

基于 gcForest-AdaBoost 的光伏储能并网系统 暂态稳定评估

王雨辰¹, 李忠艳¹, 苏 涵²

(1. 华北电力大学数理学院, 北京 102206; 2. 华北电力大学控制与计算机工程学院, 北京 102206)

摘 要: 针对当前光伏储能并网系统暂态稳定评估方面存在的预测精度不高问题, 结合深度学习与集成学习技术, 提出一种基于集成深度随机森林算法(gcForest-AdaBoost)的光伏储能并网系统暂态稳定评估方法。首先, 依据光伏储能并网系统运行过程中的稳态分量构建初始输入特征; 其次, 为减少模型过度拟合, 保证级联森林层数增加后模型仍具有较强学习能力, 结合深度随机森林算法以及 AdaBoost 算法, 提出基于 gcForest-AdaBoost 的暂态稳定评估方法; 再次, 利用输入特征集训练 gcForest-AdaBoost 模型, 并建立光伏储能并网系统暂态稳定评估模型; 最后, 在 IEEE 39 节点系统接入光伏储能机组构建算例系统展开仿真分析和数据采集, 得出评估结果。算例分析表明, 所提模型能对光伏储能并网系统暂态稳定展开有效分析, 且经过验证发现所提模型具有较好的鲁棒性和泛化能力。

关键词: 光伏发电; 储能; 暂态稳定; 随机森林; gcForest-AdaBoost; 输入特征

中图分类号: TM712

文献标志码: A

0 引 言

伴随光能、风能等可再生能源快速增长, 能源供给侧结构性改革日益深化, 促使构建以新能源为主的新型电力系统成为推动能源绿色转型、助力“双碳”目标的关键举措^[1]。然而, 光伏作为新型电力系统的主力, 存在功率输出波动较大、弃光现象严重的特征, 大规模并网后将改变传统电网拓扑结构, 使得传统机组动态特性更为复杂。对此, 2024 年 8 月, 国家发展改革委等部门联合发布《加快构建新型电力系统行动方案(2024—2027 年)》, 要求开展电力系统稳定保障行动, 提升系统安全稳定运行水平。而暂态稳定作为评估电力系统安全稳定运行的关键指标^[2], 集中反映电力系统对外界扰动的响应能力。因此, 在新型电力系统下, 如何快速、准确地对光伏储能并网系统展开暂态稳定评估, 已成为推动电网系统安全运行、实现碳达峰目标的重要着力点。

学术界针对电力系统暂态稳定展开广泛研究, 取得诸多有益成果。目前, 传统暂态稳定评估主要有直接法和时域仿真法。直接法又称解析法, 其优点在于计算速度快, 能迅速给出系统稳定性判断, 但对于复杂系统的分析能力有限, 难以全面考虑所有可能的暂态行为^[3]。时域仿真法的直观性较强^[4], 但计算量大, 对于大规模电力系统仿真效率较低, 在复

杂的现代电力系统中适用性较低。文献[5]采用时域仿真法研究低电压穿越控制参数对系统暂态稳定的影响, 得出相关优化参数。但此方法难以准确刻画控制参数的最优空间, 因而难以为光伏储能并网系统暂态稳定评估提供理论指导。

近年来, 随着人工智能的广泛应用, 机器学习算法凭借其快速、准确的特性在光伏储能并网系统暂态稳定评估中日益得到应用。目前, 用于暂态稳定评估的机器学习算法可根据其结构和复杂度大致分为浅层学习和深度学习两大类。决策树、支持向量机等浅层学习算法已在暂态稳定评估中得到成熟运用^[6]。但这些模型由于参数较少、结构较为简单, 更多适用于数据量适中、特征间关系相对简单的暂态稳定评估, 复杂成分的泛化性能相对较差^[7]。

卷积神经网络、长短时记忆网络等深度学习算法能处理图像、音频信号等高维度数据, 在数据集大且包含复杂的模式电力系统暂态稳定评估中展现出良好评估性能。文献[8]在光伏和风电场控制器中应用 PI 控制器、电力系统稳定和柔性交流输电控制器, 采用自适应速度更新松弛粒子群算法(adaptive velocity updalte relaxation particle swarm optimization, AVURPSO)、引力搜索算法(gravitational search algorithm, GSA)和遗传算法(genetic algorithm, GA)搜索最优控制器参数, 有效评估了电力系统暂态稳定情况; 文献[9]利

用长短时记忆网络算法构建新能源微电网暂态稳定评估模型,实现暂态稳定的准确评估,但该方法仅能捕捉时序关联,对空间拓扑特征的捕捉存在一定欠缺;文献[10]在 Bi-LSTM 网络中引入注意力机制识别失稳样本,通过引入焦点损失函数调整模型训练的偏向性,研究发现此模型具有较好的性能。总体而言,上述计算方法本质上均属于深度神经网络(deep neural network, DNN),与浅层机器学习相比能对高维异构的多源信息数据进行高效处理,快速实现系统暂态稳定评估。但需人为设置较多参数,花费时间成本较大,一定程度上对暂态稳定评估造成影响。

深度森林是基于深度学习模型提出的一种新算法,凭借较少的计算资源和训练时间,在处理复杂分类任务方面表现出媲美或超过 DNN 的优越性能^[11]。文献[12]构建级联深度森林模型,对暂态电压稳定展开评估,发现模型预测精度较高;文献[13]对深度森林模型进行改进,提出 gcForest 模型评估风电并网系统暂态稳定,发现该方法在快速性与准确性方面表现优异。

