得 4 个风险级别中每个类别的数据占比。输出隶属度函数如图 4 所示。

3 基于启发式规则和二阶锥规划的重要度求解

本文基于启发式规则和二阶锥规划相结合的配电网故障恢复模型对重要度进行求解计算，重要度由恢复供电过程中的负荷损失量以及失电用户数所决定，其计算式如下所示。

$$I=\varepsilon_1\sum_{i=1}^{N}\omega_i C_i+\varepsilon_2\sum_{i=1}^{N}U_i \tag{11}$$

式中：N 为故障后失电的负荷数目；ω_i 为负荷的等级($0<\omega\leq 1$)；C_i 为系统中故障失电负荷的第 i 个负荷的容量，kW；U_i 为系统中故障失电的第 i 个负荷点的用户；ε_1、ε_2 为权重系数，$\varepsilon_1+\varepsilon_2=1$，本文取 $\varepsilon_1=\varepsilon_2=0.5$。$I$ 值越大，表明线路越重要，即所面向的用户对当前供电可靠性要求越高。

SOCP 松弛能将所要处理的非凸特性问题转化为混合整数凸规划问题。凸模型最具显著的优点是可以从理论上保证得到的解为全局最优，同时计算高效。本文基于此前团队在文献[19]所做的工作建立 SOCP 凸松弛方法，并结合本团队在文献[20]中所建立的启发式规则和切负荷原则，从而构建本文的二阶锥和启发式规则相结合的(second order cone programming and heuristic rules，SOCPHR)配电网故障恢复模型。

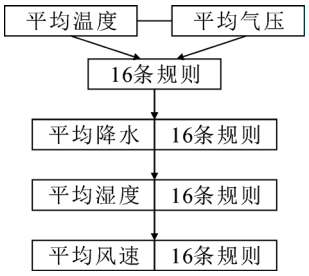


图 3 多层次模糊系统结构

Fig.3 Structure of hierarchical FIS

表 2 模糊推理规则

Table 2 Fuzzy inference rule

特征∩特征	输出	权重	特征∩等级	输出	权重
R∩R	E	0.9	R∩L	M	0.3
R∩O	E	0.9	R∩M	H	0.5
R∩P	H	0.5	R∩H	E	0.9
R∩W	E	0.5	R∩E	E	0.9
O∩R	H	0.9	O∩L	M	0.3
O∩O	H	0.5	O∩M	H	0.5
O∩P	M	0.5	O∩H	H	0.5
O∩W	H	0.3	O∩E	E	0.9
P∩R	M	0.5	P∩L	L	0.2
P∩O	L	0.3	P∩M	M	0.3
P∩P	H	0.2	P∩H	M	0.3
P∩W	H	0.5	P∩E	H	0.5
W∩R	M	0.5	W∩L	L	0.2
W∩O	L	0.3	W∩M	L	0.2
W∩P	L	0.2	W∩H	M	0.3
W∩W	E	0.2	W∩E	H	0.5

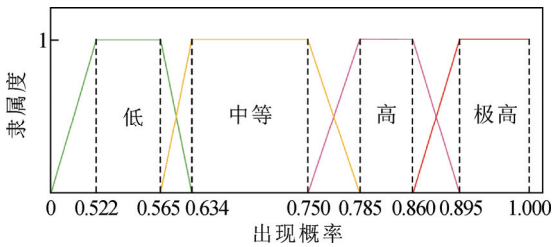


图 4 输出隶属度函数

Fig.4 Output membership functions

与文献[19-20]不同的是，本文旨在用所建立的模型求解量化的指标，即求解当调节分布式电源出力与网络重构仍不能让所有负荷恢复供电时，所采取切负荷手段过程中的负荷损失量和失电用户数，进而得到线路失电后果与线路重要度。

4 风险状态评估的整体流程与框架

通过全文的讨论，本文基于所提出的 ACHS 方

法对配电网线路运行风险状态进行评估的过程如图5所示，具体实施步骤如下所示：

1）提出健康度和重要度两个评估指标，以从外部环境和内部电气量化配电网的运行风险状态。

2）将外部环境特征分为离散特征和连续特征。

3）对于离散环境特征，通过改进的ARM模型探索其所有元素，以分离罕见元素和常见元素。具有该特征的罕见元素的记录被提取至 G_{jr} 中，并通过改进后的关联规则诊断标准计算方法(式(4)—式(5))从罕见元素中进一步挖掘HILP元素，以高频元素集与强关联规则形式体现。

