

电力机器学习分布外泛化：问题建模与分析框架

尚宇炜, 王思懿*, 郭剑波, 盛万兴, 梁英

(中国电力科学研究院有限公司, 北京市 海淀区 100192)

Out-of-distribution Generalization in Machine Learning for Power Systems: Problem Formulation and Analysis Framework

SHANG Yuwei, WANG Siyi*, GUO Jianbo, SHENG Wanxing, LIANG Ying

(China Electric Power Research Institute, Haidian District, Beijing 100192, China)

ABSTRACT: To ensure the generalization performance of machine learning (ML) models, the statistical learning theory typically assumes that data are independent and identically distributed (i.i.d.). However, due to the openness of the power systems and the dynamic evolution of actual data, data distribution shifts may be unavoidable, resulting in significantly performance degradation when well-trained models are deployed in practical applications. This paper tackles the problems of ML in scenarios where the training and test dataset are not i.i.d., known as out-of-distribution generalization (OODG). First, the analysis framework of the OODG problem is established for power systems applications. Then, through the combination of theoretical analysis and case studies, the quantitative methods for assessing data distribution shifts, strategies for improving and evaluating the MLs' OODG performance, are explored. By deepening the understanding and research of OODG, the viability and application values of ML can be further enhanced.

KEY WORDS: machine learning; power system; generalization

摘要: 为保证机器学习(machine leaning, ML)模型的泛化性能, 统计学习理论通常假设数据独立同分布。然而, 由于电力系统的开放性和实际数据的动态演化性, 数据分布的偏移难以避免, 导致训练良好的模型在应用时效果往往大打折扣。为此, 该文重点研究当数据非独立同分布时的 ML 问题, 即分布外泛化问题(out-of-distribution generalization, OODG)。首先, 结合电力 ML 特点, 建立 OODG 的分析框架; 在此基础上, 通过理论分析和案例介绍相结合, 探究数据分布偏移的量化方法、OODG 能力的提升策略及其评估技术。深化 OODG 的认识和研究, 有利于提升 ML 学习方法的生命力和实用价值。

基金项目: 中国科学技术协会第八届“青年人才托举工程”(5342PD230004); 北京市科技创新新星支持项目(5342PD230003)。

Young Elite Scientists Sponsorship Program (5342PD230004); Beijing Nova Program (5342PD230003)。

关键词: 机器学习; 电力系统; 泛化能力

0 引言

随着自然语言、图像视频等领域的数据体量和算力规模快速增长及 Transformer 等先进神经网络模型性能的稳定成熟, 以大模型为标志的机器学习(machine leaning, ML)技术发展迅猛, 成为推动经济社会发展和科学技术进步的重要力量^[1]。在能源电力领域, ML 作为数字化智能化转型的引擎, 对系统规划设计、运行管控、运营服务的支撑价值凸显^[2-3]。国家电网、南方电网分别发布“光明”、“大瓦特”大模型, 面向电力生产、科研等应用场景提供智能服务。ABB 公司持续强化面向工控系统和设备的数字化服务平台 ABB Ability™研发, 西门子公司发布了“AI+预测性维护”等系列产品, 为行业赋能。

然而, ML 迄今取得的突出成就仍主要依赖于大规模、高质量的数据集。为保证模型泛化性能, 统计学习理论通常假设数据独立同分布(independent and identically distributed, i.i.d.), 即所有数据都独立地从同一个分布采样而得^[1]。例如, 研究表明, 大语言模型训练所用的文本数据正在穷尽互联网中可用的数据总量, 由于其对海量数据分布规律的充分学习, 使模型泛化能力显著增强^[4]。与互联网中的数据特征不同^[5], 电力系统许多应用场景难以汇集大规模高质量的数据, 且由于系统开放性和数据自身的时空演化性, 导致训练阶段和应用阶段的数据并非服从 i.i.d., 而是存在不同程度上的分布偏移^[6-7]。

为研究分布偏移对 ML 泛化性能的影响, 提出

了分布外泛化(out-of-distribution generalization, OODG)这一概念^[8], 一般定义为当测试集与训练集的数据为非独立同分布时 ML 模型的适用能力。研究表明, 工业控制领域数据分布受时间、空间、环境等因素变化而变化的趋势难以避免, 导致实验室训练良好的 ML 模型在部署应用后效果大打折扣, 形成“实验-应用鸿沟(lab to real gap)”^[9]。为此, 文献[6-7,10-12]结合具体问题, 分别基于动态图学习、时间序列分析、动态环境学习等方法, 从调整数据分布差距、预测分布变化和捕捉因果关系等角度开展 OODG 研究。电力系统属于一类随机开放的复杂工控系统, 不同场景的数据往往具有动态演化性、分布不均性、异构性和噪声干扰等特点, 同时对学习结果的可行性要求严格。针对数据分布偏移问题, 现有研究立足源荷功率预测、系统运行分析优化、设备故障诊断及运维等具体场景, 从回归分析、动态规划、模式识别等层面开展研究^[13-17], 并采用统计性能评估指标对模型泛化性能进行简化分析^[18-22], 而尚未系统性研究电力 ML 的 OODG 问题^[23-24]。实际数据的分布偏移对 ML 计算结果可行性和泛化能力边界产生的影响、有限数据训练下应当如何评估和提升 ML 的 OODG 能力, 是亟待研究的一些基础问题, 也是推动电力 ML 跨越“实验-应用鸿沟”的关键。特别是在基础设施智能化背景下, 上述研究对科学系统地建设大规模 AI 基础设施、经济高效地运营云/边/端侧 ML 模型具有重要理论和实际价值。

基于以上思考, 本文开展电力 ML 的 OODG 问题研究, 并提出 OODG 的统一分析框架。鉴于所研究问题的广泛性、复杂性, 本文重点研究数据分布偏移的量化分析(what), 分布偏移对 OODG 的影响(why), OODG 的能力提升及评估策略(how)。为增加可读性, 在论述理论方法的同时, 提供真实案例和数据分析, 以期激发更多有益思考。

1 分布外泛化问题及分析框架

本节在引出 OODG 问题的基础上, 提出适用于电力 ML 的 OODG 分析框架, 并建立相应的模型, 为后文奠定基础。

1.1 分布外泛化问题

令 X 表示 ML 的特征空间, Y 为标签空间, $f_{\theta}: x \rightarrow y$ 为从原始特征 x 映射到具有可学习参数 θ 的标签 y 的映射函数 f_{θ} , 损失函数 $l: y \times y \rightarrow \mathbb{R}$ 用于刻画预

测值与真实标签之间的差异, $P_{tr}(x, y)$ 为训练集的数据分布, $P_{te}(x, y)$ 为测试集和部署应用后的数据分布。受篇幅限制, 以监督学习为例对 OODG 进行研究, 但所述分析框架亦可拓展至无监督学习、强化学习等问题中^[25]。

令 $\{(x_1, y_1), \dots, (x_n, y_n)\}$ 表示从 $P_{tr}(x, y)$ 分布中抽样的 n 个训练样本, ML 的目标是找到一个最优模型 f_{θ}^* , 对 $P_{te}(x, y)$ 中的数据分布实现最佳泛化。

$$f_{\theta}^* = \arg \min_{\theta} \mathbb{E}_{x, y \sim P_{tr}} [l(f_{\theta}(x), y)] \quad (1)$$

传统 ML 方法通常根植于数据服从 i.i.d. 的假设, 即 $P_{tr}(x, y) = P_{te}(x, y)$, 并根据经验风险最小化(empirical risk minimization, ERM), 获得 ML 模型。

$$L_{ERM}(\theta) = \frac{1}{n} \sum_{i=1}^n l(f_{\theta}(x_i), y_i) \quad (2)$$

服从式(2)的模型泛化能力可视为分布内泛化(in-distribution generalization, IDG)^[26-27]。然而, 实际问题中, 时间、空间、环境等多因素耦合下, 数据分布往往呈现出渐变、突变等动态演化特征, 数据 i.i.d. 这一基本假设面临挑战, 即 $P_{tr}(x, y) \neq P_{te}(x, y)$ 。为应对这一挑战, OODG 问题被提出并在近年来获得高度重视。当前, 通常将 OODG 问题定义为 $P_{tr}(x, y) \neq P_{te}(x, y)$ 且 $P_{te}(x, y)$ 未知的学习问题。

图 1 给出 IDG 和 OODG 两种学习范式的对比, 在 IDG 学习范式下, 获得的最优学习模型 $f_{\theta-IDG}^*$ 在一个服从 i.i.d. 的局部数据空间中进行应用泛化, 属于一类内插泛化问题; 而 OODG 学习范式下, 获得的最优模型 $f_{\theta-OODG}^*$ 同时适用于 i.i.d. 的局部数据空间及在此空间以外的学习空间的应用泛化, 属于一类外推泛化问题。

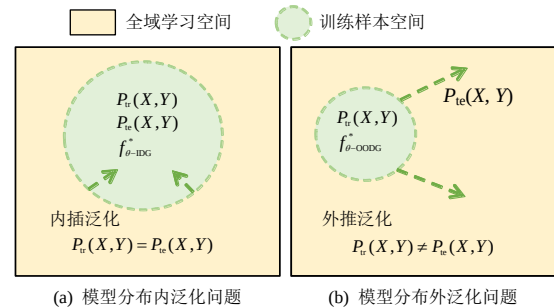


图 1 分布内泛化与分布外泛化对比示意

Fig. 1 Comparison between IDG and OODG

1.2 统一分析框架

令 $P(X)$ 表示协变量分布, $P(Y|X)$ 表示条件分布。 $P_{tr}(X, Y)$ 和 $P_{te}(X, Y)$ 之间造成分布的偏移, 本质上是由于 $P(X)$ 或 $P(Y|X)$ 的变化, 或二者的同时变化。在