综上,既有机器学习算法尽管在暂态稳定领域已取得广泛应用成果,但也存在计算成本高、数据需求量大、训练时间长、可解释性差等劣势,在一定程度上限制了其在暂态稳定评估中的有效性。而深度森林模型作为一种模仿深度学习架构设计的集成学习框架,结合了层级结构和多种粒度的特征组合,能捕捉到复杂的非线性关系,表现出高效且准确的预测能力,能在电力系统暂态稳定中表现出良好的预测能力。在光伏储能并网系统暂态稳定评估中,结合深度随机森林(gcForest)与自适应提升(AdaBoost)双重优势,引入 gcForest-AdaBoost 模型展开暂态稳定分析。与其他深度学习算法相比较而言,gcForest-AdaBoost 模型凭借多粒度级联结构和自适应权重机制,仅需较少量级的样本即可达到深度学习模型数万样本的精度水平,在光伏储能并网系统暂态稳定评估中能采用相对较少的数据达到评估效果。同时,在工程易用性方面,深度森林通过超参数的极简配置显著优于深度学习的复杂调参体系,可有效缩短光伏储能并网系统暂态稳定评估的调参耗时。相较于 LightGBM 等集成方法而言,现有集成算法受限于贪心分裂策略,多应用于静态或准稳态场景。而光伏储能并网系统的暂态过程呈现毫秒级突变和分钟级能量动态的强耦合特性,故 gcForest-AdaBoost 模型能通过多时间尺度特征融合能力,突破静态建模局限,有效提升检测灵敏度。此外,LightGBM 的特征重要性指标仅反映统计分析增益,难以关联电力系统物理量,而 gcForest-AdaBoost 通过构建“数据-物理”双向映射则能解决此种工程可解释性差的痛点。

进一步地,既有研究大多针对传统电力系统暂态稳定展开评估,对含有新能源电力系统的仿真分析相对较少。事实上,随着“双碳”目标的深入推进,加速推动能源结构绿色低

碳转型日益成为经济社会发展的刚性需求与必由之路,势必会带动高比例新能源并网。而新能源并网运行将在一定程度上产生能源消纳与供应矛盾、安全稳定运行难度激增等挑战,因而对含有新能源的电力系统进行暂态稳定分析具有较强必要性。对此,本文将光伏储能设备引入电网系统,基于 gcForest-AdaBoost 模型展开光伏储能并网系统暂态稳定评估。具体地,依据光伏储能并网系统运行特点构建初始输入特征,提出一种面向光伏储能并网系统暂态稳定评估的 gcForest-AdaBoost 评估模型,并使用该模型进行特征选择。在此基础上,在 IEEE39 节点系统接入光伏储能机组构建算例系统展开仿真分析,通过多种算例验证以及与其他模型的对比分析,验证发现所提模型具有较强精准性与可靠性,能较好用于光伏储能并网系统暂态稳定评估。

1 光伏储能并网系统的电网结构

光伏发电的功率输出具有间歇性和波动性,易造成电网电压的波动闪变。而储能系统通过对光伏发电系统电力进行调峰调谷,能对提高整体电力系统稳定性和可靠性产生积极影响。因此,在新能源大规模并网的当下,光伏储能系统接入大电网中,会对系统暂态稳定造成显著影响。在光伏集群中配置储能系统并接入大电网,可实现功率双向流动,保证系统平稳运行,便于对集群实现调节控制。图 1 为光伏储能并网系统的电网结构图。

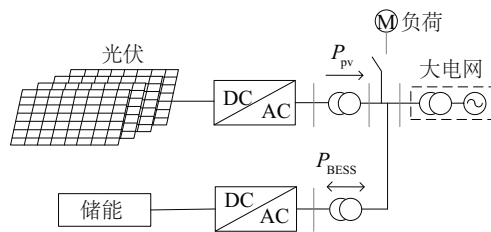


图1 光伏储能并网系统的电网结构

Fig. 1 Grid structure of photovoltaic energy storage grid-connected system

2 暂态稳定评估模型

2.1 集成深度随机森林算法

随机森林算法是一种流行且强大的机器学习算法,在处理大量数据和特征时具有出色表现。但由于随机森林模型的复杂度较高,成本和内存消耗较大,预测速度较慢。为提高随机森林算法的预测精度,本文将深度学习技术以及集成学习技术引入随机森林算法中,提出一种集成深度随机森林算法,对光伏储能并网系统暂态稳定作出评估。

2.1.1 随机森林算法

随机森林算法(random forest, RF)是由 Breiman 提出的一种集成学习模型,其原理在于建立多个决策树对样本进行独

立预测^[14],并基于所有决策树预测结果投票或取平均数得出预测结果。具体算法步骤为:

- 1) 从原始数据集中随机抽取样本,构造多个子数据集。
- 2) 对每一个子数据集构造一个决策树,并随机选择 n 个属性,随机选择最优属性进行决策树节点分裂,训练决策树。
- 3) 重复以上步骤,直至生成指定数量的决策树,构建随机森林模型。

2.1.2 深度随机森林算法

深度随机森林算法(gcForest)是由周志华等^[15]提出的由极端随机森林和随机森林两种算法组成的集成学习算法,包括多粒度扫描(multi-grained scanning)和级联森林(cascade forest)两部分,具体结构如图2所示。gcForest算法能通过随机森林和极端森林交叉验证避免过于拟合,基本调节参数较为简单,可取得良好计算效果。

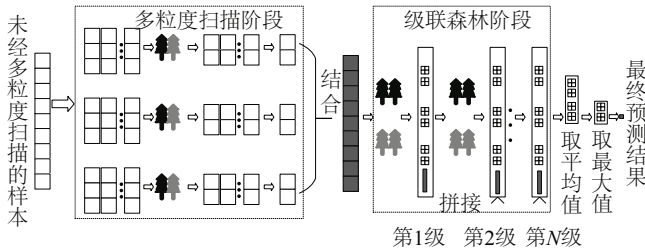


图2 gcForest模型结构

Fig. 2 gcForest model structure

1) 多粒度扫描阶段。gcForest算法利用多粒度扫描技术来最大化地提取数据特征的序贯信息,从而增强级联森林结构的预测性能。具体流程为:首先,采用一个长度为 L 的滑动窗口并设置步长为 1,对输入的 m 维特征向量样本进行处理,得到 K 个特征子样本。其次,将这些子样本分别输入到多种类型的森林模型中,以获取对应于不同类别的概率向量。最后,将分类向量按顺序 D 维的特征向量输入到级联森林模型中,强化级联森林。

