

基于特征降维与分块的输电网概率最优潮流深度学习方法

武新章,郭苏杭,代伟,王泽宇,赵子巍,石博臣,张冬冬
(广西大学 电气工程学院,广西 南宁 530004)

摘要:概率最优潮流需要对非线性最优潮流问题进行重复求解,计算量较大,从而限制了其应用。提出一种基于特征降维、分块和深度神经网络辅助预测的最优潮流两阶段求解方法。在第一阶段,提出基于深度神经网络的最优潮流部分关键决策变量的优先辨识策略,以解决深度学习中因特征维度过高而导致的数值湮没问题,进而以最优潮流的结果特征为导向,基于关联性分析和聚类分析挖掘最优潮流输入与输出特征的关联性匹配度,并构建样本数据的分块特征库,以降低学习难度。在第二阶段,利用深度神经网络完成部分关键决策变量的分块映射,基于潮流模型恢复剩余状态变量,并对计算结果不收敛、不满足约束的情况进行修正,以恢复可行性。根据最优潮流两阶段求解方法构建概率最优潮流求解方法。仿真结果表明所提方法在最优潮流、概率最优潮流的求解速度和求解精度上均有较好的表现。

关键词:最优潮流;概率最优潮流;深度神经网络;关联分析;聚类分析;潮流修正

中图分类号:TM73

文献标志码:A

DOI:10.16081/j.epae.202302004

0 引言

概率最优潮流(probabilistic optimal power flow, POPF)是一种解决电力系统运行不确定性问题的重要分析工具^[1],其本质是对非线性、非凸最优潮流(optimal power flow, OPF)问题的多次求解^[2],计算量较大。随着高渗透可再生能源接入电网,不确定源增多,这使OPF计算次数增多^[3],POPF耗时更长^[4]。因此,寻找一种兼顾精度和速度的OPF求解方法成为必要。

OPF的传统计算方法可分为数值方法和启发式方法^[5]。数值方法通常可将OPF进行线性化处理,以降低模型的复杂度,提高求解速度,但求解精度较低,且存在求解结果不可行的情况^[6]。启发式方法,如遗传算法、粒子群优化算法等,具有全局搜寻的能力,可求得全局最优解,但通常需要进行大量的迭代工作,计算速度较慢,该方法难以应用于含高比例新能源的大规模电力系统^[7]。

近年来,数据驱动方法在OPF的求解中受到广泛关注^[8-9]。相较于传统计算方法,该方法利用海量的历史数据深度挖掘负荷变量与决策变量之间的复杂函数关系,以直接映射的方式获取OPF的求解方案^[10],避免了大量的迭代计算,有效提高了计算效率^[11-12]。文献[13]以拉格朗日对偶方法构造损失函

数,并将热启动初始方案作为网络学习参考,计算速度提升明显。文献[14]利用深度神经网络(deep neural network, DNN)匹配优化器和优化参数,捕捉参数导数在DNN中的敏感度,以提高DNN的泛化能力和约束满足能力,但未考虑新能源的影响。文献[15]基于堆栈式极限学习机将OPF问题的求解过程分解为多个阶段,并在每个阶段进行偏差修正,在求解速度和精度上均取得了较好的效果,但不能保证求解可行性,且会损失系统的潮流平衡性。文献[16]将OPF中的目标函数和物理约束融入强化学习的动作值函数中,并利用解析方法更新DNN参数,以满足求解结果的最优性和可行性。然而,随着新能源的大量接入,OPF数据样本呈现出高复杂度、高波动性的特征,上述数据驱动方法缺少特征优化的提取过程,这使DNN的学习负担大幅加重,不利于函数拟合和变量预测。

为保证电力系统的潮流平衡性,降低直接映射OPF全部待求量的学习复杂度,现有数据驱动方法常采用DNN模型和潮流模型相结合的方式。文献[17]基于DNN模型分块映射出电压幅值和相角,并根据线路潮流方程重构出有功出力和无功出力,降低了学习难度。文献[18]基于DNN进行PV节点的电压预测,并基于确定的潮流模型进行其他待求量的重构。文献[19]进一步放大不等式约束范围,通过可靠的样本集提高预测结果的约束满足能力。然而,上述方法在DNN模型阶段均采用多电压映射方式,重构过程会导致电压预测误差的叠加放大,给重构量特别是线路无功功率的求解精度带来很大影响。