电力 ML 问题中, $P(X)$ 变化的原因包括训练数据不足而导致分布规律无法全面覆盖及环境变化、量测误差、数据损坏等; $P(Y|X)$ 变化的原因包括系统/对象的运行状态变化及其他内外部影响因素等。概率论中的全概率公式可将复杂事件的发生及其影响因素, 转化为不同情况下所发生简单事件的概率求和问题^[28]。本文基于全概率公式, 建立 $P(X)$ 和 $P(Y|X)$ 之间产生分布偏移的关联关系公式, 如式(3)、(4)所示。

$$\begin{aligned} Pr(D(X, Y)) = & \alpha Pr(D(X, Y)|D(X)) \cdot Pr(D(X)) + \\ & \beta Pr(D(X, Y)|D(Y|X)) \cdot Pr(D(Y|X)) \end{aligned} \quad (3)$$

式中: $D(X, Y)$ 为 $P(X, Y)$ 的偏移量; $D(X)$ 为 $P(X)$ 的偏移量; $D(Y|X)$ 为 $P(Y|X)$ 的偏移量; $Pr(\cdot)$ 为发生的概率; $Pr(D(X, Y)|D(X))$ 指已知 $X \sim D(X)$ 前提下, 发生 $D(X, Y)$ 的概率; α 、 β 为影响数据分布偏移概率的权重系数。式(3)指已知先验知识情况下的数据分布偏移; 反之, 若已知后验知识, 可由式(4)刻画数据分布偏移的关联关系。

$$\begin{aligned} Pr(D(X, Y)) = & \alpha Pr(D(X, Y)|D(Y)) \cdot Pr(D(Y)) + \\ & \beta Pr(D(X, Y)|D(X|Y)) \cdot Pr(D(X|Y)) \end{aligned} \quad (4)$$

说明 1 以式(3)为例, 不同问题中 $Pr(D(X))$ 、 $Pr(D(Y|X))$ 对 $Pr(D(X, Y))$ 的相对影响程度不同^[29]。因此, 通过引入权重系数 α 、 β 来刻画影响程度的偏差。例如, 当 $Pr(D(X))$ 为主导因素时, $\alpha \gg \beta$; 反之, 当 $Pr(D(Y|X))$ 为主导因素时, $\alpha \ll \beta$ 。

说明 2 不同问题中 $P(X)$ 与 $P(Y|X)$ 及其分布偏移的特性不同, 导致采取的分布偏移刻画方法有所区别^[29]。例如, 一些问题中, $P(X)$ 与 $P(Y|X)$ 可直接根据输入数据特征进行捕捉, 也可以通过先验/后验知识对输入数据的分布进行刻画; 在另一些问题中, 分布偏移通常难以直接通过训练样本进行特征捕捉, 但可能通过先验/后验知识进行刻画。因此, 如何利用好先验/后验知识来分析 $P(X)$ 和 $P(Y|X)$ 并进而提升 OODG 能力, 十分关键。

针对电力实际应用, 为刻画不同问题对 ML 的泛化能力要求, 文献[20]提出“Begold”学习准则并建立相应模型, “Begold”学习准则要求在 ML 全局学习空间中实现概率趋优, 同时要在关键局部学习空间中满足底线约束。本文借鉴该模型来表征 OODG 能力要求, 如式(5)、(6)所示。其中, 式(5)强调 ML 模型在全局学习空间的泛化性能具备概率趋优的性质。

$$\begin{aligned} Pr\{|R_E^r(f) - R_E^e(f)| \leq \sigma\} = 1 - \delta \\ \exists \text{small } \delta > 0, \forall x \in X \end{aligned} \quad (5)$$

式中: $R_E^r(f)$ 为给定学习模型的最小泛化风险; $R_E^e(f)$ 为实际数据分布偏移的泛化风险; σ 为泛化误差的阈值。式(5)要求 ML 模型在全局学习空间上的泛化误差依概率收敛于某泛化误差边界。

式(6)强调 ML 模型在关键的局部学习空间中满足底线约束:

$$Pr\{\|y, f_\theta(x)\|_e \leq \sigma_i\} = 1 - \exists \sigma_i \geq 0, n < \infty, \forall x \cup X_i^B \subset X \quad (6)$$

式中: X_i^B 为第 i 个具有底线约束的局部空间; $\cup X_i^B$ 为所有具有底线约束的局部空间的集合; σ_i 为在 X_i^B 上的泛化误差的阈值; n 为训练样本数量。式(6)要求 ML 方法在局部学习空间上的泛化误差以概率 1 收敛于某泛化误差边界。

式(1)、(3)–(6)构成电力 OODG 问题的分析框架。从优化理论的视角看, 式(1)为 OODG 的学习优化目标; 式(3)、(4)可视为描述数据分布规律的等式约束; 式(5)、(6)可视为描述模型泛化能力的不等式约束。不同于传统 ML 基于 i.i.d. 假设来求解式(2), 本文建立的 ML 分析框架既考虑了数据分布的演化性, 也考虑了不同问题对 ML 泛化能力的边界约束, 从而使 ML 结果更具可行性和适应力。

1.3 OODG 问题的工程案例

本节以负荷预测为例进行分析。采用某配电网实际负荷数据集, 包含 35 040 个时间序列功率数据, 时间跨度 1 年、每天 96 个采样点、采集频次为 15 min。选取长短期记忆(long short-term memory, LSTM)网络对数据进行训练, 输入为预测日前一天的负荷数据, 记为 $x(t-95), \dots, x(t)$; 输出为预测日当天 96 个时间点的功率, 记为 $\hat{x}(t+1), \hat{x}(t+2), \dots, \hat{x}(t+96)$ 。为简化分析, 按时间顺序依次将数据集划分为训练集、第 1 测试集和第 2 测试集, 对应的数据分布分别记为 $P_{\text{tr-load}}(X, Y)$ 、 $P_{\text{te1-load}}(X, Y)$ 和 $P_{\text{te2-load}}(X, Y)$, 数据占比为 6:2:2。具体来说, 使用一年中前 180 天的 17 280 条数据作为训练集, 相邻后 60 天的 5 760 条数据作为第 1 测试集, 其后 65 天的数据不使用, 最后 60 天的 5 760 条数据作为第 2 测试集。图 2(a)为训练集中某天的负荷时间序列曲线, 图 2(b)采用概率密度函数(probability density function, PDF)来表征训练集的概率密度分布, 同理绘制图 2(c)–(f)。为便于说明, 本案例以 $P(X)$ 为例来剖析分布偏移与模型性能的关系。对比图 2(b)、(d)

可知, $P_{tr-load}(X)$ 与 $P_{te1-load}(X)$ 的分布一致性较高、分布偏移小;对比图 2(b)、(f)可知, $P_{tr-load}(X)$ 与 $P_{te2-load}(X)$ 的分布偏差较大。从两个测试集中分别选取一个典型日 96 个时间点的预测结果, 其预测精度对比结果如图 2(g)所示, 其中, 红色圆点表示 $P_{tr}(X, Y)=P_{te}(X, Y)$ 条件下的模型预测精度, 绿色圆点表示 $P_{tr}(X,$

$Y) \neq P_{te}(X, Y)$ 条件下的模型预测精度, 当圆点越靠近图中 45°斜线, 表示预测精度越高, 反之则表示精度越低。可以直观地看出, 当训练和测试数据近似满足 i.i.d.时, 预测误差较小; 反之, 当二者不满足 i.i.d.时, 预测误差明显增大。这一案例表明了数据分布偏移对 ML 模型的泛化性能的影响。

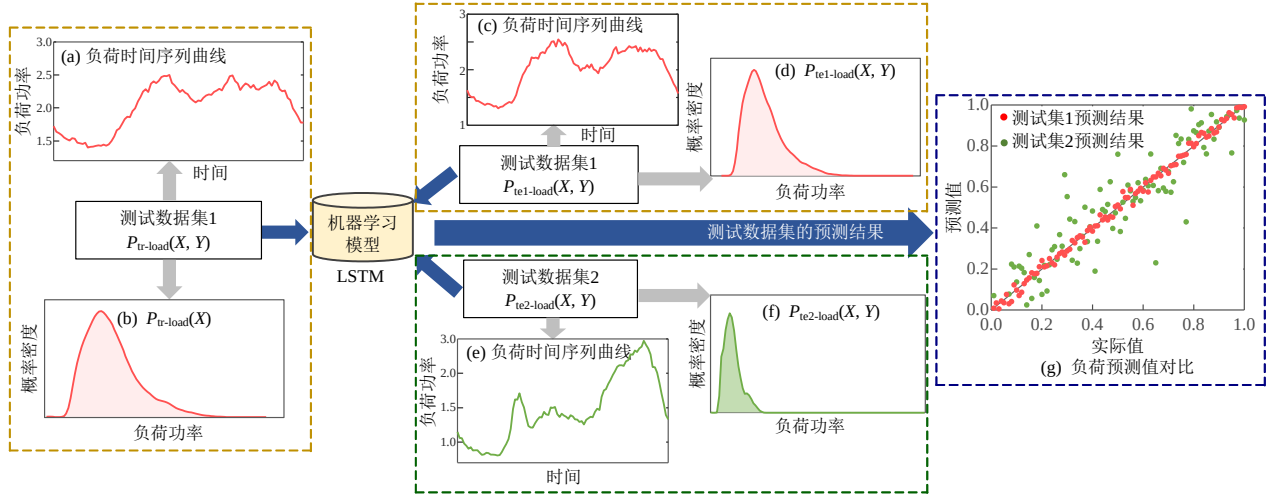


图 2 基于负荷预测案例分析数据分布偏移及模型性能影响

Fig. 2 Effects of data distribution shift and model performance are analyzed based on load forecasting case

