

11-6-2023

## Over-limit short-circuit current evaluation strategy for power grid with high penetration of renewable energy by combining data-driven and model-driven method

Zhi XIONG

*Economic and Technical Research Institute, State Grid Hubei Electric Power Co., Ltd., Wuhan 430077, China*

Moucheng ZHANG

*College of Electrical Engineering & New Energy, China Three Gorges University, Yichang 443002, China, 121965480@qq.com*

Wei YAO

*State Key Laboratory of Advanced Electromagnetic Engineering and Technology, Huazhong University of Science and Technology, Wuhan 430074, China*

Li QIAO

*Economic and Technical Research Institute, State Grid Hubei Electric Power Co., Ltd., Wuhan 430077, China*

Hongsheng ZHAO

Additional works at: <https://jepst.researchcommons.org/journal>

*Economic and Technical Research Institute, State Grid Hubei Electric Power Co., Ltd., Wuhan 430077, China*

---

### Recommended Citation

XIONG, Zhi; ZHANG, Moucheng; YAO, Wei; QIAO, Li; ZHAO, Hongsheng; LIU, Ju; and WANG, Bo (2023) "Over-limit short-circuit current evaluation strategy for power grid with high penetration of renewable energy by combining data-driven and model-driven method," *Journal of Electric Power Science and Technology*: Vol. 38: Iss. 4, Article 3.

DOI: 10.19781/j.issn.1673-9140.2023.04.003

Available at: <https://jepst.researchcommons.org/journal/vol38/iss4/3>

This Scientific Research is brought to you for free and open access by Journal of Electric Power Science and Technology. It has been accepted for inclusion in Journal of Electric Power Science and Technology by an authorized editor of Journal of Electric Power Science and Technology. For more information, please contact [dlxb04@163.com](mailto:dlxb04@163.com).

---

# Over-limit short-circuit current evaluation strategy for power grid with high penetration of renewable energy by combining data-driven and model-driven method

## Authors

Zhi XIONG, Moucheng ZHANG, Wei YAO, Li QIAO, Hongsheng ZHAO, Ju LIU, and Bo WANG

# 数据—物理融合驱动的含新能源电网 短路电流超标评估策略

熊志<sup>1</sup>, 章谋成<sup>2</sup>, 姚伟<sup>3</sup>, 乔立<sup>1</sup>, 赵红生<sup>1</sup>, 刘巨<sup>1</sup>, 王博<sup>1</sup>

(1. 国网湖北省电力有限公司经济技术研究院, 湖北 武汉 430077; 2. 三峡大学电气与新能源学院, 湖北 宜昌 443002;  
3. 华中科技大学强电磁工程与新技术国家重点实验室, 湖北 武汉 430074)

**摘要:** 目前含新能源电网短路电流超标问题日益严重, 因其状态变化更快、幅度更大, 离线的短路电流超标分析可能无法穷尽超标场景, 在线实时分析具有相当的必要性。考虑到主流物理模型计算方法难以满足在线计算速度需求, 更快速的计算具有重大意义, 因此, 提出一种数据—物理融合驱动的短路电流超标评估策略。首先, 在分析影响短路电流主要因素的基础上, 为了提高计算速度, 将原先特征降维为仅考虑负荷的影响, 再结合最优潮流及随机模拟生成大量样本集合, 通过机器学习算法训练得出数据驱动模型, 在此基础上, 使用漏检率和误检率对数据模型进行阈值整定; 然后, 利用数据驱动模型初步筛选短路电流超标场景; 最后, 通过最新研究中所提理论物理模型对初筛后的短路电流场景进行高精度校验, 并在含光伏电源的 IEEE 39 节点模型上进行验证。仿真结果表明: 该策略可在确保不遗漏超标短路电流场景的前提下有效提升校验速度。

**关键词:** 新能源; 数据—物理融合; 机器学习; 数据驱动; 短路电流超标

DOI: 10.19781/j.issn.1673-9140.2023.04.003 中图分类号: TM713 文章编号: 1673-9140(2023)04-0024-11

## Over-limit short-circuit current evaluation strategy for power grid with high penetration of renewable energy by combining data-driven and model-driven method

XIONG Zhi<sup>1</sup>, ZHANG Moucheng<sup>2</sup>, YAO Wei<sup>3</sup>, QIAO Li<sup>1</sup>, ZHAO Hongsheng<sup>1</sup>, LIU Ju<sup>1</sup>, WANG Bo<sup>1</sup>

(1. Economic and Technical Research Institute, State Grid Hubei Electric Power Co., Ltd., Wuhan 430077, China; 2. College of Electrical Engineering & New Energy, China Three Gorges University, Yichang 443002, China; 3. State Key Laboratory of Advanced Electromagnetic Engineering and Technology, Huazhong University of Science and Technology, Wuhan 430074, China)

**Abstract:** The problem of over-limit short-circuit current in power grid containing renewable energy is becoming increasingly serious. Because of its faster and larger state changes, offline over-limit short-circuit current analysis may not be able to cover all the over-limit scenarios. Therefore, online analysis is quite necessary. Considering that the mainstream physical model calculation method can hardly meet the online calculation speed demand, faster calculation is of great significance. Therefore, this paper proposes a strategy of combining data-driven and model-driven method for over-limit short-circuit current evaluation for power grid with high penetration of renewable energy. Firstly, based on the analysis of the main factors affecting the short-circuit current, in order to improve the calculation speed, the original dimension is reduced to consider only the influence of load. Then, the optimal power flow and random simulation methods are combined

收稿日期: 2022-07-15; 修回日期: 2022-12-08

基金项目: 国家自然科学基金(52107095)

通信作者: 章谋成(1998—), 男, 硕士研究生, 主要从事电力系统运行与控制等研究; E-mail: 121965480@qq.com

to generate a large set of samples, and the data-driven model is obtained through machine learning algorithm training. On this basis, the error analysis and threshold setting of the model are carried out by using false positive rate and false negative rate. Then, the data-driven model is used to screen over-limit short-circuit current scenarios; Finally, the theoretical physical model proposed in the latest research is used to verify the short-circuit current scenario after preliminary screening with high accuracy. It is verified on the IEEE 39 bus model with photovoltaic power supply. The simulation results show that this strategy can effectively improve the verification speed without omitting the over-limit short-circuit current scenarios.