2) 级联森林阶段。级联森林由多层随机决策树森林级联堆叠构成,每一层包含两个完全随机森林和两个随机森林。为提升算法分类准确度,gcForest算法将经过多粒度扫描技术处理过的特征向量输入级联森林模型,以对特征向量展开逐层分析。具体流程为:首先,将 D 维的特征向量输入到级联森林模型,经过处理后得到增强特征向量。其次,将 D 维特征向量与增强后的特征向量进行连接,然后输入到下一个层级中。再次,重复步骤 1) 和步骤 2),并将数据传递至最后一层。最后,将最后一层分类向量取平均值,并选取最大值对应输出最终预测结果。

综上,gcForest算法的整体步骤为:

- 1) 对样本数据进行预处理,借助插值法补齐缺失数据,并对无效数据进行剔除。将处理后的数据集分为训练集和测试集。

2) 将训练集样本输入 gcForest 算法中进行训练。首先将输入多粒度扫描算法中获取特征向量,进而输入级联森林模型中构建 gcForest 模型。

3) 利用测试集对 gcForest 模型进行检验。

2.1.3 集成深度随机森林算法

gcForest模型本质上是对卷积神经网络模型(convolutional neural network, CNN)的优化,但易产生过多的网络层数造成模型过度拟合。对此,为提升模型训练精度,参照文献[16]的做法,在 gcForest 算法引入 AdaBoost 技术,构建集成深度随机森林算法(gcForest-AdaBoost),以进一步对模型进行优化,从而实现对光伏储能并网系统暂态稳定进行评估。算法过程如下:

1) 初始化数据集的每个数据权重,样本集为 X , $X = \{x_1, \dots, x_n\}$; 样本数量为 n ; 样本初始化权重分布为 $D_1 = (m_{1,1}, m_{1,2}, \dots, m_{1,n})$; 样本初始化权重 $m_1(i) = 1/n$ 。

2) 设置迭代轮次为 t 次, $t = 1, 2, \dots, T$, 基于 D_1 生成弱分类器为 $h_t = f_{\text{gcForest}}(X, D_t)$ 。

3) 此轮迭代产生的弱分类器 h_t 在当前 D_t 下加权错误率为:

$$\varepsilon_n = \sum_{i: p_n(x_i) \neq y_i} m_n(i) \quad (1)$$

4) 计算 h_t 在强分类器中集合中的加权系数:

$$a_t = \frac{1}{2} \ln \left(\frac{1 - \varepsilon_n}{\varepsilon_n} \right) \quad (2)$$

5) 为下次迭代更新样本分布:

$$m_{t+1}(i) = m_{t,i} e^{-a_t y_i^2} \quad (3)$$

6) 结果输出最终强分类器,即集成深度随机森林(gcForest-AdaBoost):

$$f_{\text{gcForest-AdaBoost}}(x) = \text{round} \left[\sum_{t=1}^T a_t G(x) \right] \quad (4)$$

2.2 电力系统 TSA 实施流程

本文结合深度随机森林算法以及 AdaBoost 算法,提出基于 gcForest-AdaBoost 的光伏储能并网系统暂态稳定评估方法。图3为该方法的实施流程图,具体步骤如下:

1) 利用电力系统仿真软件设置不同故障类型展开时域仿真,以获取光伏储能并网系统暂态稳定数据。

2) 借助暂态稳定数据构建训练集与测试集。

3) 初始化 AdaBoost 弱分类器中参数。

4) 训练弱分类器 gcForest。对原始数据集中的特征向量按由高到低进行重要性排序,并通过依次加入 TSA 模型来获得在不同输入特征维度下的评估准确率。

5) 直到迭代结束,输出强分类器 gcForest-AdaBoost。

6) 输入暂态稳定数据测试集,利用训练好的强分类器 gcForest-AdaBoost 对当前光伏储能并网系统暂态稳定进行评估,输出相应结果。

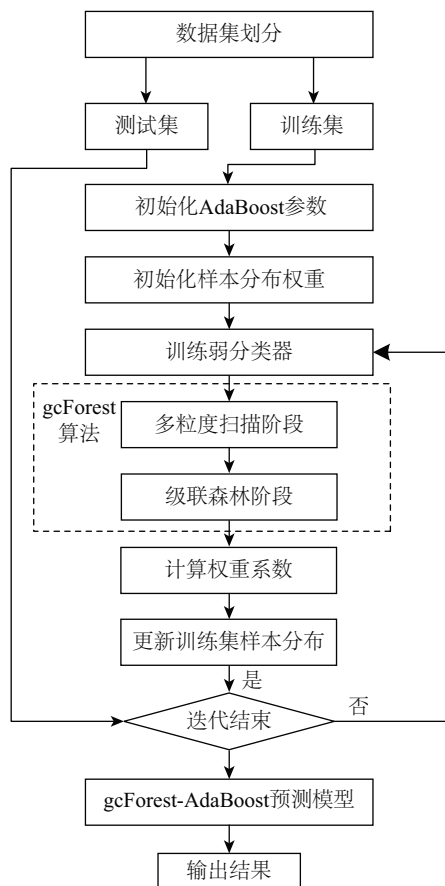


图3 电力系统TSA方法流程

Fig. 3 Flowchart of power system TSA method

3 暂态稳定评估流程

3.1 暂态稳定判定

光伏储能并网系统暂态稳定主要指的是系统在受到较大干扰时的功率稳定情况。同时引入暂态功角稳定性和暂态电压稳定性作为暂态稳定的评估指标,可更好、更全面地描述系统的稳定状态。暂态功角稳定性采用暂态功角稳定指数进行判定,参见文献[17],具体公式为:

$$T = \frac{360^\circ - |\Delta\delta_{\max}|}{360^\circ + |\Delta\delta_{\max}|} \quad (5)$$

式中: $\Delta\delta_{\max}$ ——仿真时间内任意两台发电机的最大相对功角差, ($^\circ$);若 T 大于 0,则可判定系统稳定,可输入标签 1,代表稳定样本,若 T 小于 0,则可判定系统失稳,可输入标签 0,代表失稳样本。

暂态电压稳定性实用判据为:故障后母线节点电压低于 0.75 pu 的时间不超过 1 s。条件满足则判定为稳定节点,样本标签标记为 1;条件不满足则视为失稳节点,样本标签标记为 0。每个样本可到 $1 \times m$ 维的标签。