针对上述问题,本文提出基于特征降维与分块的POPF深度学习方法。首先,提出OPF部分关键

收稿日期:2022-06-19;修回日期:2022-12-27

在线出版日期:2023-02-06

基金项目:国家自然科学基金资助项目(52107082);广西科技基地和人才专项(桂科AD22080052)

Project supported by the National Natural Science Foundation of China(52107082) and Guangxi Science and Technology Base and Talent Special Project(AD22080052)

决策变量的优先辨识策略,并基于关联性分析理论得到DNN输入与输出的相关性匹配度,进而得到部分关键决策变量的降维关联变量组;其次,以OPF的结果特征为导向,基于聚类分析深度挖掘数据样本中的潜在规律,进而构建样本数据的分块特征库,并完成关键决策变量的分块预测;然后,对OPF全部待求量进行重构,并对求解方案进行判断,若可行则求解结束,否则借助OPF数学模型进行修正,以恢复可行性;最后,根据两阶段求解思想构建POPF求解策略,并通过仿真验证所提方法在OPF、POPF的求解精度和速度以及可行性保障上均有较大优势。

1 OPF问题描述及研究方法框架

1.1 OPF求解模型

OPF数学模型可表示为:

$$\min C(\mathbf{P}_g) \quad (1)$$

$$\text{s.t. } F^p(\mathbf{P}_g, \mathbf{P}_d, \mathbf{V}, \boldsymbol{\theta}) = 0 \quad (2)$$

$$F^q(\mathbf{Q}_g, \mathbf{Q}_d, \mathbf{V}, \boldsymbol{\theta}) = 0 \quad (3)$$

$$L(\mathbf{P}_g, \mathbf{Q}_g, \mathbf{P}_d, \mathbf{Q}_d, \mathbf{V}, \boldsymbol{\theta}) \leq 0 \quad (4)$$

式中: $C(\cdot)$ 为发电机的发电成本函数; $F^p(\cdot)$ 、 $F^q(\cdot)$ 分别为有功、无功功率平衡函数; $L(\cdot)$ 为安全约束函数; \mathbf{P}_g 、 \mathbf{Q}_g 、 \mathbf{P}_d 、 \mathbf{Q}_d 、 \mathbf{V} 、 $\boldsymbol{\theta}$ 分别为发电机有功出力向量、发电机无功出力向量、有功负荷向量、无功负荷向量、节点电压幅值向量、节点电压相角向量。式(1)为OPF的目标函数;式(2)—(4)为OPF的约束,包含发电机出力约束、电压约束、功率流约束。

1.2 研究方法框架

为避免直接映射带来的过高的学习复杂度,并保证OPF求解方法的潮流平衡性,本文基于DNN模型和潮流模型设计一种两阶段求解方法,即先利用DNN预测出部分决策变量 \mathbf{X} (包括发电机有功出力 \mathbf{P}_g 、发电机无功出力 \mathbf{Q}_g 、平衡节点电压幅值 V_0),再对系统潮流进行计算,以获得OPF全部待求量。

定义 $\{\mathbf{M}, \mathbf{N}\}$ 为模型数据集, $\{\mathbf{P}_d, \mathbf{Q}_d\}$ 构成 \mathbf{M} ,为DNN输入特征, $\{\mathbf{P}_g, \mathbf{Q}_g, \mathbf{V}_s\}$ 构成 \mathbf{N} ,为DNN输出特征,其中 \mathbf{V}_s 为平衡节点及与其相连节点的电压幅值向量。本文建立由结果导向的数据预处理、DNN辅助预测和潮流重构的两阶段模型求解框架,具体如下。

1)第一阶段。基于 \mathbf{M} ,对 \mathbf{P}_g 、 \mathbf{Q}_g 、 \mathbf{V}_s 进行关联性分析,根据输入、输出关联性匹配度的大小生成3对关联变量组,以降低样本特征维度和DNN模型的复杂度。基于结果导向的聚类分析方法,对各关联变量组样本进行划分,生成模块化特征库,以捕捉大样本中潜在类别信息,降低样本的复杂性和波动性。