**Key words:** renewable energy; data-physical fusion; machine learning; data-driven; over-limit short-circuit current

电力系统短路电流的校核计算对于继电保护整定和限制短路电流的开关设备及其元件选型具有重大意义<sup>[1-2]</sup>。目前,随着以风电、光伏为代表的新能源机组大量接入电网,短路电流特性已发生潜移默化的改变,现有研究普遍认为高比例新能源接入后会对短路电流水平造成一定程度的提升<sup>[3-5]</sup>,加之新能源的不确定性,可能导致在实际运行中短路电流水平超出断路器的遮断容量。因此,亟需开展含高比例新能源机组的新型电力系统在线、实时的短路电流超标评估研究。

相关领域的学者已针对新型电力系统短路电流问题开展了大量研究。文献[6]建立了不同类型新能源的电压电流映射关系,并进一步提出在故障局部迭代的短路电流计算方法;文献[7]对故障后的稳态电流考虑逆变网络非线性特性和换流器故障穿越策略,对暂态电流则在基于稳态电流值基础上利用微分方程求解;文献[8]根据不同类型新能源机组的低电压穿越控制策略分析了新能源电源等值计算模型,提出了含不同类型新能源接入不平衡配电网的短路计算方法。上述研究主要基于模型迭代法实现了含新能源电网的短路电流计算,但对于在线运行而言,模型迭代存在计算耗时长的缺陷,难以满足工程实际需求。

数据驱动方法在计算速度方面具有显著优势<sup>[9]</sup>。目前,数据驱动方法中的机器学习(machine learning, ML)已在电力系统的诸多层面取得了丰硕的研究成果<sup>[10-15]</sup>,也为未来电网短路电流风险分析提供了新的可行的解决方案<sup>[16]</sup>。文献[17]利用不接入逆变型分布式电源(inverter-interfaced distributed generator, IIDG)时的配电网短路电流作为ML训练时的关键输入特征,得到精确度优于使用普通特征

进行训练的结果,实现对含IIDG配电网的短路电流计算;文献[18]提出了一种基于自适应混沌粒子群一方向传播神经网络的短路电流预测方法,对全网短路电流进行实时快速的预测。但是,以上基于数据驱动的方法在电力系统实际应用方面受到一定限制,其原因在于ML是完全脱离传统物理模型的方法,在准确性及可靠性方面受到相关领域传统专家的质疑。

综上,物理建模方法的优势在于准确度,而劣势在于计算速度,数据驱动方法则反之。单独的数据和物理驱动方法均无法做到在线的含新能源电网短路电流超标校核。而将数据和物理驱动方法融合,有望得到更加精确快速的方法<sup>[18-20]</sup>。文献[18]总结了各领域数据与物理方法,结合电力系统特点,梳理了数据与物理驱动的典型融合模式,即并行、串行、引导和反馈模式;文献[19]提出在频率动态预测过程中,采用系统频率响应(system frequency response, SFR)模型(保留了基本的物理因果关系)与基于极限学习机(extreme learning machine, ELM)的学习模型(用于SFR的拟合和误差修正)的集成;文献[20]提出一种基于门控循环单元神经网络的融合物理与数据知识的频率在线计算方法。而在短路电流计算这一方面,还少有基于数据物理融合方法的研究。

因此,本文将数据驱动方法和物理模型相融合,提出了数据-物理融合驱动的短路电流超标评估策略。在短路电流高精度计算模型的基础上,通过最优潮流+随机模拟的潜在运行方式估计技术生成大量数据样本,训练得到最优数据驱动模型;再利用数据驱动模型对待评估场景进行短路电流超标的快速初筛;最后利用物理模型对初筛后的场景进行高精度校验。

# 1 数据—物理融合驱动的短路电流超标快速评估策略流程

## 1.1 考虑数据驱动误差的数据—物理融合驱动流程

本文所提出的数据—物理融合驱动的短路电流超标快速评估策略流程如图1所示。该策略包括离线阶段和在线阶段。在离线阶段,首先获取历史运行数据,构建出含新能源短路电流高精度计算(物理)模型;接着,利用物理模型仿真为ML训练提

供数据集,从中提取出与短路电流高度相关的特征,在经过数据预处理后,训练基于机器学习的短路电流快速评估(数据驱动)模型;然后进入在线应用阶段,利用数据驱动模型对大量预测场景进行快速在线初步筛选,获得少量的疑似潜在短路电流超标场景集;最后,利用物理模型对疑似场景进行在线高精度校核,得出最终超标场景。

该方法同时结合了数据驱动计算速度快和物理模型计算准确度高的优势,且一定程度上消除了数据驱动方法准确度不如物理模型及物理模型面对大型网络计算速度慢的劣势。

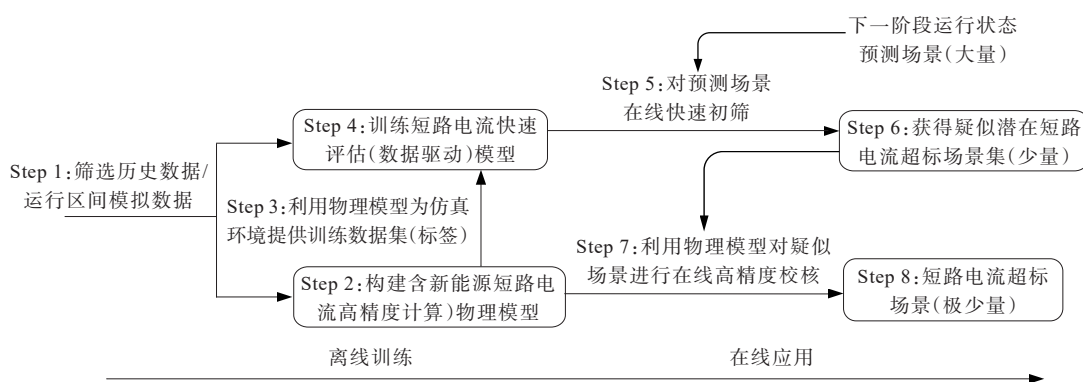


图1 数据—物理融合驱动短路电流超标评估策略

Figure 1 Strategy of data-physical-fusion driven over-limit short-circuit current evaluation

## 1.2 算法误差分析与超标阈值整定

值得注意的是,如果只是简单地利用2种方法进行初筛和精校,实质上并不能完全杜绝数据驱动方法对于超标场景的遗漏问题。

在用数据驱动模型对短路电流是否超标进行初步评估时,由于数据驱动方法的误差,会出现2种情况:①误检,即未超标场景被检测为超标场景,假阳性(false positive, FP);②漏检,即超标场景被模型检测为未超标场景,假阴性(false negative, FN)。

对于检测超标短路电流的误差大小使用误检率(false positive rate, FPR)和漏检率(false negative rate, FNR)2个指标来衡量,FPR代表未超标场景被检测为超标场景的概率,FNR代表超标场景被检测为未超标场景的概率。计算公式如下:

$$\alpha_{\text{FPR}} = \frac{\beta_{\text{FP}}}{\beta_{\text{FP}} + \beta_{\text{TN}}} \quad (1)$$

$$\alpha_{\text{FNR}} = \frac{\beta_{\text{FN}}}{\beta_{\text{FN}} + \beta_{\text{TP}}} \quad (2)$$

式(1)、(2)中, $\alpha_{\text{FPR}}$ 为误检率大小; $\beta_{\text{FP}}$ 为假阳性场景数; $\beta_{\text{TN}}$ 为真阴性(true negative, TN)场景数,TN代表被正确检测出的超标场景; $\alpha_{\text{FNR}}$ 为漏检率大小; $\beta_{\text{FN}}$ 为假阴性场景数; $\beta_{\text{TP}}$ 为真阳性(true positive, TP)场景数,TP代表被正确检测出的未超标场景。