3.2 输入特征的构建

光伏储能并网系统的暂态信息主要是发生故障时体现

的时序特征。因此,特征集作为模型的输入内容,是评估结果准确性的直接决定因素。本文出于物理意义、全面性以及便于计算性考虑,参考现有文献[18],从时间、空间、系统规模角度选择电压幅值、电压相角、有功功率、无功功率等易测量的原始特征作为输入特征,详见表 1。

表1 输入初始特征

Table 1 Initial input characteristics

特征序号	特征描述
1	母线电压幅值
2	母线电压相角
3	光伏储能并网电压
4	光伏储能电站有功功率
5	光伏储能电站无功功率
6	发电机有功
7	发电机角速度
8	发电机端电压
9	发电机端电流
10	发电机电磁功率
11	发电机有功功率
12	发电机无功功率
13	母线负荷有功功率
14	母线负荷无功功率
15	交流线有功功率
16	交流线无功功率
17	交流线电压
18	交流线相角

3.3 暂态稳定评估过程

基于上述分析,提出基于 gcForest-AdaBoost 的光伏储能并网系统暂态稳定评估模型,具体包括离线训练、在线评估和模型更新 3 个部分。

3.3.1 离线训练

1)生成离线样本。基于光伏储能并网系统历史运行数据,借助仿真系统设置一系列短路故障展开暂态稳定仿真研究,获取大量样本用于训练 gcForest-AdaBoost 模型。其中,每个样本中均包含光伏储能并网系统运行特征,通过对数据进行整理,打上相应稳定与不稳定标签。

2)训练离线模型

将输入特征随机打乱,并按 4:1 的比例划分为训练集和测试集,前者用于训练模型,后者用于评估模型性能。随后,

将这两个数据集输入模型进行训练,以完成 TSA 综合模型的构建。接着,输出评价结果完成光伏储能并网系统暂态稳定评估。

3.3.2 在线评估

在系统某个运行状态实时采集同步相量测量装置 (phasor measurement unit, PMU) 量测数据,结合 gcForest-AdaBoost 模型进行特征提取,并提供相关提取结果。若结果显示失稳,则系统会提示相关工作人员采取预防措施,避免发生大规模故障。

3.3.3 模型更新

考虑到光伏储能并网系统运行状况较为复杂,需设计模型更新机制以提升模型对系统运行状态变化的鲁棒性。

3.4 模型评估指标

为使得光伏储能并网系统暂态稳定评估更加全面,引入混淆矩阵。基于混淆矩阵加入相关评价指标,具体包括反映模型全局正确率的准确率(E_A)、能预测失稳样本占真实失稳样本比例的召回率(E_R)、预测失稳样本比例的查准率(E_P)以及可良好反映模型对失稳样本预测能力的 F_1 分数。上述评价指标的计算方式为:

$$E_A = \frac{E_{TP} + E_{TN}}{E_{TP} + E_{TN} + E_{FP} + E_{FN}} \quad (6)$$

$$E_R = \frac{E_{TN}}{E_{TN} + E_{FP}} \quad (7)$$

$$E_P = \frac{E_{TN}}{E_{TN} + E_{FN}} \quad (8)$$

$$F_1 = \frac{2E_R E_P}{E_R + E_P} \quad (9)$$

表 2 为 TSA 混淆矩阵。其中,TP 表示真稳定,FN 表示假稳定;FP 表示假失稳;TN 表示真失稳。

表 2 TSA 混淆矩阵

Table 2 Confusion matrix of TSA

评估模型	真实结果	
	暂态稳定	暂态失稳
暂态稳定	TP	FP
暂态失稳	FN	TN

4 算例分析

利用 PSASP 仿真平台建立 IEEE 39 节点系统作为仿真模型展开算例分析。该节点系统包括 10 个传统机组,39 条母线和 46 条传输线路。系统基准功率是 100 MVA,额定频率为 50 Hz。为测试光伏储能并网系统暂态稳定情况,在节点系统中接入光伏储能机组。

4.1 算例数据

以光伏储能并网系统历史运行数据为基础,通过仿真建

立初始数据集。在仿真过程中,设置负荷水平以 10% 为步长,负荷水平在 80%~130% 的区间内变化。在传输线路中设置故障类型为三相短路故障,故障发生时刻是 1 s,故障切除时刻为 3 s。故障位置分别设置在距离输电线路首端 1%、25%、50%、75% 以及 99% 处。基于上述设置内容,共计生成 5230 个有效样本,其中稳定样本为 3923 个,失稳样本有 1307 个。进一步地,对有效样本进行标记,随机选取其中的 4000 个样本作为训练集,剩余样本作为测试集,共计 1230 个样本。每个样本均包含表 1 中的 18 个输入特征。

4.2 数据预处理

数据预处理是机器学习模型构建的核心环节,直接影响模型的收敛速度、泛化能力和最终性能。采取以下方式对数据进行处理。1)采用插值法补齐缺失数据,以提升模型性能;2)滑动窗口内取均值进行滤波,以消除异常值;3)为避免数据量纲干扰评估模型,采用 min-max 归一化方法处理不同量测数据,以使得样本数据处于相同数量级。min-max 的表达式为:

$$X^* = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (10)$$

式中: X^* ——归一化处理后的样本数值; X ——样本数据; $X_{\max}$ 和 $X_{\min}$ ——样本数据的最大值和最小值。

4.3 电压幅值对比

由于本研究旨在考察基于 gcForest-AdaBoost 的光伏储能并网系统的暂态稳定情况,故在展开算例分析前,先对 IEEE 39 节点系统在引入光伏储能系统前后的电压幅值进行对比分析。图 4 为引入光伏储能系统前后的电压幅值对比。分析可知,在引入光伏储能系统前,3 项故障发生时刻 1 s 时,节点电压迅速下降,在故障切除时刻 3 s 时,电压迅速升高,并在后续时刻产生波动直至恢复正常值;在引入光伏储能系统后,电压幅值明显增大,但在 3 项故障发生时刻以及故障切除时刻电压变化趋势与未引入光伏储能系统时基本相同,最终也趋于稳定。