4）对于数据库中的每个离散环境特征，重复步骤3)，并通过式(6)计算离散特征中所有元素的健康度贡献值。

5）对所有连续特征进行预处理。

6）通过分层概率模糊推理系统计算每个连续环境特征的模糊健康度贡献值。

7）基于步骤4）和步骤6）对环境特征所有元素的健康度贡献值进行整合，并将其归一化。

8）利用所构建的SOCPHR故障恢复模型求解线路故障时的负荷损失量和失电用户数，通过式(11)计算线路的健康度，并将其归一化。

9）基于步骤7）所有环境特征的元素健康度贡献值进而求解测试数据集中每条线路故障事件的健康度，基于步骤 8）得到每条线路的重要度，两者的乘积代表线路的运行风险状态。

10）将步骤9）得到的风险状态评分(0→1：不可能发生故障→确定发生故障)与测试数据集真实事件的故障处理结果比较(0 或 1：故障处理结果为故障或未发生)，从而评估ACHS模型的评估性能。

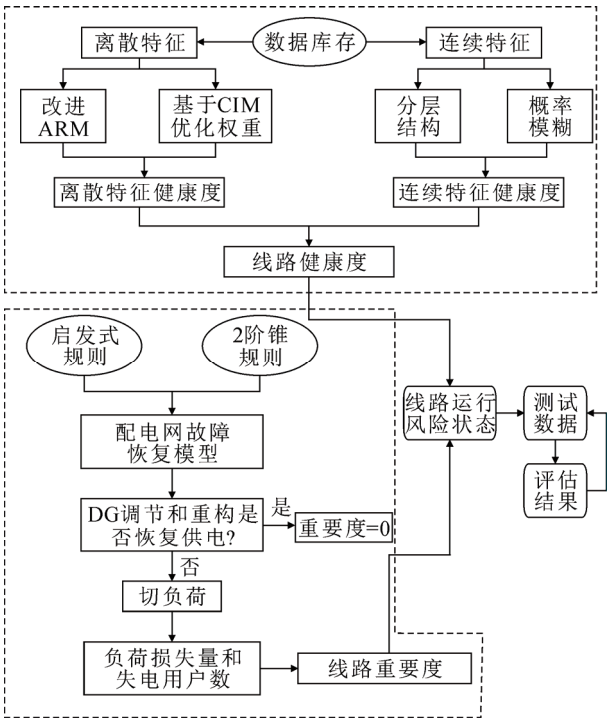
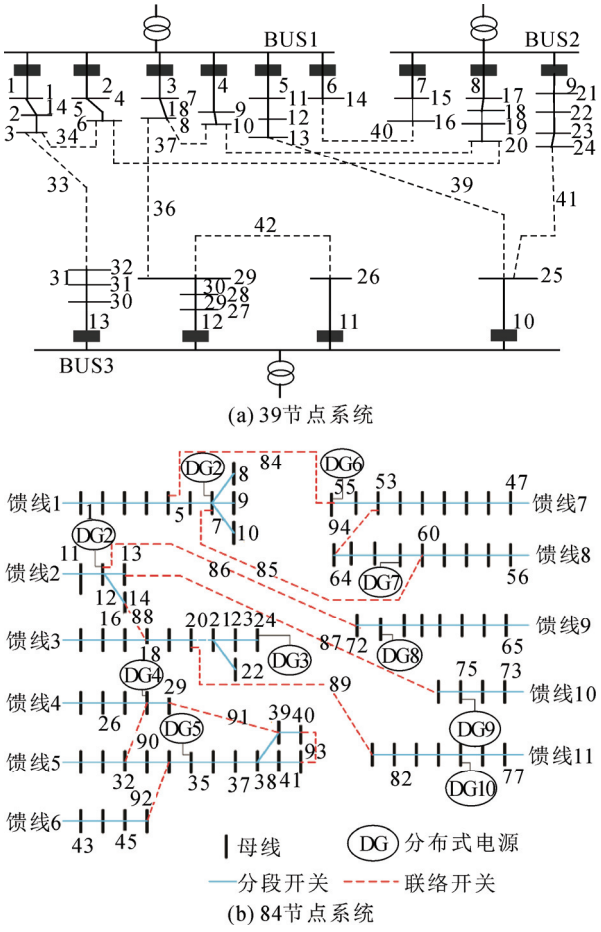