在该策略中,只会将数据驱动方法检测为超标的场景送至物理模型进行再校验,若在数据驱动方法校验时出现第2种漏检情况,后续物理模型则是校验不到的;对于第1种误检情况,后续物理模型也还能再次校验,排除误检,但随之会增加计算时间。因此,可以允许第1种误检情况出现,但要尽可能控制数量,而绝不能允许第2种漏检情况出现。

断路器最大遮断电流称为超标阈值 $I_{\text{lim}}$ ,算法为保证无遗漏,应确保不会误将超标场景判别为未超标场景,故需要将数据驱动阈值整定在超标阈值之下,此时的阈值称为保守阈值 $I_{\text{th}}$ ,其具体数值确定方法如图2所示。

将初始保守阈值 $I_{\text{th}}$ 设定为超标阈值 $I_{\text{lim}}$ ,计算出



FNR,判断其是否为0,若不为0则表示有遗漏的超标场景,则将保守阈值减去0.1 kA 重新进行筛选,再重新计算FNR,如此循环,直到FNR为0,此时得到最终的 $I_{th}$ 。

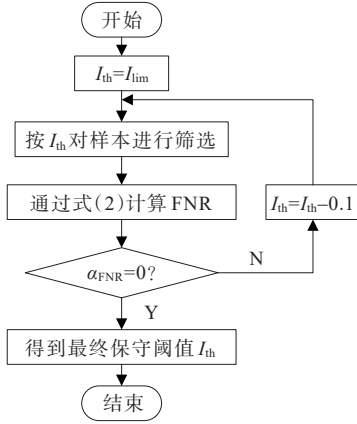


图2 保守阈值整定流程

Figure 2 Flow chart of conservative threshold setting

若按照保守阈值整定,人为降低数据驱动方法中用于判断超标与否的阈值,则可以有效调节FPR和FNR,这种情况必然会增大FP的可能性,导致出现少量的疑似超标场景。因此,对于FP的情况,本文采用物理模型进行高精度校验。

## 2 数据驱动的短路电流快速计算方法

### 2.1 与短路电流水平高度关联的特征选取

在ML中,特征就是输入变量,包含对预测结果有影响的要素,标签则为要预测的属性。在该问题中,流经短路点的短路电流即为标签,而影响含新能源短路电流水平的电网特征则包含多个方面,如新能源出力、发电机出力、节点有功和无功功率以及线路开断状况等。

以上特征虽然全面,但维度较高,特征往往存在冗余,不利于ML的快速训练及准确度提升。因此,把原先同时将负荷和发电机作为输入特征降维为仅考虑负荷的影响,而认为发电机的出力是接近于最优潮流结果的。其他如过渡电阻、短路位置及故障类型,在短路电流计算时一般均作为已知条件,不考虑作为输入特征。具体的特征标签组成如表1所示, $C_{line}$ 为线路随机开断之后的集合,反映电网拓扑结构特征的变化; $P_{new,t}$ 为第 $t$

台新能源机组容量; $P_{d,x}$ 、 $Q_{d,x}$ 分别为节点 $x$ 的有功和无功功率; $I_f$ 为电网接入新能源后流经测量点的短路电流,为样本标签。最终可由以上标签样本构成数据集。

表1 特征标签组成

Table 1 Composition of features and labels

样本	变量含义	符号
	新能源 $t$ 出力	$P_{new,t}$
特征	节点 $x$ 的有功功率	$P_{d,x}$
	节点 $x$ 的无功功率	$Q_{d,x}$
	线路开断集	$C_{line}$
标签	接入新能源后的短路电流	$I_f$

### 2.2 特征及标签样本数据生成方法

为了获得大量用于机器学习训练的数据样本,本文采用一种基于最优潮流(optimal power flow, OPF)+随机模拟的特征生成方法。首先计算系统的OPF分布获得发电机出力,然后在OPF结果上叠加小范围误差。根据Matpower软件包的设定,本文所使用的OPF模型<sup>[21]</sup>为

$$\begin{cases} \min \sum_{i=1}^w (a_k P_{g,k}^2 + b_k P_{g,k} + c_k) \\ \text{s.t. } B_{bus} \theta + P_{bus,shift} + P_d + G_{sh} - C_g P_g = 0 \quad (3) \\ \theta_{i_\theta}^{ref} \leq \theta_{i_\theta} \leq \theta_{i_\theta}^{ref}, i_\theta \in L_{ref} \\ P_{g,k}^{min} \leq P_{g,k} \leq P_{g,k}^{max} \end{cases}$$

式中, $w$ 为发电机数量; $P_{g,k}$ 为第 $k$ 台发电机的有功出力; $a_k$ 、 $b_k$ 、 $c_k$ 为第 $k$ 台发电机的耗量特性参数。 $B_{bus}$ 为节点电纳矩阵; $\theta$ 为节点电压相角向量; $P_{bus,shift}$ 为移相变压器的注入功率向量; $P_d$ 为负的节点负荷向量; $G_{sh}$ 为对地支路电导向量; $C_g$ 为发电机的位置矩阵; $P_g$ 为发电机有功出力向量; $i_\theta$ 为相位参考节点, $\theta_{i_\theta}$ 为 $i_\theta$ 的电压相角值, $\theta_{i_\theta}^{ref}$ 为 $i_\theta$ 的电压相角给定值; $L_{ref}$ 为 $i_\theta$ 的集合; $P_{g,k}^{max}$ 、 $P_{g,k}^{min}$ 分别为第 $k$ 台发电机有功输出的最大值与最小值。

式(3)中目标函数为发电机成本,等式约束条件为直流潮流方程约束,不等式约束条件包括节点电压相角及发电机出力约束。

从OPF分布可以得出各发电机出力,在实际工程中,虽然其接近OPF分布,但受到市场层面和运行技术层面的影响,实际的发电机出力结果并不会与OPF的结果完全一致,故在此基础上叠加小范围

误差  $\eta_s, \eta_s \in [-0.15P_{g,k}, 0.15P_{g,k}] (s=1, 2, \dots, \omega, \omega$  为发电机台数), 使这些误差相加为 0, 从而保持功率平衡, 使其更符合实际工程情况。具体公式为

$$P_{g,k}^* = P_{g,k} + \eta_s \quad s=1, 2, \dots, \omega \quad (4)$$

$$\eta_1 + \eta_2 + \dots + \eta_\omega = 0 \quad (5)$$

其中,  $P_{g,k}^*$  为第  $k$  台叠加误差后的发电机出力,  $P_{g,k}$  为第  $k$  台发电机出力。通过上述计算, 得出会对短路电流产生影响的发电机出力。