2 分布偏移的量化分析

分布偏移是 OODG 问题的关键影响因素, 人工智能领域已开展了一些分布偏移量化方法研究。本节在分析典型方法的基础上, 给出适用于电力场景的分布偏移量化计算流程, 并分析不同方法的优劣。

2.1 分布偏移的量化方法

数据分布偏移作为 ML 与统计学的重要概念, 一般指训练数据和测试数据之间的概率分布发生变化的现象, 常发生在模型训练与测试/部署应用阶段。令 x_1, y_1 表示从训练数据分布中的抽样, x_2, y_2 表示从测试数据分布中的抽样, 图 3 给出分布偏

移量化分析的流程示意, 主要包括数据分布刻画、数据分布的偏移程度量化 2 个步骤。步骤①中, 根据输入特征数的不同, 以二维或高维 PDF 对 $P(X)$ 进行刻画; $P(Y|X)$ 可通过输入-输出关系图来刻画。步骤②可选取附式(A1)~(A4)来量化分布偏移的程度。在具体问题中, 可参考本节给出的计算流程, 选择适当的数学方法进行分布偏移量化^[30]。

具体地, 步骤①依据 PDF 及 ML 的输入-输出关系, 建立式(7)、(8)。式(7)给出 $P(X)$ 的 PDF:

$$Pr(X) = \frac{dF(X \leq x)}{dx} \quad (7)$$

式中 $F(X \leq x)$ 表示 $X \leq x$ 的概率。设概率密度在点 x 可导, PDF 为 $F(X \leq x)$ 在该点的导数。

式(8)以神经网络为例, 给出描述 $P(Y|X)$ 关系的一种直观方式:

$$Y = f(WX + b) \quad (8)$$

式中权重 W 、偏置 b 和激活函数 f 通过 (X, Y) 数据集训练获得, 据此可绘制输入-输出关系图。

步骤②中, 在量化分布偏移程度时, 多使用 f -散度、Wasserstein 距离、最大均值差异(maximum mean discrepancy, MMD)和拟合优度测试(如: R^2)等方法^[31-32]。其中, f -散度函数用于衡量两个概率密度 $P_{tr}(X)$ 和 $P_{te}(X)$ 间的差别^[33-35]; Wasserstein 距离

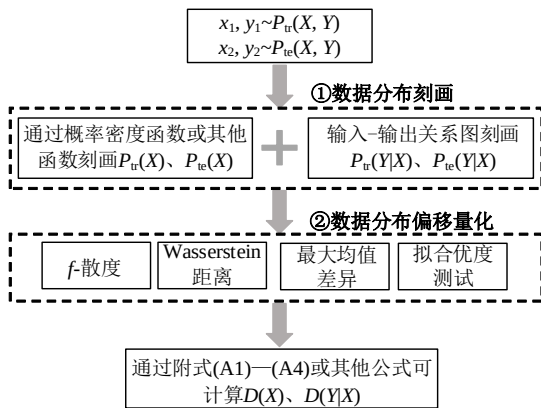


图 3 数据分布偏移量化流程示意

Fig. 3 Data distribution shift quantization flow diagram

用于衡量从 $P_{tr}(X)$ 移动至 $P_{te}(X)$ 时,所需要移动的平均距离的最小值^[36-37]; MMD 通过再生核希尔伯特空间中的均值嵌入,来比较 $P_{tr}(X)$ 与 $P_{te}(X)$ ^[38-40]; R^2 表示回归直线对观测值的拟合程度,当计算结果越接近 1,表示拟合效果越好^[41]。上述方法的异同点分析见表 1,具体计算式见附录 A。

表 1 数据分布偏移典型量化方法
Table 1 Typical quantization method of data distribution shift

方法	适用性
f -散度	通常适用于正态分布或近似正态分布样本; 不适用于数据分布无重叠情况
Wasserstein 距离	通常适用于非对称分布; 在分布无重叠时可反映两个分布的差距
最大均值差异 (MMD)	通常适用于比较两个分布的中心位置; 重点反映样本均值的差异;
拟合优度测试(R^2)	通常适用于分析不同分布下模型计算结果的变化

2.2 分布偏移量化案例分析

本节以分布式光伏功率预测为例进行分析。选取某区域 6 个分布式光伏场站的功率数据^[42],时间跨度为 1 年、采集频次为 5 min。光伏场站最大容量 1 058.4 kW、最小容量为 45.8 kW。为简化说明,本节仅选取辐照度和温度 2 个影响因素,并采用图神经网络进行训练^[43]。模型输入为各区域辐照度(G)和温度(T),记作 $X=[x_1, x_2, \cdots, x_i]$, $x_i=[G_i, T_i]$;输出为不同光伏场站功率,记作 $Y=[P_1, P_2, \cdots, P_i]$, $i=6$ 。图 4 给出 X 及 $Y|X$ 偏移数据分布的效果示意。

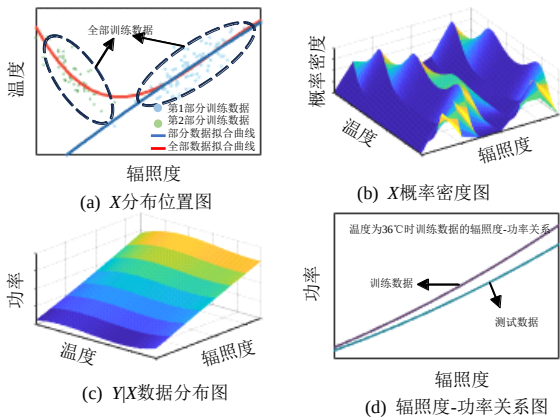


图 4 X 及 $Y|X$ 偏移数据分布示意
Fig. 4 Data distribution shift of X and $Y|X$

说明 3 图 4(a)、(b)用于对 $P(X)$ 进行刻画。图 4(a)中,蓝色点和绿色点代表全部训练数据,ML 模型的实际拟合曲线为红色线。若仅使用部分训练数据(如蓝色点集)对模型进行训练,则拟合得到蓝色直线,与实际的红色曲线差别大。因此,当训练

数据不足或分布偏移较大时,无法直接训练出有效模型。同时,图 4(a)仅能观察输入数据在图中分布位置的差异,无法直观地看出输入数据分布密度的差异。进一步,图 4(b)采用 PDF 计算输入数据的分布,图中浅色区域说明该区域的数据多、概率密度大;反之,深色区域的概率密度小。通过对比图 4(a)、(b),通过 PDF 来描述 $P(X)$,可以直观地显示数据缺失、分布情况等。

说明 4 图 4(c)、(d)用于对 $P(Y|X)$ 进行分析。图 4(c)选取单场站的辐照度、温度和光伏功率,绘制 $Y|X$ 数据分布的三维图,其中,辐照度与光伏功率间具有显著正相关性,温度与光伏功率具有相关性但图示规律不明显。图 4(d)为图 4(c)的二维断面图,表示当温度固定为 36℃时辐照度与功率的关系,随辐照度的增大,在同一温度下,ML 测试数据与训练数据的功率值差距逐渐增加,说明 2 个数据集的 $Y|X$ 相关性不同,即 $P_{tr}(Y|X)$ 与 $P_{te}(Y|X)$ 发生偏移。

图 5 为光伏预测案例中数据分布偏移的量化结果示意图。

说明 5 图 5(a)给出区域内 2 个场站的训练数据 $P_{tr-PV1}(X)$ 、 $P_{tr-PV2}(X)$ 与其测试数据 $P_{te-PV1}(X)$ 、 $P_{te-PV2}(X)$ 的概率密度图,4 个分布曲面分别用绿、蓝、紫和黄 4 种颜色表示。可以看出,4 个分布曲面差异较大,且数据分布不均匀,例如当温度为 40℃、辐照度为 800 W/m² 时,训练数据的密度明显较测试数据少,表明训练样本与测试样本的输入数据分布不同, $P(X)$ 分布偏移 $D(X)$ 是影响 $P(X,Y)$ 变化的因素之一。

说明 6 不同场站间的相关性也是影响 ML 预测性能的关键因素。图 5(b)给出 4 个时间断面下各场站间的相关性,以 I_{ij}^t 表示在 t 时刻场站 i 与 j 的相关性,每个圆圈代表一个光伏场站的天气数据,两圆圈间的连线粗细表示场站间天气相关性的强弱程度。在计算中,不同场站间天气数据的相关性作为图神经网络中邻接矩阵的输入。可以看出,随时间变化,不同场站间的相关性强弱动态变化。由此可得出,场站间的数据相关性存在分布偏移,这是影响预测性能的另一原因。

说明 7 图 5(c)和(d)分别给出区域内 2 个光伏场站的 $P(Y|X)$ 分布。可以看出, $P_{tr-PV1}(Y|X)$ 与 $P_{te-PV1}(Y|X)$ 间, $P_{tr-PV2}(Y|X)$ 与 $P_{te-PV2}(Y|X)$ 间并不完全重合,且在某些输入下偏差较大。这表明 X 、 Y 之