图5　ACHS模型流程图

Fig.5　Flow chart of ACHS model



(a) 39节点系统

(b) 84节点系统

图6　39节点和84节点系统拓扑图

Fig.6　39-node and 84-node system topology

5　案例分析

本次案例分析以实际的39节点和84节点配电系统为研究对象，如图6所示。

5.1　39节点系统

5.1.1　健康度的求解分析

由于在同一时刻，系统中各支路的气象和环境状态不相一致，所以健康度也不一致，本小节以39节点系统中 7—8 支路在某次环境场景下的健康度求解分析为例论述整个模糊健康度的求解过程，具体如图7所示。

此故障事件中的各个连续特征的输入概率由图1所获取：平均温度的概率密度 P_T=30%，平均

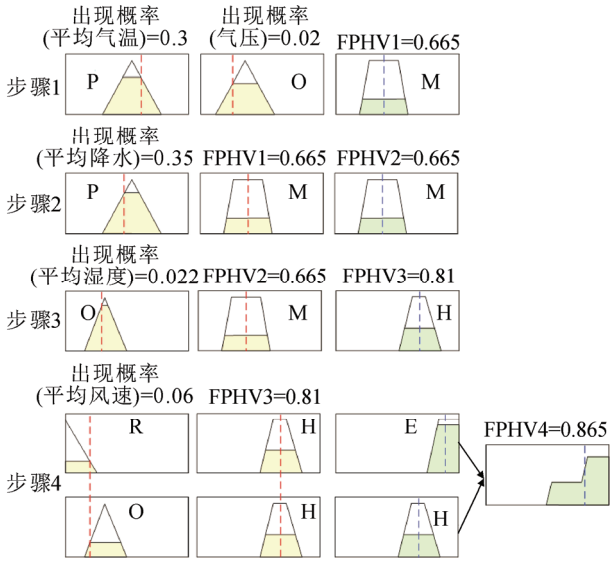


图 7 模糊健康度的求解分析示例(39 节点系统)

Fig.7 An example of fuzzy health index(39-node system)

气压的概率密度 P_p=2%，平均降水量的概率密度 P_w=35%，平均湿度的概率密度 P_{RH}=2.2%，平均风速的概率密度 P_v=5%。这些出现概率在输入隶属度函数中以红色虚线表示，并且基于模糊规则将输入概率的隶属度转化为模糊风险概率区域，接着对得到的风险概率区域进行加权和汇集。最后，累加区域需要通过质心法进行去模糊化，求解出最终的模糊风险概率，以蓝色虚线表示。在本故障事件示例中，5 个连续特征的总体模糊健康指数为 0.865。

5.1.2 重要度的求解分析

本小节同样基于 39 节点系统中 7-8 支路的故障恢复过程为例，以论述线路重要度的求解过程，7-8 支路的故障恢复过程如图 8 所示。

故障场景：如图 8 所示，在经过第 1 阶段的恢复网络连通和分布式电源并网后，通过潮流验证，支路最大负载率与节点电压最低值皆越限。触发阶段 2 的网络重构，经潮流验证节点电压仍然越限，触发切负荷寻优，切除的负荷节点包括 22、27、28、31、35。再次进行潮流验证，满足系统安全运行的要求，恢复供电过程结束。表 3 给出了此故障场景的概述，可知线路 7-8 的重要度为 0.35。

5.1.3 39 节点系统的线路运行风险状态二维映射

当 39 总线系统各支路的健康度和重要度都得以求解后，可得到如图 9 所示的运行风险状态二维映射图。根据工程经验，将图 9 分为Ⅰ、Ⅱ、Ⅲ 3 个区域，Ⅰ→Ⅲ区域的每个风险映射点代表不同支路的运行风险(风险程度低→高对应区域Ⅰ→Ⅲ)。