对于一个含有  $h$  台新能源机组、 $z$  个负荷节点的电力系统而言, 为了获取足够多的随机样本数据集, 可以根据需要设置随机生成多种运行场景。以某种场景为例, 在此场景下, 新能源基础出力为  $P_{b,a}$  ( $a=1, 2, \dots, h$ ), 各节点基础有功、无功功率分别为  $P_{d,b}, Q_{d,b}$  ( $b=1, 2, \dots, z$ )。设置以风电、光伏为代表的新能源出力在  $\alpha_t \in [0, 1] (t=1, 2, \dots, h)$  范围内随机波动, 靠近新能源机组的节点负荷在  $\beta_j \in [0.8, 1.2] (j=1, 2, \dots, j^*, j^*$  为靠近新能源机组的节点个数) 范围内随机波动, 靠近电源的节点负荷在  $\gamma_u \in [0.4, 1.5] (u=1, 2, \dots, u^*, u^*$  为靠近电源的节点个数) 范围内随机波动, 其他负荷节点在  $\delta_v \in [0.7, 1.4] (v=1, 2, \dots, v^*, v^*$  为其他节点个数) 范围内随机波动。具体计算公式为

$$P_{\text{new},t} = P_{b,a} \cdot \alpha_t \quad (6)$$

$$P_{d,x} = \begin{cases} P_{d,b} \cdot \beta_j, & \text{靠近新能源的节点} \\ P_{d,b} \cdot \gamma_u, & \text{靠近电源的节点} \\ P_{d,b} \cdot \delta_v, & \text{其他节点} \end{cases} \quad (7)$$

$$Q_{d,x} = \begin{cases} Q_{d,b} \cdot \beta_j, & \text{靠近新能源的节点} \\ Q_{d,b} \cdot \gamma_u, & \text{靠近电源的节点} \\ Q_{d,b} \cdot \delta_v, & \text{其他节点} \end{cases} \quad (8)$$

而对于线路开断集合矩阵  $C_{\text{line}}$  的生成, 可通过如下形式表示:

$$l = \{l_1, \dots, l_b, \dots, l_B\} \quad (9)$$

$$l_b = \begin{cases} 0, & \text{线路 } b \text{ 正常运行} \\ 1, & \text{线路 } b \text{ 断开一回} \\ 2, & \text{线路 } b \text{ 双回全断开} \end{cases}$$

式中,  $B$  为系统中输电线路的总数;  $l$  为各线路运行状态;  $l_b$  为第  $b$  条线路的运行状态。断线前所有线路处于正常运行状态。

将随机生成的数据样本输入至物理模型中进行计算, 生成大量短路电流样本数据。

## 2.3 机器学习训练

### 2.3.1 机器学习方法

本文主要通过比较前馈神经网络 (feedforward neural network, FNN), 支持向量回归 (support vector regression, SVR) 和 Bagging 这 3 种 ML 算法的综合表现后, 选取其中效果最好的模型。

1) FNN 属于神经网络的一种, 是单向多层结构, 每一层由多个神经元互相连接组成。每一个神经元以上一层各节点的输出作为输入, 通过线性变换和非线性函数激活, 得出此节点的输出并转送给下一节点。若第  $r$  层有  $M$  个节点, 第  $r+1$  层有  $N$  个节点, 则可得:

$$x_n^{r+1} = f_n^{r+1} \left( \sum_{m=1}^M (\omega_{nm}^{r+1} x_m^r + b_n^{r+1}) \right) \quad (10)$$

式中,  $x_n^{r+1}$  为第  $r+1$  层第  $n$  个节点的输出;  $\omega_{nm}^{r+1}$  为第  $r$  层第  $m$  个节点到第  $r+1$  层第  $n$  个节点的权重;  $f_n^{r+1}$  为第  $r+1$  层第  $n$  个节点的激活函数;  $x_m^r$  为第  $r$  层第  $m$  个节点的输出;  $b_n^{r+1}$  为第  $r+1$  层第  $n$  个节点的偏差。

FNN 的网络结构如图 3 所示。

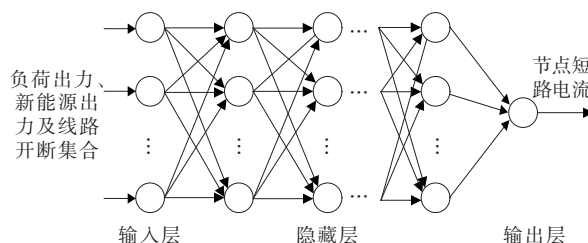


图 3 FNN 结构

Figure 3 Structure of FNN

2) SVR 也是一种用于回归预测问题的算法。在给定的样本数据集  $\{(x_i, y_i)\} (i=1, 2, \dots, m)$  中, 通过非线性映射函数  $\varphi(x)$  将  $m$  维的输入向量映射到  $l$  维的特征空间, 在其中构造一个最优超平面  $f(x) = W^T \varphi(x) + b$ , 其中,  $W$  和  $b$  分别为回归系数和截距。若所有样本点到最优超平面距离  $|y_i - f(x_i)| \leq \epsilon$ , 通过引入非负松弛变量  $\xi_i, \xi_i^*$  和惩罚因子  $C$  后, 将其转化为求解凸二次优化问题:

$$\begin{cases} \min \frac{1}{2} \|W\|^2 + C \sum_{i=1}^m (\xi_i + \xi_i^*) \\ \text{s.t. } y_i - f(x_i) \leq \epsilon + \xi_i \\ f(x_i) - y_i \leq \epsilon + \xi_i^* \\ \xi_i \geq 0, \xi_i^* \geq 0 \end{cases} \quad (11)$$

引入拉格朗日乘子  $\alpha_i$  和  $\alpha_i^*$ , 将式(11)转化为其对偶形式:

$$\begin{cases} \max Q(\alpha, \alpha^*) = - \sum_{i=1, j=1}^M (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) \mathbf{g}^T \cdot \\ \quad (\mathbf{x}_i) \mathbf{g}(\mathbf{x}_j) - \epsilon \sum_{i=1}^M (\alpha_i + \alpha_i^*) + \sum_{i=1}^M y_i (\alpha_i - \alpha_i^*) \\ \text{s.t. } \sum_{i=1}^M (\alpha_i - \alpha_i^*) = 0, 0 \leq \alpha_i, \alpha_i^* \leq C \end{cases} \quad (12)$$

接着采用核函数  $k(x_i, x)$  代替  $\mathbf{g}^T(x_i) \mathbf{g}(x)$ , 得到回归函数为

$$f(x) = \sum_{i=1}^M (\alpha_i - \alpha_i^*) k(x_i, x) + b \quad (13)$$

其中,  $\alpha_i, \alpha_i^*$  和  $b$  可通过样本数据训练多次迭代得出。

3) Bagging 是集成学习中的一种, 集成学习主要思想是分别训练多个基学习器, 在回归问题中通过取平均值将这些学习器结合, 得到最终模型输出。对于 Bagging 而言, 其各个学习之间相互独立, 分开训练, 如图 4 所示。具体为首先通过自助采样法从数据集中有放回的重复随机抽取样本, 形成  $n$  个子训练集, 分别训练得到  $n$  个基学习器, 并通过不同学习器得到的函数值求平均得到最终结果。