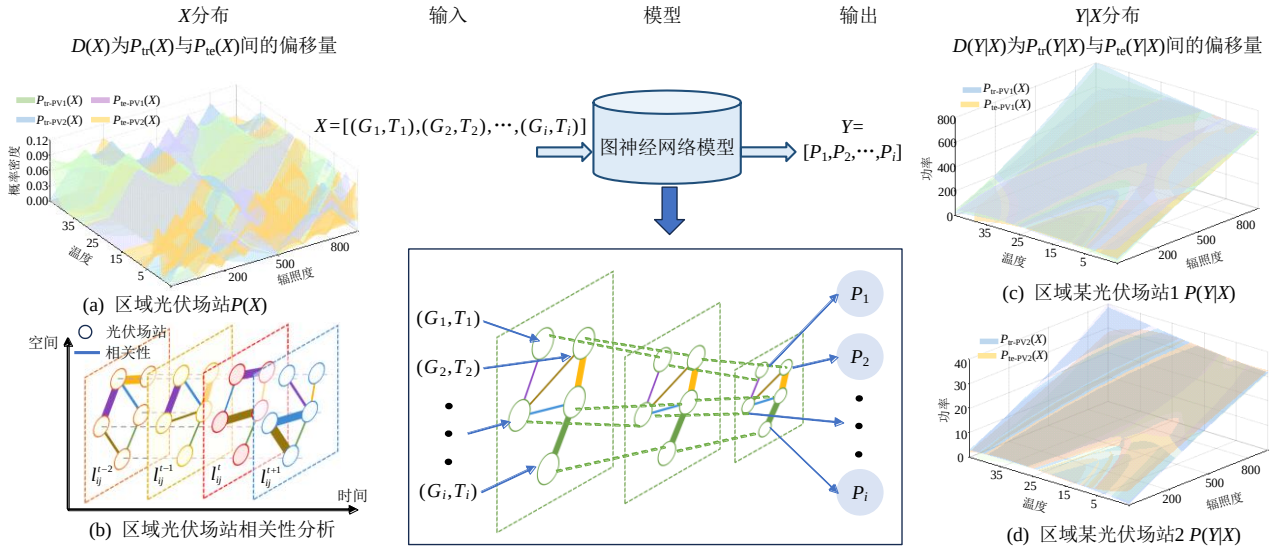


图 5 区域光伏预测案例中数据分布偏移量化示意图

Fig. 5 Data distribution shift quantification in a regional photovoltaic forecasting case

间的映射关系发生变化，导致 $D_{PV1}(Y|X)$ 和 $D_{PV2}(Y|X)$ ，因此， $P(Y|X)$ 分布的偏移是影响 $P(X, Y)$ 的一个主要原因。

说明 8 表 2 使用 4 种典型量化分布偏移的方法来计算图 5(a)、(c)、(d) 中的数据分布偏移程度，即计算 $D_{PV}(X)$ 与 $D_{PV}(Y|X)$ 。以使用 MMD 计算 $D_{PV1}(X)$ 的过程为例，选取高斯核作为核函数，分别计算 $P_{tr-PV1}(X)$ 、 $P_{te-PV1}(X)$ 所对应样本的格拉姆矩阵及 $P_{tr-PV1}(X)$ 、 $P_{te-PV1}(X)$ 所对应样本的交叉格拉姆矩阵，从而求得 MMD^[38-40]。其他分布量化方法的计算过程可参考文献[33-37, 41]。由表 2 可以看出不同方法在本案例中的效果差异。其中， f -散度计算的 2 组数据分布偏移差距较小，这是由于本问题中数据分布重叠较少。 R^2 计算结果存在负值。Wasserstein 距离、MMD 的计算结果数值不同，但反映数据分布偏移程度的趋势一致，在本问题中均能够有效量化分布偏移程度。

表 2 不同分布偏移量化方法的计算结果对比

Table 2 Computation results comparison of different distribution shifts quantization methods

	f -散度	Wasserstein 距离	MMD	R^2
$D_{PV1}(X)$	748.66	0.429 7	0.285 5	0.996 9
$D_{PV2}(X)$	748.66	0.590 5	0.285 7	0.991 5
$D_{PV1}(Y X)$	1.26	0.005 7	0.047 9	-0.388 7
$D_{PV2}(Y X)$	1.26	0.002 9	0.044 9	-0.566 8

3 提高学习模型分布外泛化能力的策略

根据“没有免费午餐定理(No Free Lunch Theorem)”^[44]，当 ML 模型缺少针对学习任务的结

构性假设和先验/后验知识时，难以通过有限数据的学习直接提升 OODG 能力。为此，本节在对比分析 OODG 与其他相关方法的基础上，从 ML 训练优化的全过程，建立提升模型 OODG 能力的策略集。

3.1 OODG 能力提升策略

根据问题特征、样本分布的不同，OODG 与多任务学习、迁移学习、零样本学习等研究领域具有相关性。例如，文献[45]指出，多任务学习的目标是使用单个模型同时学习多个相关任务且表现良好；迁移学习的目标是将旧模型应用到新领域中以解决新问题；域泛化的目标是将模型推广到因协变量变化而导致的不可见的分布；域自适应可通过获得目标数据得到数据分布情况。面向 OODG 的 ML 算法研究，可借鉴改进已有的算法。而从行为学视角来看，1 节建立的 OODG 分析框架突破了传统的端到端学习范式，从目标驱动和约束要求 2 个层面，促使模型具备自我评估验证能力，为拓展模型训练和动态优化的机制提供可能性。达到这一效果的核心是如何实现先验/后验知识与数据驱动的学习过程自洽融合。为此，本节从数据处理、模型设计、训练调优等 ML 训练优化的全过程，提出提升模型 OODG 的方法策略集，如图 6 所示。

1) 数据处理阶段。输入特征的选择和数据质量是影响 OODG 能力的基础因素。目前，输入特征的选取通常依赖于经验或相关性公式；数据缺失、错误等问题处理则多采用插值法、概率填充法等。这些方法往往忽略特征选择偏差和数据分布演化的客观因素。因此，本阶段可通过引入先验和后验

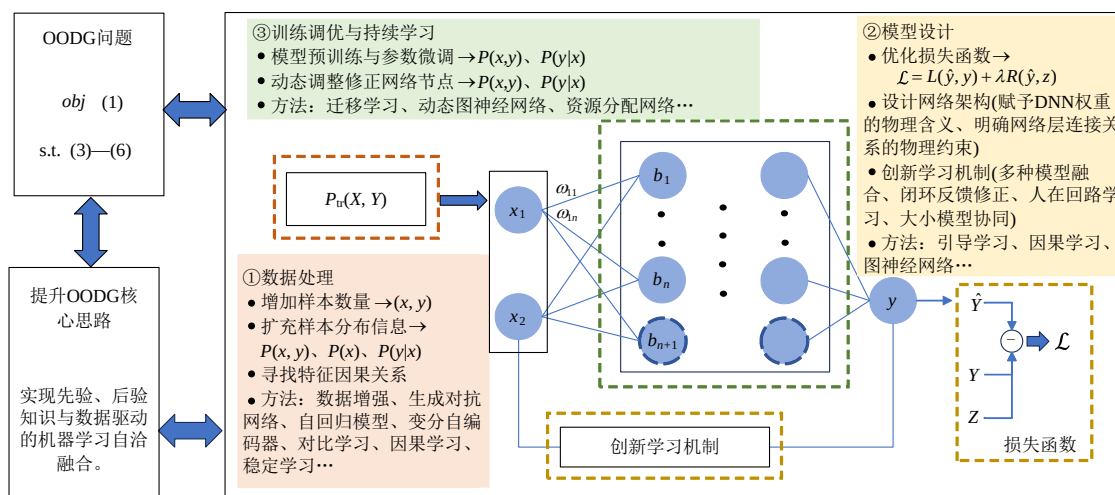


图6 电力机器学习的分布外泛化能力提升策略示意

Fig. 6 OODG performance improving strategies for electric machine learning models

知识, 优选特征集合并准确刻画数据分布特征。例如, 通过分析数据的关联/因果关系, 可避免特征混淆, 同时为捕捉数据分布规律奠定基础。具体可采取3类策略来提升OODG能力。

①增加样本数量。其核心是通过数据合成和样本增值, 方法主要包括: 合成少数类样本、生成对抗网络(generative adversarial network, GAN)和变分自编码器等。例如, 变分自编码器通过编码输入数据为潜在空间的概率分布, 并从该分布中采样以重建数据集, 实现数据扩充或 $P(X)$ 的捕捉。②扩充样本分布信息。其关键在于根据先验/后验知识, 建立适用于工程应用场景的物理约束, 以扩充符合客观实际的数据分布信息, 方法主要包括: 知识迁移学习、物理规律约束嵌入等。例如, 从已有的物理模型中提取知识, 来直接扩充数据或改变 $P(Y|X)$ 。③优选输入特征。其核心是寻找输入输出之间的相关性关系, 以避免混淆特征对模型性能的影响, 方法主要包括: 不变学习、稳定学习等。例如, 稳定学习提供了一种将因果推理与ML相结合的方法, 以寻找不变性为核心, 探索一种不随环境变化的稳定结构, 可视为精准捕捉 $P(Y|X)$, 通过样本重加权来提升模型泛化能力^[46-47]。

2) 模型设计阶段。传统基于端到端的学习过程简洁, 但对数据损失和偏差较敏感, 易导致计算结果不稳定。此外, 大多数模型采用“黑盒”架构, 未赋予神经网络结构以物理意义, 限制了OODG性能的提升。因此, 本阶段可通过引入先验、后验知识, 对模型进行设计优化, 重点采用3类策略实现。