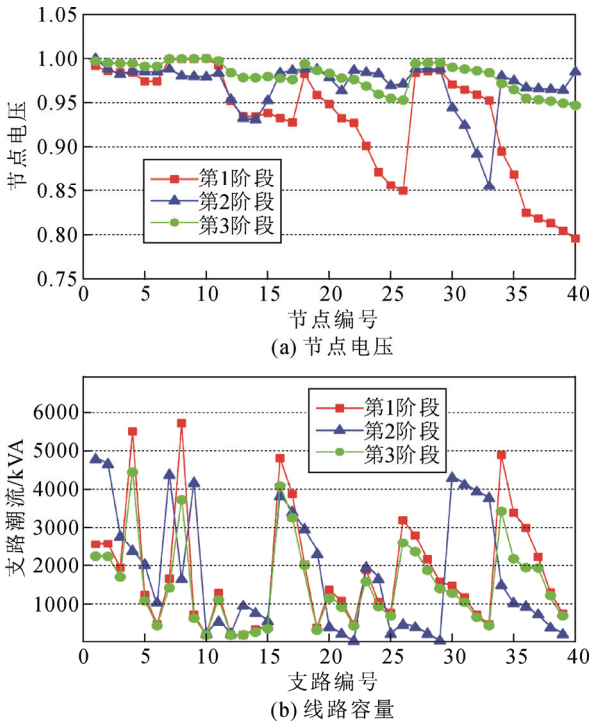


图 8 故障恢复过程的节点电压和线路容量

Fig.8 Node voltage and line capacity distribution during fault recovery

表 3 故障场景描述

Table 3 Fault scenario description

故障支路	切除负荷	重要度
7-8	22,27,28,31,35	0.35

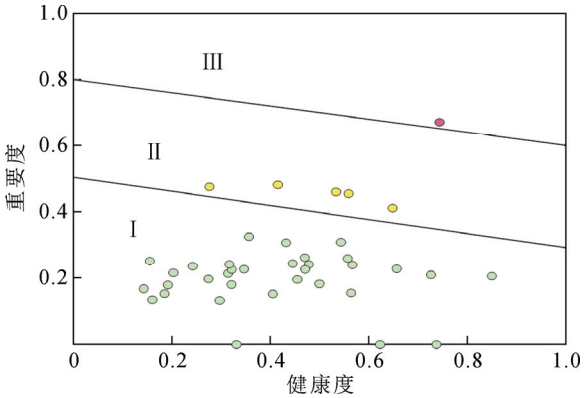


图 9 风险状态的二维映射

Fig.9 Two-dimensional mapping of risk states

由图 9 可知有 6 条线路因风险值较高落入警戒区域(Ⅱ、Ⅲ)，结果如表 4 所示，其中 14 号线路(节点 1-2)的健康度和重要度指数相对较高，表明该线路的运行状态最恶劣，同时该线路所连接用户对供电可靠性的要求最高，因而其发生故障的风险最大。

表 4　警戒区风险评估结果

Table 4　Results of the line operation risk assessment located in the alert area

区域	线路	健康度	重要度
Ⅲ	14(节点 1—2)	0.7447	0.6700
Ⅱ	19(节点 10—11)	0.4157	0.4804
Ⅱ	21(节点 14—15)	0.5595	0.4538
Ⅱ	24(节点 19—20)	0.6491	0.4106
Ⅱ	25(节点 20—21)	0.5340	0.4591
Ⅱ	34(节点 33—34)	0.2769	0.4743

基于图 9 可以直观反映支路在不同角度的运行状态风险情况，表明不同支路承担风险的能力差异，便于调度员判别线路的风险分布并进行相关操作，进而优先对高风险支路(如区域Ⅲ)进行维护检修。

5.1.4　39 节点系统的结果验证

本文通过接受者操作特性曲线(receiver operating characteristic, ROC)评价所构建的模型性能，同时应用 ROC 曲线下面积(area under the ROC, AUROC)作为评价方法，AUROC 的数值越大，模型的性能越好。此外，为减小 ROC 运算过程中的不确定性的影响，在分析 AUROC 时，额外加入准确度(accuracy, ACC)以及约登指数(Youden's index, YI)两个评价标准。ACC 数值越高，模型可信程度越高。YI 数值越高，实验事件的真实性越大。在本案例分析中综合 AUROC、ACC 以及 YI 来对比