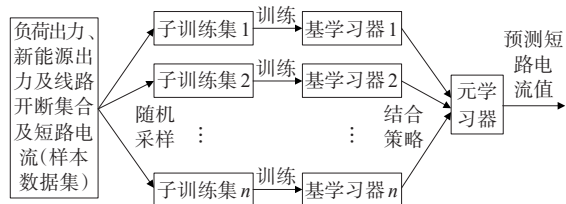


图 4 Bagging 集成学习结构

Figure 4 Structure of bagging ensemble learning

### 2.3.2 数据预处理

在得到样本数据集后, 还需要对数据集进行预处理, 从而提升预测模型的精确度和训练速度并防止过拟合等。本文对新能源出力、节点无功和有功率进行标准化处理, 即将每个特征向量缩放成平均值为 0、方差为 1。去除数据量纲, 将每组特征化为同一数量级, 确保在训练时为特征分配合适的权重, 避免依赖量纲较大的特征数据而造成模型的不精确。具体公式为

$$y_i = \frac{x_i - \bar{x}_i}{s_{id}} \quad (14)$$

式中,  $y_i$  为标准化后数据;  $x_i$  为某一特征数据;  $\bar{x}_i$  为该组特征数据的平均值;  $s_{id}$  为该组输入特征数据的标准差。

开断线路矩阵  $C_{line}$  是通过编号 0、1、2 表征不同的线路开断情况, 但将开断情况转换为数字表示后, 默认成为连续变量有违本质含义, 会对模型精度产生影响, 故对  $C_{line}$  进行独热编码, 以处理表征线路开断情况这种非数值的类别属性。具体编码结构如表 2 所示。

表 2 开断类型编码

Table 2 Coding of break type

开断类型	代表序号	独热码
正常运行	0	100
两回断一回	1	010
双回全断开	2	001

### 2.3.3 训练过程

数据驱动的短路电流评估模型生成整体流程如图 5 所示。

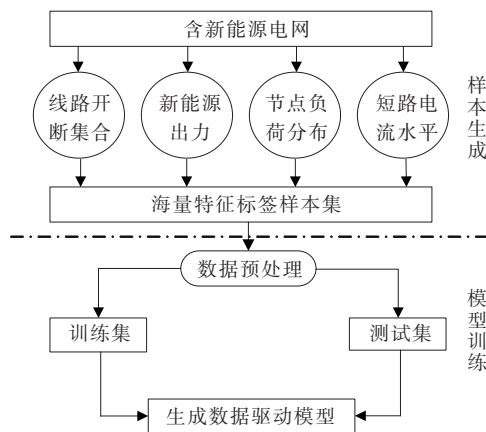


图 5 数据驱动的短路电流评估模型生成流程

Figure 5 Flow chart of data-driven short-circuit current evaluation model generation

在对生成的数据样本集预处理后, 将样本按 8:2 划分为训练集和测试集, 使用交叉验证法利用上述 3 种 ML 算法进行训练, 初步确定性能较好的模型, 最后用贝叶斯优化法对其进行超参数优化, 并将得到模型在测试集上测试其泛化能力, 选取其中表现最佳的模型。其模型的评价指标采用目前常用于回归问题的决定系数 (R-squared) 和均方根误差 (root mean squared error, RMSE) 来综合考量, 两者的数学表达式为



$$R_{\text{MSE}} = \sqrt{\frac{1}{e} \sum_{i=1}^e (y_{\text{actual}} - y_{\text{predict}})^2} \quad (15)$$

$$R^2 = 1 - \frac{\sum (y_{\text{actual}} - y_{\text{predict}})^2}{\sum (y_{\text{actual}} - y_{\text{mean}})^2} \quad (16)$$

式(15)、(16)中,  $y_{\text{actual}}$ 、 $y_{\text{predict}}$  分别为样本标签的预测值和真实值;  $e$  表示样本总数;  $y_{\text{mean}}$  为预测值的平均值。

RMSE用来衡量预测值同真实值之间的偏差, 值越小, 说明模型预测的效果越好; 而  $R^2$  越接近1, 说明模型预测的效果越好。

### 3 新能源电网的短路电流物理模型计算方法

#### 3.1 新能源短路电流等效模型

本文仅考虑光伏电源(photovoltaic, PV)的短路电流等效模型, 采用的光伏电源等效模型为文献[6]中所述, 具体如下。

光伏电源的短路电流解析主要分为低电压穿越和外环控制2个阶段。在外环控制阶段, 由于输出电流变化很小, 故一般定为额定电流  $I_N$ 。在LVRT控制阶段, 当并网点电压  $U < 0.2$  p.u. 时, 光伏电源不同于双馈风机脱网状态, 此时光伏电源需保持不脱网运行且需要向电网提供更多的无功电流, 用于支撑电压恢复。

结合分析以及实际故障录波数据, 归纳出光伏电源适应于工程实用化的电压电流映射关系为

$$I_{\text{PV}} = \begin{cases} 1.05I_N, & U < 0.2 \\ 1.1I_N, & 0.2 \leq U \leq 0.9 \\ I_N, & U > 0.9 \end{cases} \quad (17)$$

式中,  $I_{\text{PV}}$  为光伏电源的输出短路电流;  $U$  为并网点电压, p.u.。

#### 3.2 短路电流快速迭代计算方法

采用文献[6]所提故障区域迭代的方法。整体迭代思想流程如图6所示。首先, 通过传统短路电流计算方法得出故障前、后的各节点电压  $U_0$ 、 $U_1$ , 然后, 求其故障后节点电压相对于故障前电压的幅值比, 作为该节点  $x$  的电压跌落系数  $K_x$ 。设置一个区域临界电压  $U_{\text{lim}} \in (0, 0.9)$ , 若网络的  $K_x$  小于该临界电压, 即  $K_x < U_{\text{lim}}$ , 则认为该节点为故障相关节点而保留,

形成节点集合  $D_0$ 。再保留  $D_0$  中相互连通的节点形成节点集  $D$ , 其余独立的节点则不计入故障相关节点。

当故障区域节点集  $D$  中没有光伏电源节点时, 无需进行局部迭代计算; 当含有光伏电源节点时, 则进行故障区域局部迭代。首先形成节点集  $D$  的节点导纳矩阵  $Y_D$ , 然后通过  $U_1$  中节点集  $D$  所对应故障区域节点的节点电压  $U_D^{(x)}$  以及式(15), 求出光伏电源的输出电流  $I_{\text{PV}}^{(x)}$ , 并求出光伏电源节点注入电流的修正量:

$$\Delta I_{\text{PV}}^{(x)} = \begin{cases} I_{\text{PV}}^{(x)} - 1.2I_{\text{PV}}^{(0)}, & I_{\text{PV},i} = 1.2I_{\text{PV}}^{(0)} \\ I_{\text{PV}}^{(x)}, & I_{\text{PV},i} = 0 \end{cases} \quad (18)$$