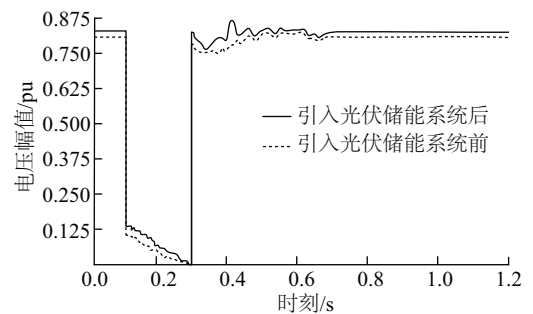


图 4 电压幅值对比

Fig. 4 Voltage amplitude comparison

4.4 不同评估模型性能对比

为验证本文所提 gcForest-AdaBoost 模型对光伏储能并网

系统暂态稳定的评估性能,引入其他几种常用的暂态稳定评估模型展开对比分析。选取的模型主要有决策树(decision tree, DT),算法采用的是 CART 算法,最大深度设置为 10; CNN 应用 3×3 的卷积核进行卷积操作,输出特征图的尺寸是 2×2;人工神经网络(artificial neural network, ANN),利用 Adm 算法进行训练,神经元数量设置为 100 个;深度置信网络(deep belief network, DBN),初始学习率设置为 0.001;支持向量机(support vector machines, SVM),采用高斯核函数解决分类问题。上述模型的暂态稳定相关测试结果如表 3 所示。LightGBM(light gradient boosting machine),默认迭代次数为 100 次,设置 200 个决策树作为基学习器。

表3 不同模型的测试结果

Table 3 Test results of different models

模型名称	准确率/%	召回率/%	差准率/%	F_1 分数
gcForest-AdaBoost	98.56	97.82	97.41	98.25
LightGBM	98.23	97.49	96.99	97.82
DT	96.51	96.12	95.93	96.17
CNN	97.46	96.87	96.71	97.45
ANN	96.84	96.76	96.08	96.44
DBN	97.13	96.34	96.26	96.98
SVM	96.08	95.73	95.42	95.86

表 3 中,LightGBM 是与本文所提模型一样的集成学习模型,其余为个体学习模型。可看出,LightGBM 模型和本文提出的 gcForest-AdaBoost 模型均较个体学习模型具有更好的评估性能,这表明在输入特征数据的挖掘与分析能力上,集成学习模型优于个体学习模型。同时,相较于其他模型,gcForest-AdaBoost 模型在各方面测试中均表现优异,评估精度明显高于其他模型,这主要得益于 gcForest-AdaBoost 对原有模型进行多方面改进,因而展现出良好的评估效果。在个体学习模型中,CNN 作为得到广泛应用的深度学习模型,擅长提取数据的局部特征,数据挖掘能力较强,因此也展现出良好的评估性能。DBN 作为一种深度学习模型,具有逐层训练的特性,能有效处理大规模数据集,展现出良好测试性能。DT 和 ANN 整体的评估数据弱于 CNN,原因是模型结构相对简单,对庞大数据的非线性映射关系挖掘能力不强。SVM 是一种浅层学习个体,结构较为简单,因而整体评估性能较低。

4.5 数据处理速度

在电力系统实际运行过程中,PMU 装置以 50 次/s 的频率采集数据^[19],因此需所提暂态稳定评估模型对单个样本的处理时间小于 0.02 s,才能达到对电力系统的良好监测。基于此,进一步考察本文所提 gcForest-AdaBoost 模型是否具备时效性,表 4 给出了模型的数据处理速度测试结果。分析显

示,gcForest-AdaBoost 模型在训练和评估阶段对单个样本的处理时间均短于 PMU 量测数据所需的最小处理时间 0.02 s。这说明,本文所提 gcForest-AdaBoost 模型能有效实现对光伏储能并网系统的实时监测。

表4 模型的数据处理速度

Table 4 Data processing speed of model

测试系统	模型训练阶段		模型评估阶段	
	样本数量	时间/s	样本数量	时间/s
IEEE 39 节点系统	4000	36.72	1230	0.041

4.6 鲁棒性测试

在光伏储能并网系统实际运行过程中,PMU 量测装置采集数据时可能受到造成干扰或数据缺失情况。因此,开展噪声干扰测试和数据缺失测试,对系统鲁棒性展开分析。

4.6.1 噪声干扰测试

向初始数据集的特征中加入噪声干扰,以开展噪声鲁棒性测试。在噪声的选择和设置方面,考虑到高斯白噪声便于量化和评价算法效果,因此为更好模拟未知的真实噪音,将 5 种信噪比的高斯白噪声信号输入光伏储能并网系统中,模拟实际数据受到干扰的情形。同时,为更加直观体现本文所提 gcForest-AdaBoost 模型在受到噪声干扰后的性能,将其他 6 个模型一同进行分析。图 5 为不同模型受噪干扰后的精确度。分析图 5 可知,gcForest-AdaBoost 模型在受到不同噪声影响情况下,仍表现出较高评估精度,说明相较于其他模型,本文所提模型可靠性较强,在筛选含有噪声干扰的电气特征方面具有较强适用性。

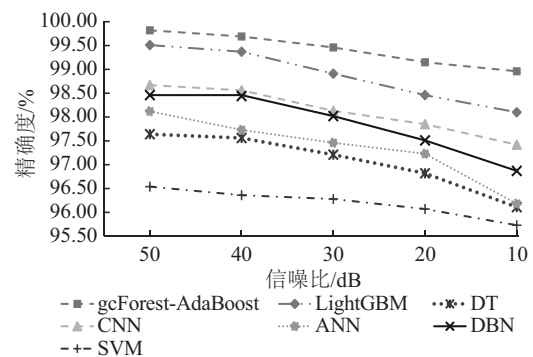


图5 不同模型受噪后的精确度

Fig. 5 Accuracy of different models after noise exposure

4.6.2 数据缺失测试

由于传感器故障、通信间断、通信延迟等在电力系统实际运行过程中时常发生,导致量测装置采集到的电力系统数据会存在一定缺失^[20]。此情形下,验证暂态稳定评估模型在数据缺失时的鲁棒性具有较强必要性。基于此,本文借助文献[21]的做法,随机丢弃数据集的某些母线运行数据,设