ACHS、HPFIS 以及传统 ARM 模型的评估性能，验证的结果如图 10 所示。

由图 10(a)可知，相比于 HPFIS 模型和 ARM 模型，ACHS 模型表现出更高的 AUROC，表明 ACHS 具有更优异的性能。并且由图 10(b)—图 10(c)可以得知，ACHS 在 ACC 和 YI 两种评价标准上同样表现出更优越的性能效果，表明结果有更好的可信度和真实性。

综合 3 个模型评价标准的验证，证明了本文所构建的风险状态评估模型的有效性和可靠性。

5.2　84 节点系统

5.2.1　健康度的求解分析

同样以 84 节点配电系统中支路 17—18 的一次故障记录事件为例，求解此次故障事件中连续特征的模糊健康度贡献值的示例如图 11 所示。

在此次故障事件中，各连续特征的输入概率分别为 P_T=4%，P_p=1%，P_w=50%，P_{RH}=15%，P_v=37%。如图 11 所示，在本次故障事件中，5 个连续特征的总体模糊健康度为 0.625。

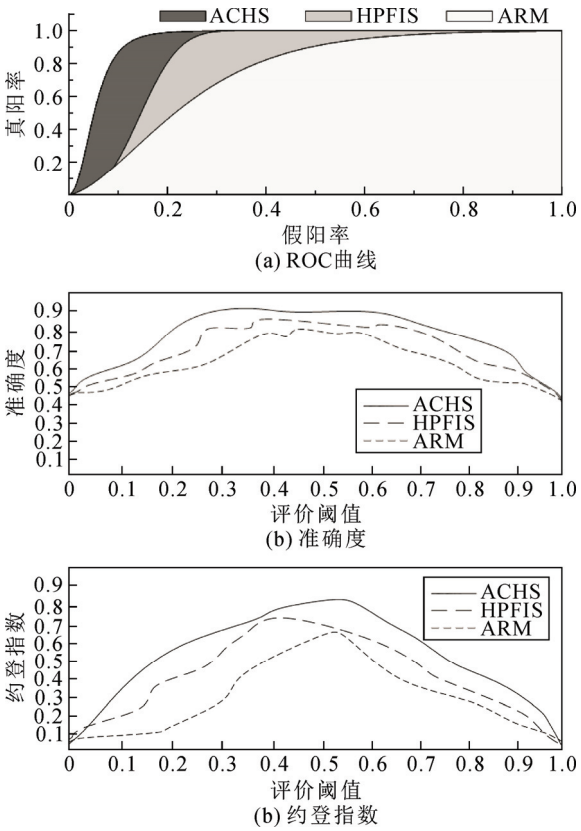


图 10　模型测试效果对比

Fig.10　Test effect comparison of the model

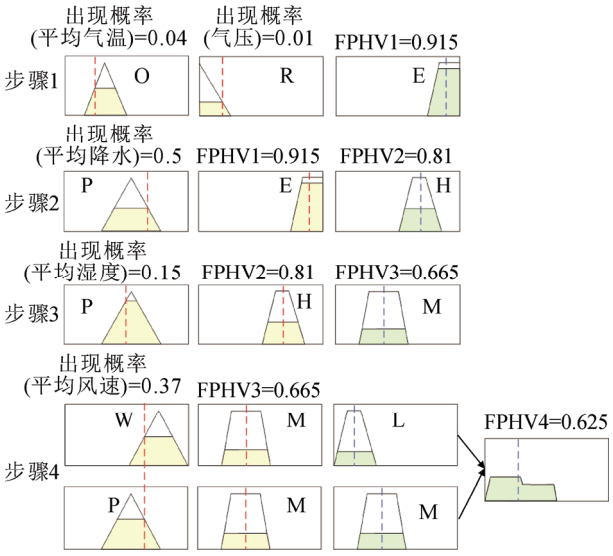


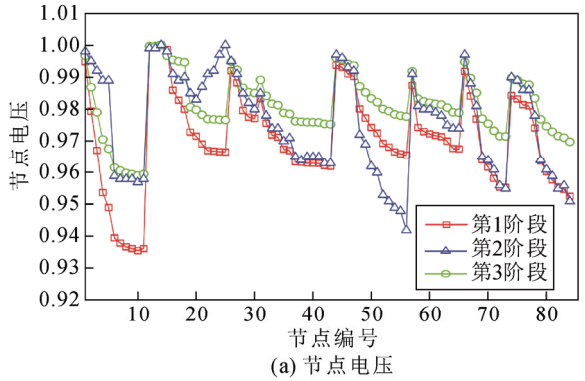
图 11　模糊健康度的求解分析示例(84 节点系统)