式中,  $I_{\text{PV},i}$  为传统短路电流计算中光伏电源节点的注入电流;  $I_{\text{PV}}^{(0)}$  为光伏电源的额定电流。

利用  $Y_D$  在故障区域内局部迭代求解节点电压的修正量:

$$\Delta U_D^{(x)} = \frac{\Delta I_D^{(x)}}{Y_D} \quad (19)$$

式中,  $\Delta I_D^{(x)}$  为故障区域节点的注入电流列向量, 其中对应光伏电源节点的电流为  $\Delta I_{\text{PV}}^{(x)}$ , 非光伏电源节点的电流为0。

通过在故障区域内不断迭代计算, 直到满足收敛条件  $|\Delta U_D^{(x)} - \Delta U_D^{(x-1)}| < \epsilon$  或达到最大迭代次数  $L_0 = 10$ , 则停止迭代。而后将  $\Delta U_D^{(x)}$  与  $U_1$  对应节点相加, 得到最终节点电压  $U_i$ , 由此求得网络短路电流  $I_f$ 。

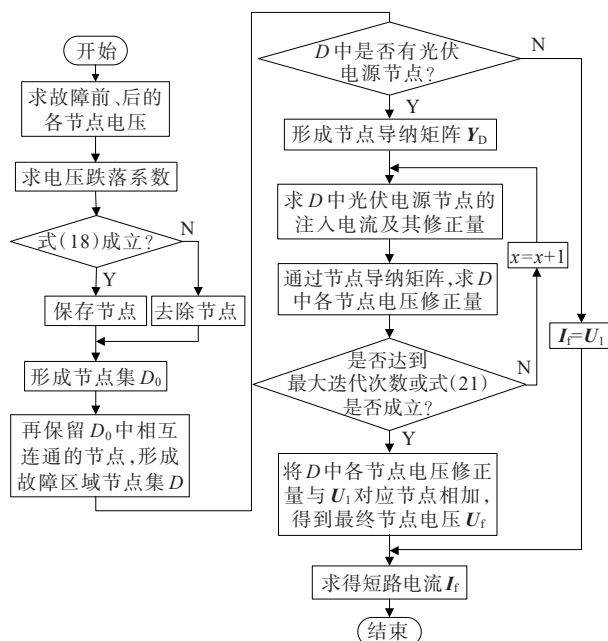


图6 局部迭代法流程

Figure 6 Flow chart of local iteration method

## 4 仿真实验及分析

### 4.1 仿真参数设置

为验证上述策略的有效性,基于 Matpower 计算程序中 IEEE 39 节点系统作为仿真场景,该系统含有 10 台发电机组、46 条输电线路,基准容量选定  $100 \text{ MV} \cdot \text{A}$ ,其结构如图 7 所示。

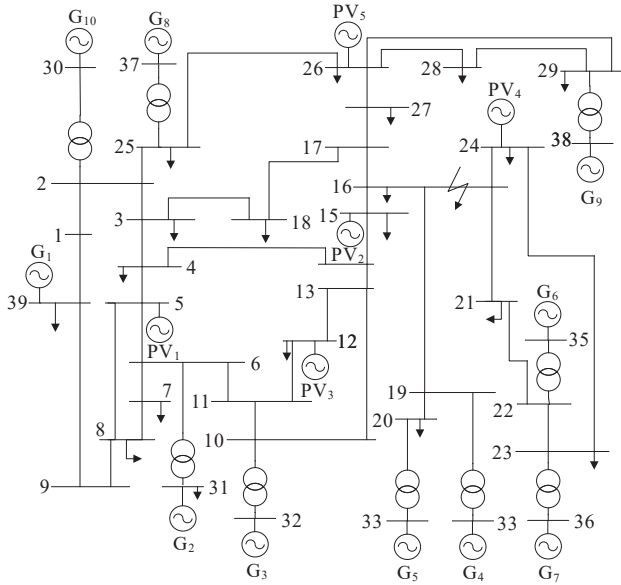


图 7 含光伏电源 IEEE 39 节点系统结构

Figure 7 Structure of IEEE 39 node system with photovoltaic power supply

设置母线 16 处发生三相对称接地故障,测量点位于母线 16 上。分别在第 5、15、21、24、26 节点接入基础容量均为 300 MW 的光伏电源,将全网的负荷波动分为 3 个层级,如文 2.2.1 中所述。再考虑线路开断导致的网络拓扑变化,根据基础 IEEE 39 节点系统结构选取 26-28、21-24、4-14、17-18 这 4 条线路作为备选开断线路,每次随机选择其中几条开断。通过物理模型随机生成 6 000 组特征标签数据集,并再对特征数据集进行预处理。在预处理完样本数据集后,预留出 20% (1 200 组) 的数据作为测试集,其余 4 800 组样本数据作为训练集。使用 10 折交叉验证,用 FNN、Bagging 和 SVR 算法分别进行训练数据,选取  $R^2$  和 RMSE 值最佳的模型作为最终数据驱动模型。

### 4.2 局部迭代法的误差分析

为了突出所选物理模型局部迭代法的精确度

优势,在 PSCAD/EMTDC 中搭建含光伏电源 IEEE 39 节点系统(图 7),设置 0.3 s 时母线 16 处发生三相金属性短路故障,短路电流实际波形如图 8 所示,通过计算得到发生故障后 40 ms 断路器分闸时刻处的电流有效值为 33 kA,而通过局部迭代法在相同场景下计算所得的短路电流为 32.8 kA,其结果误差是很小的。

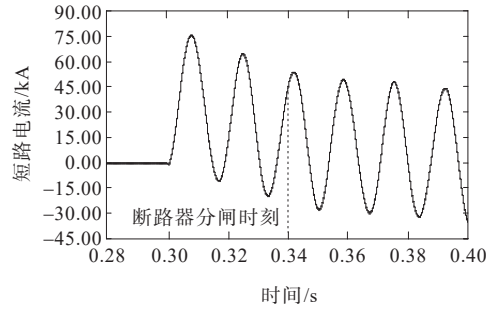


图 8 短路电流实际波形

Figure 8 Actual waveform of short-circuit current

### 4.3 算法精度分析与优选

对各算法结果进行分析,每种算法进行超参数优化后的表现如表 3 所示,可以看出,FNN 在训练集/验证集和测试集上的表现最好,SVR 的表现仅次于 FNN, Bagging 的表现稍差一点,但其误差也在允许的范围之内。

表 3 机器学习算法结果

Table 3 Results of machine learning algorithm

算法	训练集/验证集		测试集	
	$R^2$	RMSE	$R^2$	RMSE
Bagging	0.97	0.593 3	0.97	0.592 8
FNN	0.99	0.149 9	0.99	0.128 0
SVR	0.98	0.292 4	0.98	0.289 9