置 6 种数据缺失场景,并构建对应测试集展开模拟分析。表 5 为数据缺失测试结果。分析可知,在母线有所缺失且缺失数量逐步增加的情形下,gcForest-AdaBoost 模型准确率虽有所下降,但仍处于较高水平,说明模型仍然具备较好评估性能。特别是在母线 2、母线 14 和母线 19 均缺失场景下,gcForest-AdaBoost 模型准确率仍达到 97% 以上,仅较完整数据情形下的准确率低 1.14%。以上分析表明,文章所提 gcForest-AdaBoost 模型在数据缺失情形下仍具有一定鲁棒性,对光伏储能并网系统暂态稳定能给予精度较高的评估结果。

表 5 数据缺失测试结果

Table 5 Data missing test results

场景编号	缺失情形	准确率/%
1	母线 2	98.31
2	母线 7	98.37
3	母线 2 和母线 7	98.16
4	母线 5 和母线 19	98.05
5	母线 14 和母线 16	98.11
6	母线 2、母线 14 和母线 19	97.42

4.7 泛化能力测试

泛化能力是指模型在不同的电力系统网络拓扑中,对于新样本的适应能力,即经过训练后,模型能对新数据给出符合预期的评估结果^[22]。考虑到由于故障或检修等情况导致电力系统拓扑结构发生改变,为验证 gcForest-AdaBoost 模型在未知拓扑工况下的适应能力,本文设置 6 种新的拓扑场景展开分析,场景描述如表 6 所示。

表 6 拓扑结构变化

Table 6 Topology changes

场景编号	拓扑结构
1	线路 4~15 断开
2	2 号发电机退出运行
3	线路 12~17 断开;线路 19~21 断开
4	5 号发电机退出运行;线路 23 和 24 断开
5	线路 3 和 4 断开;线路 16 和 17 断开;线路 27~29 断开
6	8 号发电机退出运行;线路 5 和 6 断开;线路 25~31 断开

依据 4.1 节所述方法,每种拓扑结构生成 500 个新的测试样本,并输入评估模型中。表 7 为不同拓扑结构下的测试结果。分析可知,在电力系统拓扑结构发生不同程度变化后,gcForest-AdaBoost 模型的评估准确率虽受到一定影响,但

精度仍然处于较高水平。在场景 1 和 2 中,由于拓扑结构改变较小,gcForest-AdaBoost 模型的评估准确率仍可达到 98% 以上,即使在拓扑结构改变较大的场景 5 和 6 中,gcForest-AdaBoost 模型的评估准确率仍达到 96% 以上。由此可见,本文所提 gcForest-AdaBoost 模型在电力系统拓扑结构发生变化时具有良好的泛化能力。

表 7 不同拓扑结构下的测试结果

Table 7 Test results under different topologies

场景编号	准确率/%
1	98.43
2	98.32
3	97.58
4	97.62
5	96.84
6	96.49

4.8 多任务 TSA 模型性能分析

前述研究主要针对暂态过程后的功角稳定性进行分析,此部分进一步针对系统暂态功角稳定性(任务 1)和暂态电压稳定性(任务 2)建立多任务评估模型。考虑到暂态电压稳定性的模型输出并非为单一数值而是整数集合,采用杰卡德系数 J 作为评估指标。表 8 为测试结果。可看出,所提模型杰卡德系数为 96.12%,远高于其他模型,说明在多任务暂态稳定评估中所提模型表现出优越性能。

表 8 多任务暂态稳定评估

Table 8 Multitask transient stability evaluation

模型	任务 1	任务 2
	准确率/%	杰卡德系数/%
gcForest-AdaBoost	98.56	96.12
LightGBM	98.23	94.38
DT	96.51	93.11
CNN	97.46	94.25
ANN	96.84	93.74
DBN	97.13	93.76
SVM	96.08	92.57

4.9 数据规模对模型训练时间的影响

为进一步分析样本数据规模对 gcForest-AdaBoost 模型训练时间的影响,将原有样本随机选择 20% 作为测试集,剩余样本分别按 80%、60%、40%、20% 设置训练集。图 6 为数据规模对模型训练时间的影响。由图可知,随着训练集数据规

模的不断增加, CNN 和 LightGBM 的耗时增加幅度远高于本文所提模型,说明在大规模数据中,集成深度随机森林算法具有较强速度优势,未来可更好地适用于大规模电网的实时评估需求。

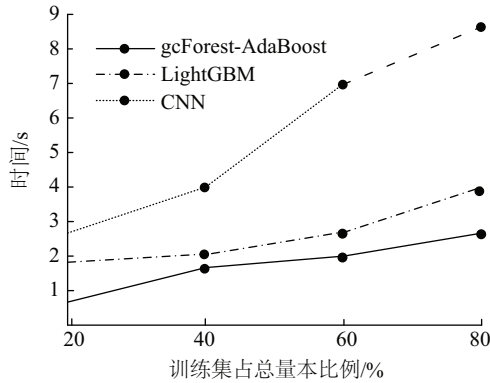


图6 数据规模对模型训练时间的影响

Fig. 6 Effect of data size on model training time

5 省级电网算例分析

本节通过中国某区域省级互联网 500 kV 主网部分展开实际系统算力分析,以验证所提方法在实际系统中的适用性。该系统共包含 168 条母线,42 台发电机和 186 条支路。在传输线路中设置故障类型为三相短路故障,故障发生时刻是 1 s,仿真时常设置为 5 s。故障位置分别设置在距离输电线首端 1%、25%、50%、75% 以及 99% 处。基于上述设置内容,共计生成 7872 个有效样本,其中稳定样本为 5746 个。进一步地,对有效样本进行标记,随机选取其中的 4500 个样本作为训练集,剩余样本作为测试集。基于表 1 中输入特征评估模型性能。表 9 为实际电网中 gcForest-AdaBoost 模型与其他模型效果对比。可看出,gcForest-AdaBoost 模型在实际电网中仍具有良好性能,其评估准确率为 98.12%,高于其他

表9 实际电网中 gcForest-AdaBoost 模型与其他模型效果对比

Table 9 Comparison of effect between gcForest-AdaBoost model and other models in actual power grid