Fig.11　An example of fuzzy health index(84-node system)

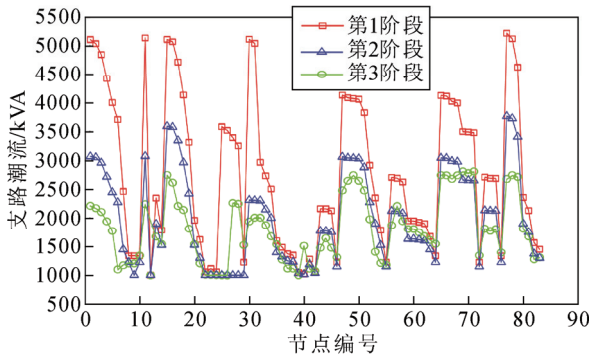
5.2.2　重要度的求解分析

同样可求解 17—18 支路的重要度，图 12 展示了 17—18 支路的故障恢复过程。

当经过 3 个阶段的恢复操作后，系统已满足安全运行的要求，恢复供电过程结束，表 5 给出了此次故障场景的概述，可知 17—18 支路的重要度为 0.32。



(a) 节点电压



(b) 线路容量

图 12 节点电压和线路容量分布

Fig.12 Node voltage and line capacity distribution during fault recovery

表 5 故障场景描述

Table 5 Fault scenario description

故障支路	切除负荷	重要度
17—18	28,32	0.32

5.2.3 84 节点系统的线路运行风险状态二维映射

84 节点配电系统各支路的风险状态二维映射如图 13 所示，可知有 7 条支路在警戒区域内，说明这些支路的运行状态都相对较为恶劣，位于警戒区域的线路的健康度与支持度如表 6 所示。

5.2.4 84 节点系统的结果验证

最后，同样基于 AUROC、ACC 和 YI 3 种模型评价指标，验证 ACHS 模型在 84 节点系统上的评估结果可靠性，并分别和 HPFIS 模型和 ARM 模型进行对比，模型评估结果的对比数据如表 7 所示。

通过表 7 可知，相比于 ARM 模型和 HPFIS 模型，ACHS 模型在 AUROC 的得分上实现了平均 16.7%和 5.1%的提升，在 ACC_{max} 上平均 19.8%和 8%的改善，以及在 YI_{max} 上平均 20.1%和 6.6%的改善。由此看出，本文所提出的 ACHS 模型均获得了更好的评估效果。

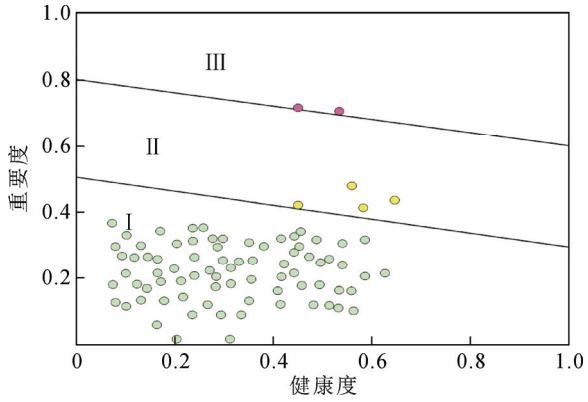


图 13 风险状态的二维映射

Fig.13 Two-dimensional mapping of risk states

表 6 警戒区风险评估结果

Table 6 Results of the line operation risk assessment located in the alert area

区域	线路	健康度	重要度
III	13(节点 12—13)	0.44	0.73
III	19(节点 18—19)	0.53	0.72
II	29(节点 28—29)	0.43	0.45
II	79(节点 78—79)	0.57	0.51
II	72(节点 71—72)	0.58	0.44
II	32(节点 31—32)	0.63	0.47