①优化损失函数。其核心是通过修改损失函

数, 如添加正则项等^[48], 将先验/后验知识融入学习目标, 以降低数据分布偏移对模型的影响, 方法主要包括: 因果学习、引导学习、物理信息网络(physics-informed neural networks, PINN)等。例如, 引导学习通过将领域知识经验形式化建模并作为知识正则项加入损失函数, 从而优化模型训练、降低噪声干扰, 提高捕捉 $P(Y|X)$ 的质效^[49]。②设计网络架构。其关键在于增强黑盒模型与物理现象的关联性, 方法主要包括: 赋予神经网络中权重参数的物理含义, 设计层次化/模块化的神经网络架构并明确相关模块的物理约束等。例如, 使用PINN预测电网运行的安全域时, 可将第一隐藏层的物理意义赋予为故障边界数量、第二隐藏层赋予故障逻辑关系, 输出层为最终的稳定性评估结果^[14]。③创新学习机制。其核心是将领域知识与ML模型融合, 以自适应修正ML模型结果, 方法主要包括模型融合、反馈、人在回路等机制^[50]。

值得一提的是, 工程实际中具有生命力的ML模型并非总是要求模型庞大且复杂。如何设计简单高效的学习机制对于提升OODG十分重要^[18,51]。例如, 比例-积分-微分(proportional-integral-derivative, PID)控制是一种简单高效实用的反馈机制, 能够自适应动态调节相关参量, 达到期望的控制效果。借鉴PID反馈机制, 本文提出一种反馈学习机制, 如图7所示, 通过引入先验/后验知识驱动模型, 从①数据分布层面、②ML模型层面、③学习结果层面这3个层面来提升OODG。具体的, 输入 x 作用于数据驱动的ML模型和先验知识驱动模型, 分别得到 $\hat{y}_D$ 和 $\hat{y}_K$, 在前馈学习过程中, 先验知识驱

动的模型可直接作用于 ML 模型。 $\hat{y}_D$ 和 $\hat{y}_K$ 作用于判别模块，以数据分布特征/变化规律等作为判别条件，通过动态聚类等方法，来判断基于数据驱动的 ML 模型预测结果是否与先验/后验知识一致或自洽。若 ML 输出结果与先验/后验知识一致，则按照一定的输出策略，如融合数据驱动与知识驱动 2 类方法的结果，作为最后的输出 y 。反之，若 ML 输出结果与先验/后验知识不一致，则进入反馈环节，进而通过①、②、③3 个层面进行自适应修正。其中，①通过反馈到数据分布层面，变化数据分布条件，本质是调整 $P(X)$ ；②通过反馈到模型层面，自适应调整模型架构，本质是调整 $P(Y|X)$ ；③通过反馈到结果层面，根据经验进行结果重加权，本质是调整 $P(X, Y)$ 。通过自适应反馈学习过程优化模型，直至 $\hat{y}_D$ 和 $\hat{y}_K$ 判别一致。

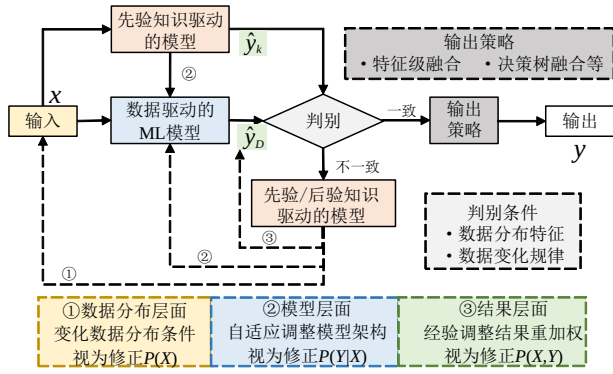


图 7 反馈机制结构

Fig. 7 Feedback mechanism structure

3) 训练调优与持续学习阶段。传统学习模型难以自适应地适应学习环境和数据分布的动态变化，通常需要根据数据特征进行针对性的微调甚至重新训练。因此，如何实现高效训练和持续地学习尤为重要，在适应动态环境变化的同时降低训练成本。为此，本阶段可通过先验/后验知识调整模型，使其更好地拟合 $P(Y|X)$ ，从而在动态演化的实际场景中具有泛化性和稳定性。

训练调优主要用于模型的训练阶段，其核心是通过高效训练和微调等技术来调整模型参数/结构，以适应数据分布的变化，方法主要有迁移学习、域自适应等。持续学习主要用于模型部署应用后，其关键在于根据数据特征来动态调整神经网络，以适应时间和环境造成的数据分布偏移，方法主要有资源分配网络(resource-allocating network, RAN)、动态图神经网络等。例如，基于 RAN 的实时调整技术可动态修改隐藏节点，通过捕捉 $x, y \sim P_{te}(X, Y)$ ，实现

在线预测模型的实时动态学习，可视为调整 $P(Y|X)$ 。为避免噪声干扰，还可在动态调整策略中建立缓冲区，以记录误差大的样本及其出现的频率，而后根据频率选取样本对模型进行动态调整^[52-53]。

此外，还可借鉴大模型的一些新技术新思路，如测试时适应(Test-time Adaptation)、测试时扩展(Test-time Scaling)等，推动 ML 模型从“静态训练”向“动态适应”转变。

3.2 OODG 提升策略的适用性分析

本节以 2.2 节光伏预测案例为例，从数据处理、模型设计和训练调优与持续学习 3 个阶段，对所述 OODG 能力提升策略进行实例化分析，以验证其可行性和适用性。

如图 8 所示，在数据处理阶段，通过物理融合的生成对抗网络(physics-informed GAN, PIGAN)，将辐照度、温度和发电功率间的物理关系式嵌入损失函数，通过 PIGAN 生成满足数据分布特征与客观规律的新样本，来降低对原始样本的依赖^[54]。在模型设计阶段，建立反馈学习机制，将客观物理知识与数据驱动的模型相结合，有利于避免数据损坏造成的影响。先验知识驱动模型的输出 $P_t^1=[P_1^1, P_2^1, \dots, P_t^1]$ ，数据驱动模型的输出 $P_t^2=[P_1^2, P_2^2, \dots, P_t^2]$ ，基于动态聚类算法对 2 种输出结果进行比较，当输出结果聚为同一类时，以数据驱动模型的结果作为最终输出；反之，若 2 种结果不一致，则调整数据分布重新计算，直至二者结果趋于一类。在训练调优与持续学习阶段，可采用 RAN 来适应环境变化造成的数据分布偏移，动态调整神经网络结构以实时适应数据分布情况。若输出 $P_t^1=[P_1^1, P_2^1, \dots,$

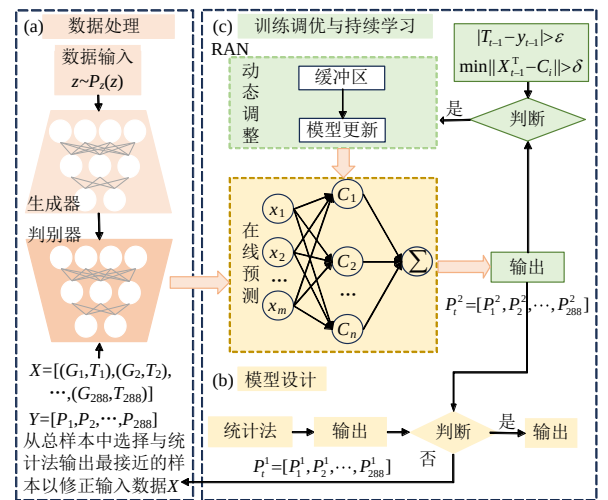


图 8 提升模型 OODG 能力的策略实例示意

Fig. 8 Example of improving ML's OODG performance

P_1^j 与实际光伏功率偏差大于某临界值 ε ,且新样本与历史样本的平均距离大于临界距离阈值 δ 时,将该样本存入缓冲区。当缓冲区数值达到某限值时,在网络中添加新的节点 C_n ,并归一化当前缓冲区数据来训练新隐藏层节点,从而适应数据分布的动态变化情况。通过上述3个阶段,ML模型可视为具有自推理自校核能力,从而更好地适应数据分布偏移。

4 分布外泛化能力评估

传统泛化性能评估方法多基于统计指标或经验进行简化分析。本节结合电力应用特点,针对性地选取 OODG 评估方法,并建立模型训练和部署应用两阶段联动的 OODG 持续评估策略。

4.1 OODG 能力评估

以往针对 ML 模型泛化性能的评估多采用均方误差(mean squared error, MSE)、均方根误差(normalized root mean squared error, NRMSE)、平均绝对误差(mean absolute error, MAE)、 R^2 等统计指标进行简化分析,难以科学反映 ML 模型的泛化性能。而 OODG 评估具有内在复杂性,学界一般从 OODG 性能测试、OODG 性能预测和 OODG 内在特性表征等方面开展研究^[26]。结合电力工程应用特点,本节从 ML 训练和应用两阶段着手,提出 OODG 评估策略,主要思路是将理论与经验分析相结合,如图 9 所示。在模型训练阶段,重点分析模型的鲁棒性、稳定性、不变性、平稳性和一致性“5性”^[55],具体计算过程见附录 A;在模型应用阶段,随着所获得实际数据不断丰富,可在分析“5性”基础上,分析数据分布偏移程度和模型置信度这“2度”,持续修正对 OODG 性能的评估。