模型名称	准确率/%	召回率/%	差准率/%	F_1 分数
gcForest-AdaBoost	98.12	97.45	97.29	98.17
LightGBM	97.75	97.11	96.58	97.49
DT	96.23	95.86	95.69	96.11
CNN	97.39	96.45	96.28	97.09
ANN	96.56	96.12	95.78	96.21
DBN	96.84	95.75	95.91	96.46
SVM	95.85	95.47	95.13	95.56

模型,能满足实际工程需要。综上,在大规模实际系统中,所提模型依然保持良好的评估性能,不会因电网规模的扩大导致性能明显降低。在实时预警和故障预警中,光伏储能并网系统可引入 gcForest-AdaBoost 模型,凭借其多源数据融合、暂态特征提取和动态权重优化的能力,与现有的电力系统调度和控制策略相结合,在省级电网中实现工程化应用。

6 结论

为提高数据驱动的电力系统暂态稳定评估的准确性,本文提出一种基于 gcForest-AdaBoost 模型的光伏储能并网系统暂态稳定评估方法,并在 IEEE39 节点系统展开算例分析,得到以下结论:

1) 将光伏储能系统接入 IEEE39 节点系统后,电压幅值发生明显变化,说明光伏储能系统对电力系统暂态稳定性具有一定影响。

2) 所提 gcForest-AdaBoost 模型在光伏储能并网系统暂态稳定评估中展现出优越性能。通过与 LightGBM、DT、CNN、ANN、DBN、SVM 等集成学习模型和个体学习模型对比评估结果,发现 gcForest-AdaBoost 模型在各方面测试中均表现优异,评估精度明显高于其他模型。

3) gcForest-AdaBoost 模型延续了集成学习模型的良好性能,在可能遇到的噪声干扰以及数据缺失测试中均展现出较强鲁棒性。同时,在面对未知拓扑结构改变时,gcForest-AdaBoost 模型也展现较强泛化能力,说明此模型具有良好评估性能。

未来,可进一步探索改进 gcForest-AdaBoost 模型,并深入挖掘光伏储能指标与电压暂态稳定裕度的关联性,以加强模型的可解释性与评估准确率的验证研究。同时,未来还可将模型扩展至电-氢联合储能系统,在多能源电网中测试模型的通用性,并根据发现的问题对模型进行改进与更新。

[参考文献]

- [1] 文继锋, 刘子俊, 周专, 等. 新型电力系统下高渗透新能源接入的次同步振荡问题研究[J]. 太阳能学报, 2024, 45(11): 50-60.
WEN J F, LIU Z J, ZHOU Z, et al. Research on sub-synchronous oscillation problem with high permeability new energy access in new power system[J]. Acta energiae solaris sinica, 2024, 45(11): 50-60.
- [2] 李汐, 高阳, 马斌, 等. 基于 PEF-PinSVM 的含可再生能源并网电力系统电压暂态稳定裕度评估[J]. 太阳能学报, 2024, 45(8): 240-248.
LI X, GAO Y, MA B, et al. Voltage transient stability margin assessment of grid connected power systems containing renewable energy based on PEF-PinSVM [J].

- Acta energiae solaris sinica, 2024, 45(8): 240-248.
- [3] REN C, XU Y. A fully data-driven method based on generative adversarial networks for power system dynamic security assessment with missing data [J]. IEEE transactions on power systems, 2019, 34(6): 5044-5052.
- [4] 田春胜, 任永峰, 胡志帅, 等. 基于自恢复型下垂控制的微电网运行控制策略研究[J]. 太阳能学报, 2024, 45(8): 71-77.
- TIAN C S, REN Y F, HU Z S, et al. Research on microgrid operation control strategy based on self-recovery droop control [J]. Acta energiae solaris sinica, 2024, 45(8): 71-77.
- [5] 林伟芳, 任晓钰, 张桂红, 等. 考虑功角稳定和暂态过电压的新能源电压穿越控制参数优化[J]. 电网技术, 2023, 47(4): 1323-1330.
- LIN W F, REN X Y, ZHANG G H, et al. Optimization of voltage ride-through control parameters of renewable energy considering power angle stability and transient overvoltage [J]. Power system technology, 2023, 47(4): 1323-1330.
- [6] 张玉冰, 申彦波, 姚鑫, 等. 青海高原光伏适宜性评价的不同决策树算法的比较研究[J]. 太阳能学报, 2024, 45(12): 30-39.
- ZHANG Y B, SHEN Y B, YAO X, et al. Comparative study of different decision tree algorithms for PV suitability evaluation in Qinghai plateau [J]. Acta energiae solaris sinica, 2024, 45(12): 30-39.
- [7] 朱乔木, 陈金富, 李弘毅, 等. 基于堆叠自动编码器的电力系统暂态稳定评估[J]. 中国电机工程学报, 2018, 38(10): 2937-2946.
- ZHU Q M, CHEN J F, LI H Y, et al. Transient stability assessment based on stacked autoencoder [J]. Proceedings of the CSEE, 2018, 38(10): 2937-2946.
- [8] MOVAHEDI A, NIASAR A H, GHAREHPETIAN G B. Designing SSSC, TCSC, and STATCOM controllers using AVURPSO, GSA, and GA for transient stability improvement of a multi-machine power system with PV and wind farms [J]. International journal of electrical power & energy systems, 2019, 106: 455-466.
- [9] 谈赢杰, 向真, 杨昆, 等. 时序数据驱动的微电网暂态稳定运行评估[J]. 南方电网技术, 2023, 17(7): 125-134.
- TAN Y J, XIANG Z, YANG K, et al. Time series data-driven transient stability assessment for microgrid [J]. Southern power system technology, 2023, 17(7): 125-134.
- [10] 李志兵, 肖健梅, 王锡淮. 基于多粒度NRS和改进Bi-LSTM的电力系统暂态稳定评估[J]. 电气工程学报, 2023, 18(3): 232-241.
- LI Z B, XIAO J M, WANG X H. Transient stability assessment of power system based on multi-granularity neighborhood rough set and improved Bi-directional long-short-term memory network [J]. Journal of electrical engineering, 2023, 18(3): 232-241.
- [11] 李森, 雷鸣, 周挺, 等. 基于深度森林的电力系统暂态稳定评估方法[J]. 电测与仪表, 2021, 58(2): 53-58.
- LI M, LEI M, ZHOU T, et al. Transient stability assessment method for power system based on deep forest [J]. Electrical measurement & instrumentation, 2021, 58(2): 53-58.
- [12] 朱瑞金, 董亚丽, 唐波. 基于改进深度森林的暂态电压稳定快速评估[J]. 电网与清洁能源, 2022, 38(6): 24-30, 43.
- ZHU R J, DONG Y L, TANG B. Fast assessment of transient voltage stability based on improved deep forest [J]. Advances of power system & hydroelectric engineering, 2022, 38(6): 24-30, 43.
- [13] 陈康, 王泽, 郭永吉. 基于grcForest模型的风电并网系统暂态电压稳定评估[J]. 智慧电力, 2023, 51(1): 31-37.
- CHEN K, WANG Z, GUO Y J. Transient voltage stability assessment of wind power grid-connected system based on grcForest model [J]. Smart power, 2023, 51(1): 31-37.
- [14] 叶进, 卢泉, 王钰淞, 等. 基于级联随机森林的光伏故障诊断模型研究[J]. 太阳能学报, 2021, 42(3): 358-362.
- YE J, LU Q, WANG Y S, et al. Research on PV fault diagnosis model based on cascaded random forest [J]. Acta energiae solaris sinica, 2021, 42(3): 358-362.
- [15] ZHOU Z H, FENG J. Deep forest: towards an alternative to deep neural networks [C]//Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence. Melbourne, Australia, 2017: 3553-3559.
- [16] 庄保乾, 韩路, 李晓虎, 等. 基于集成深度随机森林算法的智能电厂设备健康评估方法[J]. 计算机测量与控制, 2024, 32(8): 322-328.
- ZHUANG B Q, HAN L, LI X H, et al. Health status assessment method of intelligent power plant equipment based on integrated deep random forest algorithm [J]. Computer measurement & control, 2024, 32(8): 322-328.
- [17] 史法顺, 吴俊勇, 吴昊衍, 等. 基于深度学习的电力系统暂态功角与暂态电压稳定裕度一体化评估[J]. 电网技术, 2023, 47(2): 731-740.
- SHI F S, WU J Y, WU H Y, et al. Integrated evaluation of power system transient power angle and transient voltage stability margin based on deep learning [J]. Power system