表 7 评估结果对比

Table 7 Assessment results test comparison

模型	AUROC	ACC_{max}	YI_{max}
ARM	0.758	0.745	0.738
HPFIS	0.842	0.826	0.835
ACHS	0.885	0.893	0.890

6 结论

1）所建立的 ACHS 模型能深入挖掘外界环境因子并将 HRLP 变量区分，同时优化每个元素的相对权重计算，能直观评价单一要素的出现或变化对系统整体影响的相关性，使评估可靠性得以提升。此外，ACHS 模型同时采用了层次模糊结构和概率模糊，减少了运算复杂性并能考虑连续变量中的稀有元素，提高了诊断精度。

2）由电气特征所影响的风险状态也能通过 ACHS 模型进行分析，能有效求解故障时的负荷损失量和失电用户数，进而反映内部用户对供电需求的差异性。

3）通过对 39 节点和 84 节点配电系统进行算例分析，验证了所设计的 ACHS 模型的准确性和有效性，所提出的二维风险状态图可以直观地区分薄弱安全线路，便于检查和维护。

参考文献　References

[1] STOTT P A, CHRISTIDIS N, OTTO F E L, et al. Attribution of extreme weather and climate - related events[J]. WIREs Climate Change, 2016, 7(1): 23-41.

[2] 孟凡成，郭　琦，康宏伟，等. 计及集中式和分布式新能源的电力系统连锁故障模拟[J]. 高电压技术，2022，48(1)：189-198.
MENG Fancheng, GUO Qi, KANG Hongwei, et al. Cascading failure simulation for power system with utility-scale and distributed renewable energy[J]. High Voltage Engineering, 2022, 48(1): 189-198.

[3] 周智行，石立宝. 考虑风电不确定性及系统调频作用的电网连锁故障风险评估[J]. 中国电机工程学报，2021，41(10)：3305-3316.
ZHOU Zhixing, SHI Libao. Risk assessment of power system cascading failure considering wind power uncertainty and system frequency modulation[J]. Proceedings of the CSEE, 2021, 41(10): 3305-3316.

[4] 张　稳，盛万兴，刘科研，等. 高渗透率分布式电源按节点关键性接入配电网的运行风险评估[J]. 高电压技术，2021，47(3)：937-945.
ZHANG Wen, SHENG Wanxing, LIU Keyan, et al. Operation risk assessment of high penetration distributed generation connected to distribution network according to node criticality[J]. High Voltage Engineering, 2021, 47(3): 937-945.

[5] 李梓锋，张　娜，徐熙林，等. 考虑风险偏好的快速频率响应备用容量规划[J]. 高电压技术，2021，47(10)：3475-3484.
LI Zifeng, ZHANG Na, XU Xilin, et al. Reserved capacity planning of fast frequency response considering risk preference[J]. High Voltage Engineering, 2021, 47(10): 3475-3484.

[6] 刘科研，吴心忠，石　琛，等. 基于数据挖掘的配电网故障风险预警[J]. 电力自动化设备，2018，38(5)：148-153.
LIU Keyan, WU Xinzhong, SHI Chen, et al. Fault risk early warning of distribution network based on data mining[J]. Electric Power Automation Equipment, 2018, 38(5): 148-153.

[7] 欧阳森，许伟东. 考虑外部环境信息的电网故障识别与原因分析方法[J]. 高电压技术，2021，47(3)：1075-1082.
OUYANG Sen, XU Weidong. Power grid fault identification and cause identification method considering external environmental information[J]. High Voltage Engineering, 2021, 47(3): 1075-1082.

[8] 郑国鑫，雷　霞，王　湘，等. 地震灾害模拟及配电网的风险评估[J]. 电工技术学报，2020，35(24)：5218-5226.
ZHENG Guoxin, LEI Xia, WANG Xiang, et al. Earthquake simulation and risk assessment of distribution network[J]. Transactions of China Electrotechnical Society, 2020, 35(24): 5218-5226.

[9] 侯　慧，俞菊芳，黄　勇，等. 台风侵袭下输电线路风偏跳闸风险评估[J]. 高电压技术，2019，45(12)：3907-3915.
HOU Hui, YU Jufang, HUANG Yong, et al. Risk assessment of transmission line trip caused by Windage Yaw under typhoon[J]. High Voltage Engineering, 2019, 45(12): 3907-3915.