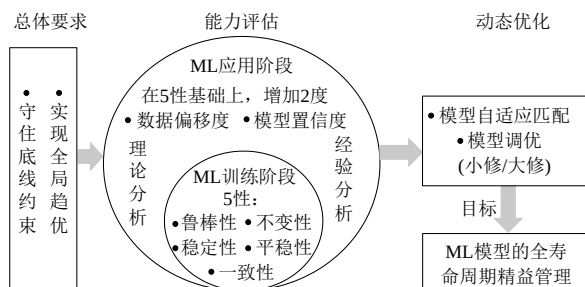


图9 ML训练和应用阶段的 OODG 能力评估策略

Fig. 9 OODG performance evaluation strategy in ML training and application stage

图9中,模型鲁棒性指模型在面对异常数据、噪声或不完整数据时,仍能准确预测或分类的能力。模型稳定性指模型能够持续地输出一致的、可

预测的结果。模型不变性指在通过设计新的学习算法来引导模型利用数据中的不变关系,使输入与输出间的关系在所有环境下都成立。模型平稳性指一个时间序列在统计性质上不随时间变化。模型一致性指一个模型在不同的数据集、不同的运行情况下能保持相似的预测性能。数据分布偏移程度详见2.1节式(7)、(8)及附式(A1)~(A4)^[56-57]。模型置信度指模型对预测结果的确信程度。通常,模型在具有高置信水平范围内的性能优越,通过平均阈值置信度可估计误差。

根据实际问题需求和数据特征,在 ML 的不同阶段,可针对性地选择不同的方法和指标。OODG 的评估结果,可进一步为优化调整 ML 模型性能提供参考。例如,若某类问题中,既有的 ML 模型库已十分丰富,可直接依据数据分布的相似度,自适应匹配相关的 ML 模型进行应用;反之,若既有 ML 模型较少时,可结合3节所述方法对 ML 模型进行“运维”,以持续提升其泛化性能,相应的运维策略包括“小修”(如参数微调)、“大修”(如神经网络结构调整),或模型退役/更换等。值得一提的是,针对 ML 模型的全寿命周期运维管理策略,对未来大规模部署应用 ML 的技术经济性十分重要。

4.2 OODG 能力评估策略适用性分析

本节通过配电网电压控制和光伏功率预测案例,分析所述 OODG 能力评估策略在动态规划、时序预测等经典学习问题中的适用性。

案例1: 配电网电压控制。

1) 问题背景:通过控制分布式电源和储能,以避免节点电压越限,并最大化配电网运行经济效益。采用强化学习中的软行动者-评论家(soft actor-critic, SAC)算法对问题进行求解,模型输入为配电网拓扑结构、功率等,输出为分布式电源无功及储能的出力大小,该问题数据及采用的 ML 模型的详细介绍可参考文献[13]。由于分布式电源的运行不确定性,通过评估 OODG 能力,有助于掌握 SAC 算法的实际泛化性能,确保电网运行安全可靠。

2) OODG 评估:本案例重点分析模型输出结果是否满足底线约束,以确保配电网运行安全性。一种可行的评估方法如下。首先,据式(6),采用4.1节所述的鲁棒性评估开展理论分析:基于策略梯度、最大熵目标函数等数学方法,重点论证 SAC 算法具有鲁棒性^[58]。其次,当理论分析通过后,进一步结合实际问题和数据开展经验分析:传统评估

仅考虑在特定数据下的决策效果；鲁棒性评估通过对分布式电源添加随机噪声，即模拟数据分布偏移的不同程度，噪声幅度越大表明训练-测试数据间的分布变化越显著。基于添加随机噪声，生成多个决策结果，并统计结果分布情况及上下限，参考式(A5)，分析 SAC 决策结果能否确保配电网节点电压标么值处于[0.95, 1.05]的可行范围内，以此获得该模型在数据分布随机变化下的 OODG 能力边界。如表 3 所示，对分布式电源出力数据随机添加 10%、20%高斯噪声，并对配电网 24h 的电压越限率进行分析发现，传统评估难以刻画数据分布偏移下算法的泛化性能；本文采用的鲁棒性评估指标表明，随着数据分布偏移程度增加时，算法的泛化能力呈现下降趋势，在更加立体地展示学习算法性能的同时，也体现了该算法在实际应用中的效果。

表 3 配电网调压模型的鲁棒性评估

Table 3 Robustness assessment of voltage regulation models for distribution networks

输入条件	电压越限率	
	传统评估	鲁棒性评估
无噪声	0%	0%
10%噪声	未考虑	2.3%±0.6%
20%噪声	未考虑	4.7%±1.1%

案例 2：光伏功率预测。

1) 问题背景：设 ML 模型的输入为辐照度、温度，输出为光伏功率，选取前馈神经网络进行光伏功率预测，问题详细描述见于 2.2 节案例^[59]。同时，建立 3.2 节提出的反馈机制来动态优化模型的 OODG 性能。

2) OODG 评估：本案例重点分析模型输出结果的精准性，即是否具有全局趋优的泛化特性。一种可行的评估方法如下。首先，据式(5)，采用 4.1 节所述的置信度评估开展理论分析：融合深度学习和统计推断的方法，其核心是通过神经网络的非线性能力捕捉复杂关系，利用非线性分位数回归/贝叶斯神经网络建模结果分布^[60]。其次，在此基础上，进一步开展经验分析：传统评估仅基于测试数据计算误差反映模型性能；置信度评估则通过设置不同百分比，动态计算不同置信区间，参考附式(A9)，并通过平均覆盖误差(average coverage error, ACE)和区间平均宽度(prediction interval normalized average width, PINAW)指标评估模型在应用阶段的性能，反映模型预测结果的有效性。如表 4 所示，基于前馈神经网络结合非线性分位数回归技术计

算 90%置信度，对 24h 的光伏预测值进行分析发现，传统评估方法难以刻画模型在实际运行中的泛化性能；本文采用的置信度评估指标表明，90%置信度的区间宽度和预测结果的覆盖误差能有效反映预测结果的可靠性，从而为电网调度提供决策依据。

表 4 光伏功率预测模型的置信度计算

Table 4 Calculation of the confidence level for photovoltaic power forecasting models

结果形式	传统评估	置信度评估	
	NRMSE	ACE	PINAW
精确预测	9.3%	—	—
90%置信度	—	0.107	0.216

5 结论

1) 由于电力系统运行的开放性和实际数据的动态演化性，数据分布偏移难以避免，从而对 ML 的泛化性能产生影响，造成“实验-应用鸿沟”。对 ML 分布外泛化能力的系统性认识和研究不足，或将成为电力 ML 大规模部署和应用的关键阻碍。

2) 从假设数据服从 i.i.d. 的 IDG 学习范式转变为 OODG 学习范式，是跨越“实验-应用鸿沟”的关键。本文从优化理论视角，建立了 OODG 分析框架，基于全概率模型分析数据分布偏移的等式约束，形成“全局概率趋优、局部守住底线”的 OODG 不等式约束。该框架同时适用于监督学习、无监督学习、强化学习等学习任务。

3) 分布偏移的量化可通过刻画数据分布、量化数据分布偏移程度 2 个主要步骤完成。其中，数据分布可通过 PDF 等方法刻画，偏移可通过 f -散度、Wasserstein 距离、MMD 和拟合优度测试等方法和公式量化。

4) 当 ML 模型缺少针对学习任务的结构性和先验、后验知识时，难以通过有限数据的学习直接提升 OODG。为此，要根据学习任务合理构建数据、结构等方面的假设条件，并结合实际问题校验假设的符合性，这是研究 OODG 的前提；其次，从样本优化、结构设计、动态调整等维度，将先验、后验知识自洽地融入数据驱动的学习过程，这是提升 OODG 的核心。

5) 随着 ML 技术的进步，泛化能力的定义也在不断发展，对于 OODG 的评估需要结合具体问题具体研究。一般地，在训练阶段可重点考虑模型的“5 性”；在应用阶段可增加评估模型的“2 度”。

OODG 的评估可为 ML 模型的可行性和能力边界提供分析依据,为 ML 模型的全寿命周期科学精益管理提供决策依据。

6) OODG 研究重点包括分布偏移的量化分析、OODG 能力提升策略、OODG 能力评估等方面。需结合具体问题结构,考虑复杂环境学习、实时数据处理等实际需求,将电气、统计学、运筹学等理论结合,强化基础理论创新;同时,顺应电力信息-物理系统融合发展趋势,强化在功率预测、系统调度、设备运维等方面的应用创新。