进一步将短路电流预测值和真实值绘制成散点图,如图 9 所示。对比 3 种 ML 算法模型,发现其预测-真实值分布基本上都能接近于对角直线,并且各个样本数据点距离对角线都在一个较小的范围内,说明各个样本的预测误差相对较小。Bagging 模型相较于 FNN 和 SVR 模型,其整体样本数据点更为分散,部分样本点离对角线的垂直距离较远,即预测不够精确。而 FNN 和 SVR 模型相对而言有更多的样本点落在对角线附近,即预测更为精确。综合考虑,最终选择精度最高的 FNN 模型作为数据驱动模型。

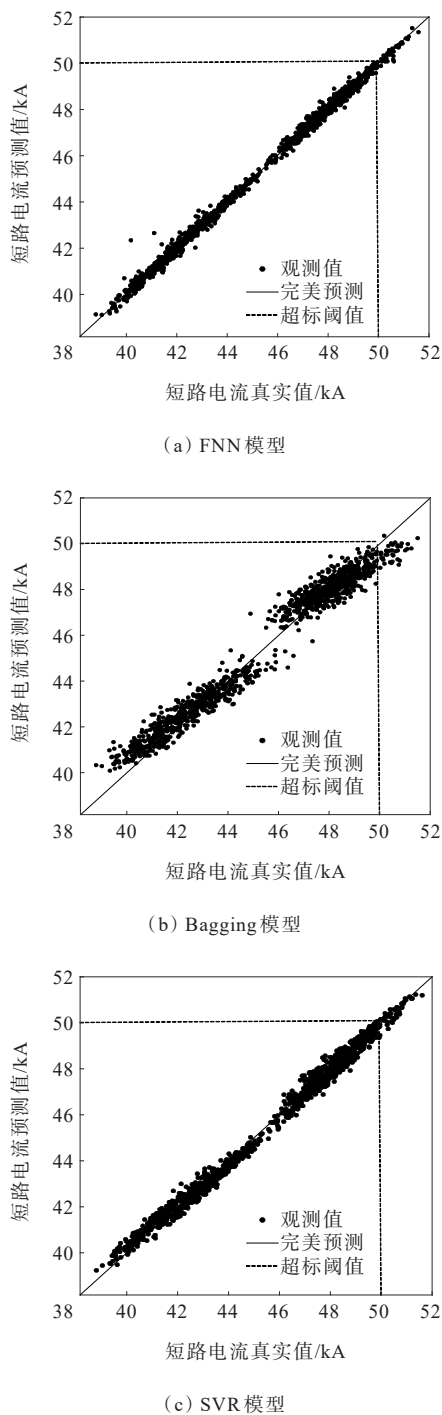


图9 模型预测值—真实值

Figure 9 Figure of predicted value vs. real value of the model

#### 4.4 短路电流超标判断视角下的算法性能分析

对于数据驱动模型,初步筛选超标短路电流样本的保守阈值整定,需要满足不能有FN的前提下尽可能控制FP,即误检率FPR为0且漏检率FNR尽可能小,避免给后续物理模型增加不必要的计算负担,保守阈值的确定通过仿真数据来决定,FNN模型下不同阈值整定结果如图10所示,随着阈值设

定的降低FPR越来越高,更多的实际未超标短路电流场景被FNN模型检测为超标;与此同时,FNR则越来越低,但阈值降低到49.6 kA时FNR为0,可以保证不会有超标的场景被FNN模型所遗漏,故将保守阈值整定为49.6 kA。

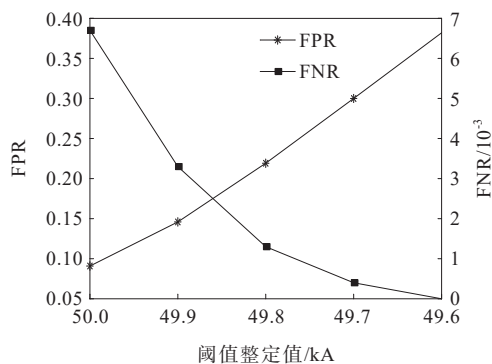


图10 不同阈值下的FNR与FPR

Figure 10 FNR and FPR under different thresholds

需要注意的是,本文所采用的含光伏电源的IEEE 39节点系统所计算的短路电流结果并没有达到一般断路器的极限开断电流50 kA这一量级,故将短路电流结果采取等比例放大6倍。接着使用所提方法对待预测的4 777组短路电流数据场景进行快速超标评估。为了体现阈值整定的必要性,将不设置保守阈值,即直接按超标阈值50 kA整定的结果与设定48 kA保守阈值进行对比,其结果如表4所示。

表4 数据—物理融合驱动短路电流超标评估结果

Table 4 Results of data-physical-fusion driven short-circuit current over-limit evaluation

FNN模型 阈值/kA	FNN模型初 筛后超标数	物理模型精校 后最终超标数	FNR	FPR
49.6	309	191	0.000 0	0.381 9
50.0	176	160	0.006 7	0.090 9

从表4可以看出,当阈值直接按断路器最大遮断电流50 kA整定时,则会遗漏31组超标场景,造成后续物理模型无法对这些遗漏场景进行再次校验;而当阈值整定为49.6 kA时,FNR为0,则不会出现遗漏的超标场景,能够准确地筛选出191组超标的场景。

为了进一步体现该方法在计算速度方面的优越性,现将其与单纯的物理建模方法进行对比,采

用的物理模型如文3中所述。通过实验研究发现4 777组场景物理模型完成计算校验需要67.15 s,而融合模型只需要4.65 s即可完成校验,主要在于数据驱动模型离线应用仅需0.11 s便完成超标短路电流场景的初筛,留待物理模型校验的场景数量极大缩减。2种方法均不会遗漏超标短路电流场景。

## 5 结语

在含新能源电网短路电流计算方面,针对目前数据驱动方法结果不够精确可靠和物理迭代模型面对大型复杂网络计算速度较慢的难题,本文提出数据-物理融合驱动的短路电流超标评估策略。通过含光伏电源的IEEE 39节点系统算例分析,对3种不同ML算法进行了对比分析,得到FNN模型为其中最佳的模型,且所提融合策略能够精确预测超标短路电流场景。并通过将数据-物理融合驱动模型与物理模型的计算时间进行对比,得到所提融合模型具有明显速度优势。多方面仿真验证了所提方法能够精确快速地实现含新能源电网的超标短路电流场景评估。