- technology, 2023, 47(2): 731-740.
- [18] 李鹏, 董鑫剑, 孟庆伟, 等. 基于Fisher Score特征选择的电力系统暂态稳定评估方法[J]. 电力自动化设备, 2023, 43(7): 117-123.
- LI P, DONG X J, MENG Q W, et al. Transient stability assessment method for power system based on Fisher Score feature selection [J]. Electric power automation equipment, 2023, 43(7): 117-123.
- [19] 刘聪, 刘颂凯, 刘礼煌, 等. 基于TCN-CCRELMS的电力系统暂态稳定评估[J]. 电力系统及其自动化学报, 2023, 35(7): 36-44, 82.
- LIU C, LIU S K, LIU L H, et al. Transient stability assessment of power system based on TCN-CCRELMS[J]. Proceedings of the CSU-EPSA, 2023, 35(7): 36-44, 82.
- [20] 谢瀚阳, 彭泽武, 唐重阳, 等. 基于数据挖掘技术的电网时序数据质量维护研究[J]. 电测与仪表, 2022, 59(2): 38-44.
- XIE H Y, PENG Z W, TANG C Y, et al. Research on power grid time-sequence data quality maintenance based on data mining technology [J]. Electrical measurement & instrumentation, 2022, 59(2): 38-44.
- [21] 刘颂凯, 胡竞哲, 杨超, 等. 基于改进CDBN的电力系统暂态稳定评估[J]. 智慧电力, 2023, 51(6): 8-14, 92.
- LIU S K, HU J Z, YANG C, et al. Transient stability assessment of power systems based on improved CDBN[J]. Smart power, 2023, 51(6): 8-14, 92.
- [22] 张玉敏, 孙鹏凯, 叶平峰, 等. 基于CNN的配电网快速重构方法[J]. 智慧电力, 2022, 50(11): 112-118.
- ZHANG Y M, SUN P K, YE P F, et al. Fast reconfiguration method of distribution network of based on convolutional neural network [J]. Smart power, 2022, 50(11): 112-118.

TRANSIENT STABILITY EVALUATION OF GRID-CONNECTED PHOTOVOLTAIC ENERGY STORAGE SYSTEM BASED ON gcForest-AdaBoost

Wang Yuchen¹, Li Zhongyan¹, Su Han²

(1. College of Mathematics and Physics, North China Electric Power University, Beijing 102206, China;

2. School of Control and Computer Engineering, North China Electric Power University, Beijing 102206, China)

Abstract: It is of great significance to realize transient stability evaluation accurately and quickly for stable and safe operation of power system. Aiming at the problem of low prediction accuracy in the current transient stability evaluation of grid-connected photovoltaic energy storage systems, a transient stability evaluation method based on integrated deep Random forest algorithm (gcForest-AdaBoost) is proposed by combining deep learning and ensemble learning techniques. Firstly, the initial input features are constructed according to the steady-state components of the grid-connected photovoltaic energy storage system during operation. Secondly, in order to reduce the overfit of the model and ensure that the model still has strong learning ability after the number of cascades increases, a transient stability evaluation method based on gcForest-AdaBoost is proposed by combining deep random forest algorithm and AdaBoost algorithm. Thirdly, the gcForest-AdaBoost model is trained with input feature set, and the transient stability evaluation model of grid-connected PV energy storage system is established. Finally, the IEEE 39-node system is connected to the photovoltaic energy storage unit to build a case system for simulation analysis and data acquisition, and the evaluation results are obtained. Numerical examples show that the proposed model can effectively analyze the transient stability of grid-connected photovoltaic energy storage systems, and it is found that the proposed model has good robustness and generalization ability.

Keywords: solar power generation; energy storage; transient stability; random forests; gcForest-AdaBoost; input characteristics