参考文献

- [1] LIU Jiashuo, SHEN Zheyang, HE Yue, et al. Towards out-of-distribution generalization: a survey [DB/OL]. (2023-07-27)[2025-01-28]. <https://doi.org/10.48550/arXiv.2108.13624>.
- [2] 国家能源局. 国家能源局关于加快推进能源数字化智能化发展的若干意见[EB/OL]. (2023-03-28) [2025-01-28]. https://zfxxgk.nea.gov.cn/2023-03/28/c_1310707122.htm. National Energy Administration. Several Opinions of the National Energy Administration on Accelerating the Digital and Intelligent Development of Energy [EB/OL]. (2023-03-28)[2025-05-01]. https://zfxxgk.nea.gov.cn/2023-03/28/c_1310707122.htm(in Chinese).
- [3] 王新迎, 闫冬, 施展, 等. 机器学习赋能的优化算法及其在新型电力系统中的应用与展望[J]. 中国电机工程学报, 2024, 44(16): 6367-6384.
WANG Xinying, YAN Dong, SHI Zhan, et al. Machine learning empowered optimization algorithms and their applications and prospects in new type power system [J]. Proceedings of the CSEE, 2024, 44(16): 6367-6384(in Chinese).
- [4] JONES N. The AI revolution is running out of data. What can researchers do?[J]. Nature, 2024, 636(8042): 290-292.
- [5] YIN Dong, LOPES R G, SHLENS J, et al. A Fourier perspective on model robustness in computer vision[C]// Proceedings of the 33rd International Conference on Neural Information Processing Systems. Vancouver: Curran Associates Inc., 2019: 13276-13286.
- [6] YUAN Haonan, SUN Qingyun, FU Xingcheng, et al. Environment-aware dynamic graph learning for out-of-distribution generalization [DB/OL]. (2023-11-18)[2025-01-28]. <https://doi.org/10.48550/arXiv.2311.11114>.
- [7] MAO Yixiu, WANG Qi, QU Yun, et al. Doubly mild generalization for offline reinforcement learning [DB/OL]. (2024-11-12)[2025-01-28]. <https://doi.org/10.48550/arXiv.2411.07934>.
- [8] CHEN Yongqiang, ZHANG Yonggang, BIAN Yatao, et al. Learning causally invariant representations for out-of-distribution generalization on graphs[DB/OL]. (2022-10-11)[2025-01-28]. <https://doi.org/10.48550/arXiv.2202.05441>.
- [9] HSU K C, REN A Z, NGUYEN D P, et al. Sim-to-lab-to-real: safe reinforcement learning with shielding and generalization guarantees[J]. Artificial Intelligence, 2023, 314: 103811.
- [10] WANG Zijian, LUO Yadan, QIU Ruihong, et al. Learning to diversify for single domain generalization[C]// Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision (ICCV). Montreal: IEEE, 2021: 814-823.
- [11] FAN Shaohua, WANG Xiao, SHI Chuan, et al. Generalizing graph neural networks on out-of-distribution graphs[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2024, 46(1): 322-337.
- [12] WANG Lu, WANG Jingdong, SUN Xinwei, et al. Out-of-distribution representation learning for time series classification[DB/OL]. (2022-09-15)[2025-01-28]. <https://doi.org/10.48550/arXiv.2209.07027>.
- [13] WANG Siyi, SHENG Wanxing, SHANG Yuwei, et al. Distribution network voltage control considering virtual power plants cooperative optimization with transactive energy[J]. Applied Energy, 2024, 371: 123680.
- [14] REN Junzhi, ZENG Yuan, QIN Chao. Practical dynamic security region model: a hybrid physical model-driven and data-driven approach[J]. IEEE Transactions on Power Systems, 2025, 40(1): 728-739.
- [15] LIU Wei, REN Chao, XU Yan. PV generation forecasting with missing input data: a super-resolution perception approach[J]. IEEE Transactions on Sustainable Energy, 2021, 12(2): 1493-1496.
- [16] ZHU Yumeng, ZI Yanyang, LI Jie, et al. PhysiCausalNet: a causal- and physics-driven domain generalization network for cross-machine fault diagnosis of unseen domain[J]. IEEE Transactions on Industrial Informatics, 2024, 20(6): 8488-8498.
- [17] KAKIMOTO M, ENDOH Y, SHIN H, et al. Probabilistic solar irradiance forecasting by conditioning joint probability method and its application to electric power trading[J]. IEEE Transactions on Sustainable Energy, 2019, 10(2): 983-993.
- [18] 尚宇炜, 郭剑波, 吴文传, 等. 电力脑初探: 一种多模

- 态自适应学习系统[J]. 中国电机工程学报, 2018, 38(11): 3133-3143.
- SHANG Yuwei, GUO Jianbo, WU Wenchuan, et al. Preliminary study of electric power brain: a multimodal adaptive learning system[J]. Proceedings of the CSEE, 2018, 28(11): 3133-3143(in Chinese).
- [19] 尚宇炜, 郭剑波, 吴文传, 等. 数据-知识融合的机器学习(1): 模型分析[J]. 中国电机工程学报, 2019, 39(15): 4406-4415.
- SHANG Yuwei, GUO Jianbo, WU Wenchuan, et al. Machine learning methods embedded with domain knowledge (part I): model analysis[J]. Proceedings of the CSEE, 2019, 39(15): 4406-4415(in Chinese).
- [20] 尚宇炜, 郭剑波, 吴文传, 等. 数据-知识融合的机器学习(2): 泛化风险[J]. 中国电机工程学报, 2019, 39(16): 4641-4649.
- SHANG Yuwei, GUO Jianbo, WU Wenchuan, et al. Machine learning methods embedded with domain knowledge(part II): generalization risk[J]. Proceedings of the CSEE, 2019, 39(16): 4641-4649(in Chinese).
- [21] 尚宇炜, 马钊, 彭晨阳, 等. 内嵌专业知识和经验的机器学习方法探索(二): 引导学习的应用与实践[J]. 中国电机工程学报, 2017, 37(20): 5852-5861.
- SHANG Yuwei, MA Zhao, PENG Chenyang, et al. Study of a novel machine learning method embedding expertise (part II): applications and practices of guiding learning[J]. Proceedings of the CSEE, 2017, 37(20): 5852-5861(in Chinese).
- [22] SHANG Yuwei, WU Wenchuan, HUAI Xu, et al. Loss of life estimation of distribution transformers considering corrupted AMI data recovery and field verification [J]. IEEE Transactions on Power Delivery, 2021, 36(1): 180-190.
- [23] GUILLORY D, SHANKAR V, EBRAHIMI S, et al. Predicting with confidence on unseen distributions[C]// Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision(ICCV). Montreal: IEEE, 2021: 1114-1124.
- [24] ZHOU Kaiyang, LIU Ziwei, QIAO Yu, et al. Domain generalization: a survey[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(4): 4396-4415.
- [25] ADA S E, OZTOP E, UGUR E. Diffusion policies for out-of-distribution generalization in offline reinforcement learning[J]. IEEE Robotics and Automation Letters, 2024, 9(4): 3116-3123.
- [26] YU Han, LIU Jiashuo, ZHANG Xingxuan, et al. A survey on evaluation of out-of-distribution generalization [DB/OL]. (2024-03-04)[2025-01-28]. <https://doi.org/10.48550/arXiv.2403.01874>.
- [27] KAWAGUCHI K, KAELBLING L P, BENGIO Y. Generalization in deep learning [DB/OL]. (2017-07-06)[2025-01-28]. <https://doi.org/10.48550/arXiv.1706.08947>.
- [28] 李航. 机器学习方法[M]. 北京: 清华大学出版社, 2022.
- LI Hang. Machine learning methods[M]. Beijing: Tsinghua University Press, 2022(in Chinese).
- [29] LU Wang, WANG Jindong, SUN Xinwei, et al. Out-of-distribution representation learning for time series classification[DB/OL]. (2023-02-08)[2025-01-28]. <https://doi.org/10.48550/arXiv.2209.07027>.
- [30] LU Jie, LIU Anjing, DONG Fan, et al. Learning under concept drift: a review[J]. IEEE Transactions on Knowledge and Data Engineering, 2019, 31(12): 2346-2363.
- [31] CREAGER E, JACOBSEN J H, ZEMEL R. Environment inference for invariant learning[DB/OL]. (2021-07-15)[2025-01-28]. <https://doi.org/10.48550/arXiv.2010.07249>.
- [32] CASTRO D C, WALKER I, GLOCKER B. Causality matters in medical imaging[J]. Nature Communications, 2020, 11(1): 3673.
- [33] DRAGOMIR S. A survey of reverse inequalities for f-divergence measure in information theory[M]//DARAS N J, RASSIAS M T. Computation, Cryptography, and Network Security. Cham: Springer, 2015: 209-251.
- [34] DRAGOMIR S S. A refinement of Jensen's inequality with applications for f-divergence measures[J]. Taiwanese Journal of Mathematics, 2010, 14(1): 153-164.
- [35] NAMKOONG H, DUCHI J C. Stochastic gradient methods for distributionally robust optimization with f-divergences[C]//Proceedings of the 30th International Conference on Neural Information Processing Systems. Barcelona: Curran Associates Inc., 2016: 2216-2224.
- [36] ESFAHANI P M, KUHN D. Data-driven distributionally robust optimization using the wasserstein metric: performance guarantees and tractable reformulations [J]. Mathematical Programming, 2018, 171(1): 115-166.
- [37] BLANCHET J, KANG Yang, MURTHY K. Robust wasserstein profile inference and applications to machine learning[J]. Journal of Applied Probability, 2019, 56(3): 830-857.
- [38] GRETTON A, BORGWARDT K M, RASCH M, et al. A kernel method for the two-sample-problem[C]// Proceedings of the 20th International Conference on Neural Information Processing Systems. Cambridge: MIT