### 参考文献:

- [1] 魏繁荣,马啸,林湘宁,等.应对不可预知型短路电流超标的主动响应型站域保护策略[J].中国电机工程学报,2018,38(2):484-496.  
WEI Fanrong, MA Xiao, LIN Xiangning, et al. Strategy on the active response substation integrated protection for dealing with the unpredictable excessive short-circuit current[J]. Proceedings of the CSEE, 2018, 38(2): 484-496.
- [2] 杨振纲,李力,李扬絮,等.广东电网短路电流超标问题及对策[J].南方电网技术,2011,5(5):90-93.  
YANG Zhengang, LI Li, LI Yangxu, et al. The issue of short-circuit current being out of limitation in guangdong power grid and related countermeasures[J]. Southern Power System Technology, 2011, 5(5): 90-93.
- [3] 傅旭,李想,王笑飞.新能源发电接入对电网短路电流的影响研究[J].分布式能源,2018,3(1):58-63.  
FU Xu, LI Xiang, WANG Xiaofei. Short circuit current of electric power system with new energy power generation [J]. Distributed Energy, 2018, 3(1): 58-63.
- [4] 南东亮,王维庆,彭寅章,等.故障限流器在新能源并网中最佳安装位置和参数配置研究[J].太阳能学报,2022,43(1):307-312.  
NAN Dongliang, WANG Weiqing, PENG Yinzhang, et al. Research on optimum installation position and parameter configuration of fault current limiter in new energy grid[J]. Acta Energetica Solaris Sinica, 2022, 43(1): 307-312.
- [5] 郑超铭,陈义宣,李玲芳.大规模新能源接入对云南电网短路电流影响分析[J].云南电力技术,2021,49(4):21-23+28.  
ZHENG Chaoming, CHEN Yixuan, LI Lingfang, et al. Analysis of the impact of large-scale new energy integration on the short-circuit current of yunnan power grid[J]. Yunnan Electric Power, 2021, 49(4): 21-23+28.
- [6] 贾科,侯来运,毕天姝,等.基于故障域局部迭代的工程实用化新能源短路电流计算[J].电力系统自动化,2021,45(13):151-158.  
JIA Ke, HOU Laiyun, BI Tianshu, et al. Practical short-circuit current engineering calculation for renewable energy based on local iteration in fault area[J]. Automation of Electric Power Systems, 2021, 45(13): 151-158.
- [7] 匡晓云,方煜,关红兵,等.适用于含新能源逆变电源网络的全时域短路电流计算方法[J].电力自动化设备,2020,40(5):113-122.  
KUANG Xiaoyun, FANG Yu, GUAN Hongbing, et al. Full-time domain short circuit current calculation method suitable for power network with inverter-interfaced renewable energy source[J]. Electric Power Automation Equipment, 2020, 40(5): 113-122.
- [8] 刘慧媛,肖繁,张哲,等.新能源电源接入不平衡配电网的短路计算方法[J].电力系统自动化,2019,43(21):177-186.  
LIU Huiyuan, XIAO Fan, ZHANG Zhe, et al. Short-circuit calculation method for unbalanced distribution network with integration of renewable energy[J]. Automation of Electric Power Systems, 2019, 43(21): 177-186.
- [9] 乔黎伟,王静怡,郭炜,等.基于随机森林算法的中短期用电量预测[J].电力科学与技术学报,2020,35(2):150-156.  
QIAO Liwei, WANG Jingyi, GUO Wei, et al. Medium and short term electricity demand prediction based on random forests algorithm. Journal of Electric Power Science and Technology, 2020, 35(2): 150-156.
- [10] 赵洋,王瀚墨,康丽,等.基于时间卷积网络的短期电力负荷预测[J].电工技术学报,2022,37(5):1242-1251.



- ZHAO Yang, WANG Hanmo, KANG Li, et al. Temporal convolution network-based short-term electrical load forecasting[J]. Transactions of China Electrotechnical Society, 2022, 37(5): 1242-1251.
- [11] 厉仄平, 伍冲翀, 熊来红, 等. 基于机器学习和雷电感应电压的输电线路雷击定位方法[J]. 高压电器, 2022, 58(12): 109-116.
- LI Zeping, WU Chongchong, XIONG Laihong, et al. Lightning location method for transmission lines based on machine learning and lightning induced voltage[J]. High Voltage Apparatus, 2022, 58(12): 109-116.
- [12] 周楠, 张平, 郑征, 等. 基于机器学习的电力通信网带宽分配算法[J]. 电网与清洁能源, 2021, 37(5): 67-73.
- ZHOU Nan, ZHANG Ping, ZHENG Zheng, et al. Bandwidth allocation algorithm for power communication network based on machine learning[J]. Power System and Clean Energy, 2021, 37(5): 67-73.
- [13] YANG L T, YANG H G. Analysis of different neural networks and a new architecture for short-term load forecasting[J]. Energies, 2019, 12(8): 1433-1455.
- [14] LIU T X, JIN Y, GAO Y Y. A new hybrid approach for short-term electric load forecasting applying support vector machine with ensemble empirical mode decomposition and whale optimization[J]. Energies, 2019, 12(8): 1520-1539.
- [15] 王子晗, 高红均, 高艺文, 等. 基于深度强化学习的城市配电网多级动态重构优化运行方法[J]. 电力系统保护与控制, 2022, 50(24): 60-70.
- WANG Zihan, GAO Hongjun, GAO Yiwen, et al. Multi-level dynamic reconfiguration and operation optimization method for an urban distribution network based on deep reinforcement learning[J]. Power System Protection and Control, 2022, 50(24): 60-70.
- [16] 任柯政, 徐泰山, 郭瑾程, 等. 数据驱动的新型电力系统断面极限估算方法[J]. 智慧电力, 2022, 50(11): 9-16.
- REN Kezheng, XU Taishan, GUO Jincheng, et al. A new data-driven method for estimating the section limit of power system[J]. Smart Power, 2022, 50(11): 9-16.
- [17] 郑翔, 王慧芳, 姜宽, 等. 机理与数据融合驱动的含IIDG配电网短路电流计算方法[J]. 电力自动化设备, 2021, 41(1): 41-48.
- ZHENG Xiang, WANG Huifang, JIANG Kuan, et al. Calculation method of short circuit current in distribution network with IIDG driven by mechanism and data fusion[J]. Electric Power Automation Equipment, 2021, 41(1): 41-48.
- [18] 李峰, 王琦, 胡健雄, 等. 数据与知识联合驱动方法研究进展及其在电力系统中应用展望[J]. 中国电机工程学报, 2021, 41(13): 4377-4390.
- LI Feng, WANG Qi, HU Jianxiong, et al. Combined data-driven and knowledge-driven methodology research advances and its applied prospect in power systems[J]. Proceedings of the CSEE, 2021, 41(13): 4377-4390.
- [19] WANG Q, LI F, TANG Y, et al. Integrating model-driven and data-driven methods for power system frequency stability assessment and control[J]. IEEE Transactions on Power Systems, 2019, 34(6): 4557-4568.
- [20] 张涵, 王程, 毕天姝. 融合物理与数据知识的电力系统扰动后频率在线快速计算方法[J]. 电网技术, 2022, 46(11): 4325-4335.
- ZHANG Han, WANG Cheng, BI Tianshu. On-line fast frequency calculation method after power system disturbance based on fusion of physics and data knowledge[J]. Power System Technology, 2022, 46(11): 4325-4335.
- [21] ZIMMERMAN R D, MURILLO-SANCHEZ C E, THOMAS R J. MATPOWER: steady-state operations, planning and analysis tools for power systems research and education[J]. IEEE Transactions on Power Systems, 2011, 26(1): 12-19.