[10] 王守相，黄仁山，潘志新，等. 极端冰雪天气下配电网弹性恢复力指标的构建及评估方法[J]. 高电压技术，2020，46(1)：123-132.
WANG Shouxiang, HUANG Renshan, PAN Zhixin, et al. Construction and evaluation of resilience restoration capability indices for distribution network under extreme ice and snow weather[J]. High Voltage Engineering, 2020, 46(1): 123-132.

[11] 吴素我，张　焰，苏　运. 基于配用电数据关联的中压配电网断线故障诊断方法[J]. 电力自动化设备，2017，37(7)：101-109.
WU Suwo, ZHANG Yan, SU Yun. Open-line fault diagnosis based on data association of MV distribution network[J]. Electric Power Automation Equipment, 2017, 37(7): 101-109.

[12] 胡丽娟，刁赢龙，刘科研，等. 基于大数据技术的配电网运行可靠性分析[J]. 电网技术，2017，41(1)：276-282.
HU Lijuan, DIAO Yinglong, LIU Keyan, et al. Operational reliability analysis of distribution network based on big data technology[J]. Power System Technology, 2017, 41(1): 276-282.

[13] 徐特威，鲁宗相，乔　颖，等. 基于典型故障与环境场景关联识别的城市配电网运行风险预警方法[J]. 电网技术，2017，41(8)：2577-2584.
XU Tewei, LU Zongxiang, QIAO Ying, et al. A risk warning method for urban distribution network based on associated recognition of typical fault and environment scenario[J]. Power System Technology, 2017, 41(8): 2577-2584.

[14] 杨越文，侯慧娟，杨　祎，等. 基于贝叶斯模型的输电线路运行参量关联规则挖掘及预测方法[J]. 电网技术，2017，41(11)：3648-3654.
YANG Yuewen, HOU Huijuan, YANG Yi, et al. Association rule mining and prediction method for transmission line operation parameters based on bayesian model[J]. Power System Technology, 2017, 41(11): 3648-3654.

[15] WANG Y, XU Z, ZHOU K, et al. On the correlation model between extreme disaster weather and distribution transformer fault types[C]//2021 IEEE 10th Data Driven Control and Learning Systems Conference (DDCLS). Suzhou, China: IEEE, 2021: 935-940.

[16] 梁永亮，郭汉琮，薛永端. 基于特征气体关联特征的变压器故障诊断方法[J]. 高电压技术，2019，45(2)：386-392.
LIANG Yongliang, GUO Hancong, XUE Yongduan. Transformer fault diagnosis method based on association characteristics of characteristic gases[J]. High Voltage Engineering, 2019, 45(2): 386-392.

[17] VESELY W E, DAVIS T C. Two measures of risk importance and their application[J]. Nuclear Technology, 1985, 68(2): 226-234.

[18] CHANG C W, TAO C W. A simplified implementation of hierarchical fuzzy systems[J]. Soft Computing, 2019, 23(12): 4471-4481.

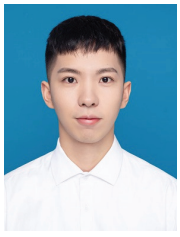
[19] LI Y, XIAO J X, CHEN C, et al. Service restoration model with mixed-integer second-order cone programming for distribution network with distributed generations[J]. IEEE Transactions on Smart Grid, 2019, 10(4): 4138-4150.

[20] WANG F, CHEN C, LI C B, et al. A multi-stage restoration method for medium-voltage distribution system with DGs[J]. IEEE Transactions on Smart Grid, 2017, 8(6): 2627-2636.



陈　春
1987—，男，博士，副教授
主要研究方向为智能配电网规划；配电网故障特性分析与自愈控制，配电网大数据分析与运行控制
E-mail: chch3266@126.com

CHEN Chun
Ph.D.
Associate professor



黄俊贤(通信作者)
1998—，男，硕士生
主要研究方向为大数据下配电网故障风险的评估与预测
E-mail: 1127285013@qq.com

HUANG Junxian
Corresponding author

收稿日期　2023-01-11　修回日期　2023-03-13　　编辑　曹昭君