- Press, 2006: 513-520.
- [39] GRETTON A, BORGWARDT K M, RASCH M J, et al. A kernel two-sample test[J]. *Journal of Machine Learning Research*, 2012, 13: 723-773.
- [40] BIŃKOWSKI M, SUTHERLAND D J, ARBEL M, et al. Demystifying MMD GANs[DB/OL]. (2021-01-14) [2025-01-28]. <https://doi.org/10.48550/arXiv.1801.01401>.
- [41] KUTNER M, NACHTSHEIM C, NETER J, et al. *Applied linear statistical models*[M]. 5th ed. New York: McGraw-Hill/Irwin, 2004: 1-9.
- [42] Yulara Solar System DKA Solar Centre [EB/OL]. (2020-01-14)[2025-01-28]. <https://dkasolarcentre.com.au/locations/yulara/>.
- [43] LIAO Wenlong, BAK-JENSEN B, PILLAI J R, et al. A review of graph neural networks and their applications in power systems[J]. *Journal of Modern Power Systems and Clean Energy*, 2022, 10(2): 345-360.
- [44] WOLPERT D H. The lack of a priori distinctions between learning algorithms[J]. *Neural Computation*, 1996, 8(7): 1341-1390.
- [45] FU Jian, ZHONG Yadong, YANG Feng. Adversarial domain generalization with MixStyle[C]//2022 International Conference on Advanced Robotics and Mechatronics (ICARM), Guilin, China, 2022: pp. 379-385.
- [46] XU Renzhe, CUI Peng, SHEN Zheyang, et al. Why stable learning works? A theory of covariate shift generalization [DB/OL]. (2023-10-17)[2025-01-28]. <https://doi.org/10.48550/arXiv.2111.02355>.
- [47] CUI Peng, ATHEY S. Stable learning establishes some common ground between causal inference and machine learning[J]. *Nature Machine Intelligence*, 2022, 4(2): 110-115.
- [48] LI Jie, WANG Yu, ZI Yanyang, et al. Whitening-Net: a generalized network to diagnose the faults among different machines and conditions[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2022, 33(10): 5845-5858.
- [49] KARPATNE A, JIA Xiaowei, KUMAR V. Knowledge-guided machine learning: current trends and future prospects[DB/OL]. (2024-03-01)[2025-01-28]. <https://doi.org/10.48550/arXiv.2403.15989>.
- [50] 乔骥, 郭剑波, 范士雄, 等. 人在回路的电网调控混合增强智能初探: 基本概念与研究框架[J]. *中国电机工程学报*, 2023, 43(1): 1-14.
- QIAO Ji, GUO Jianbo, FAN Shixiong, et al. Human-in-the-loop hybrid-augmented intelligence method for power system dispatching: basic concept and research framework[J]. *Proceedings of the CSEE*, 2023, 43(1): 1-14(in Chinese).
- [51] KHARAZI S, AMJADY N, NEJATI M, et al. A new closed-loop solar power forecasting method with sample selection[J]. *IEEE Transactions on Sustainable Energy*, 2024, 15(1): 687-698.
- [52] PLATT J. A resource-allocating network for function interpolation[J]. *Neural Computation*, 1991, 3(2): 213-225.
- [53] ZHANG Tengfei, MA Fumin, PENG Chen, et al. A very-short-term online PV power prediction model based on RAN with secondary dynamic adjustment[J]. *IEEE Transactions on Artificial Intelligence*, 2023, 4(5): 1214-1224.
- [54] 肖白, 黄钰茹, 姜卓, 等. 数据匮乏场景下采用生成对抗网络的空间负荷预测方法[J]. *中国电机工程学报*, 2020, 40(24): 7990-8001.
- XIAO Bai, HUANG Yuru, JIANG Zhuo, et al. The method of spatial load forecasting based on the generative adversarial network for data scarcity scenarios [J]. *Proceedings of the CSEE*, 2020, 40(24): 7990-8001(in Chinese).
- [55] MILLER J P, TAORI R, RAGHUNATHAN A, et al. Accuracy on the line: on the strong correlation between out-of-distribution and in-distribution generalization [DB/OL]. (2021-10-07)[2025-01-28]. <https://doi.org/10.48550/arXiv.2107.04649>.
- [56] BAEK C, JIANG Yiding, RAGHUNATHAN A, et al. Agreement-on-the-line: predicting the performance of neural networks under distribution shift[C]//Proceedings of the 36th International Conference on Neural Information Processing Systems. New Orleans: Curran Associates Inc., 2022: 19274-19289.
- [57] LIU Jiashuo, WANG Tianyu, CUI Peng, et al. Rethinking distribution shifts: empirical analysis and inductive modeling for tabular data[DB/OL]. (2024-11-13) [2025-01-28]. <https://doi.org/10.48550/arXiv.2307.05284>.
- [58] HAARNOJA T, ZHOU A, HARTIKAINEN K, et al. Soft actor-critic algorithms and applications [DB/OL]. (2019-01-29)[2025-01-28]. <https://doi.org/10.48550/arXiv.1812.05905>.
- [59] 朱琼峰, 李家腾, 乔骥, 等. 人工智能技术在新能源功率预测的应用及展望[J]. *中国电机工程学报*, 2024, 43(8): 3027-3047.
- ZHU Qiongfeng, LI Jiateng, QIAO Ji, et al. Application and prospect of artificial intelligence technology in

renewable energy forecasting[J]. Proceedings of the CSEE, 2024, 43(8): 3027-3047(in Chinese).

[60] YU Yixiao, WANG Mengxia, YAN Fangqing, et al. Improved convolutional neural network-based quantile regression for regional photovoltaic generation probabilistic forecast[J]. IET Renewable Power Generation, 2020, 14(14): 2712-2719.

附录 A 数据分布量化及 OODG 能力评估方法说明

1) 数据分布偏移量化方法说明。

f -散度、Wasserstein 距离、MMD 和拟合优度测试(如: R^2)等方法公式如下:

$$D_F[P_{tr}(X)|P_{te}(X)] = \int P_{te}(X) f\left[\frac{P_{tr}(X)}{P_{te}(X)}\right] dX \quad (A1)$$

$$W(P_{tr}(X), P_{te}(X)) = \inf_{\gamma \sim \Pi(P_{tr}(X), P_{te}(X))} \mathbb{E}_{x, y \sim \gamma} [\|x - y\|] \quad (A2)$$

$$MMD[F, P_{tr}(X), P_{te}(X)] = \sup_{\|f\|_{\mathcal{H}} \leq 1} \mathbb{E}_{P_{tr}(X)}[f(x)] - \mathbb{E}_{P_{te}(X)}[f(y)] \quad (A3)$$

$$R^2 = 1 - \frac{\sum (X_{tr} - X_{te})^2}{\sum (X_{tr} - X_{mean})^2} \quad (A4)$$

式中: $\inf$ 表示为求下界; $\Pi P_{tr}(X), P_{te}(X)$ 为分布 $P_{tr}(X)$ 和 $P_{te}(X)$ 的所有可能的联合分布集合, 可以从每个可能的联合分布 γ 中得到一个样本 x 和 y ; $\|x - y\|$ 表示样本的距离; $\sup$ 表示为求上界; $x_{tr}, x_{te} \sim P_{tr}(X), P_{te}(X)$; X_{mean} 为 X_{tr} 的平均值。

2) OODG 能力评估方法说明。

①模型鲁棒性: 使用分布式鲁棒定义不确定性集中最坏情况的风险。

$$\mathfrak{R}(\theta) = \sup_{Q \in \mathcal{P}(P_{tr})} \mathbb{E}_{X, Y \sim Q} [l(f_{\theta}(X), Y)] \quad (A5)$$

式中 $\mathcal{P}(P_{tr}) = \{Q: \text{Dist}(Q, P_{tr}) \leq \rho\}$ 表示围绕训练分布 P_{tr} 的不确定集。直观地说, 将测试数据分布封装在不确定集合中来反映模型的泛化性能。

②模型稳定性: 测量不稳定性。

$$s(\theta, P_{tr}) = \inf \{-D_{KL}(P||P_{tr}): \mathbb{E}_Q(l) \geq t\} \quad (A6)$$

式中: $D_{KL}(\cdot||\cdot)$ 为 KL 散度; $t > 0$ 为阈值; l 为模型预测误差。

③模型不变性: 不变风险最小化使用一阶信息来量化任何违反不变性原则的行为, 不变性惩罚定义为

$$\sum_{e \in \mathcal{E}_{tr}} \left\| \nabla_{\omega} R^e(\omega, \Phi) \right\|^2 \quad (A7)$$

式中: $\mathcal{E}_{tr}$ 为训练环境集合; R^e 为环境 e 中的预测风险; ω, Φ 为线性分类器。当度量为 0 时, 表明学习到的模型在整个训练环境中最优。

④模型平稳性: 最广泛采用的平稳度度量是最小化锐度 (sharpness-aware minimization, SAM), 定义为给定点和邻域点之间的最大损失间隙。

$$R_{\rho}^{SAM}(\theta) = \max_{\theta' \in B(\theta, \rho)} [\hat{L}(\theta') - \hat{L}(\theta)] \quad (A8)$$

式中: θ 为给定的参数点; ρ 为邻域球半径; $\hat{L}$ 为经验损失。

⑤模型置信度: 置信度是模型得分低于阈值 t 的测试样本的期望分数。

$$ATC_{P_{te}}(s) = \mathbb{E}_{X \sim P_{te}} [I[s(f_{\theta}(x)) < t]] \quad (A9)$$

式中: f_{θ} 为输出结果所对应的概率向量; s 为分数函数, 一般选择为向量中最大置信度值或负熵。



尚宇炜

在线出版日期: 2025-05-13。

收稿日期: 2025-02-05。

作者简介:

尚宇炜(1991), 男, 博士, 高级工程师, 研究方向为智能配用电与分布式能源, shangyuwei@epri.sgcc.com.cn;

*通信作者: 王思懿(1996), 女, 博士研究生, 研究方向为分布式电源及负荷预测、配电网运行与控制, siyiwang006@163.com;

郭剑波(1960), 男, 博士, 中国工程院院士, 教授级高级工程师, 研究方向为电力系统规划、运行分析和电网可靠性等, guojb@epri.sgcc.com.cn;

盛万兴(1965), 男, 博士, 研究员, 研究方向为配电网运行与控制、分布式发电和微电网、能源信息物理系统等, wxsheng@epri.sgcc.com.cn;

梁英(1975), 女, 博士, 教授级高级工程师, 研究方向为智能配用电与分布式能源, liangy@epri.sgcc.com.cn。

(编辑 乔宝榆, 张蕾)