2)第二阶段。在每个特征库中进行DNN训练,

并预测生成 \mathbf{X} 。根据 \mathbf{X} ,基于潮流模型恢复剩余状态变量。对于计算结果不收敛、不满足约束的样本,基于OPF数学模型进行修正,以恢复可行性。

2 结果导向的数据预处理

2.1 两阶段组合方式的选取

DNN与潮流模型的结合可降低直接映射全部待求量带来的复杂性,同时保留系统潮流平衡性,且潮流模型并不会引入新的误差,有利于求解精度的提高。现有两阶段求解方法主要有以下2种组合。

1)组合1。将负荷变量 $\{\mathbf{P}_d, \mathbf{Q}_d\}$ 作为DNN的输入,预测出系统中PV节点的有功出力、电压幅值以及平衡节点的电压幅值,此时PQ节点为负荷节点,注入的有功功率和无功功率已知,满足潮流计算条件,再通过有功、无功功率平衡方程式(2)、(3)进行计算即可获得全网待求量^[18]。

2)组合2。将负荷变量 $\{\mathbf{P}_d, \mathbf{Q}_d\}$ 作为DNN的输入,预测出系统中全部节点的电压幅值和相角,此时通过对节点电压幅值和相角与支路功率的关系方程式(5)、(6)进行计算即可获得支路有功功率和无功功率,进而根据节点功率平衡求得发电机的有功出力和无功出力^[12]。

$$P_{ij} = V_i^2 G_{ij} - V_i V_j (\cos \theta_{ij} G_{ij} + \sin \theta_{ij} B_{ij}) \quad (5)$$

$$Q_{ij} = -V_i^2 (B_{ij} + B_{i0}) - V_i V_j (\sin \theta_{ij} G_{ij} - \cos \theta_{ij} B_{ij}) \quad (6)$$

式中: P_{ij} 、 Q_{ij} 分别为节点 i 、 j 间支路的有功、无功功率; V_i 、 V_j 分别为节点 i 、 j 的电压幅值; G_{ij} 、 B_{ij} 分别为节点 i 、 j 间支路的电导、电纳; θ_{ij} 为节点 i 、 j 间的电压相角差; B_{i0} 为节点 i 的接地电纳。

在OPF的求解过程中,节点电压幅值通常采用标幺值,因此OPF其他待求量对节点电压幅值的准确度较为敏感,较小的电压幅值偏差就会引起其他待求量较大的求解误差。另外,当预测的节点电压幅值数量较多时,各幅值的映射误差会相互叠加,使求解精度更低,例如,式(5)、(6)中的线路有功、无功功率误差会因节点电压幅值平方、节点电压幅值间的乘积而升高。

针对以上问题,本文提出组合3。将负荷变量 $\{\mathbf{P}_d, \mathbf{Q}_d\}$ 作为DNN的输入,预测出发电机的有功出力 \mathbf{P}_g 、发电机的无功出力 \mathbf{Q}_g 、平衡节点电压幅值 V_0 ,此时系统仅含有平衡节点和PQ节点且平衡节点电压已知,OPF求解问题变成了潮流计算问题,再基于潮流模型,根据有功、无功功率平衡方程式(2)、(3)求得全网功率流。

组合3中采用单电压的映射方式,仅选择波动性较小的平衡节点电压作为电压预测输出,并作为全网电压参考进行潮流重构,这样易于DNN学习和映射,避免了潮流重构过程中因节点电压误差叠加

而造成的误差放大。

虽然组合3中只需计算平衡节点的电压幅值,但考虑到节点电压幅值间具有一定的关联性,且往往网络拓扑中直接相连的节点电压关联性较大,因此本文利用DNN预测平衡节点及与其直接相连节点的电压幅值(记为 V_s),但仅使用平衡节点电压。

2.2 关联性分析和结果导向的聚类处理

若将 P_g, Q_g, V_s 三者组合作为输出,则输出维度过高,学习负担过重,而将 P_g, Q_g, V_s 三者拆分并分别进行训练和预测,可降低DNN的输出维度以及输入、输出之间的复杂性。

为进一步降低输入维度对网络学习的影响,避免无效特征湮没输入、输出间的映射关系,基于典型相关性理论,本文计算训练集输入量 M_x 中各变量与 $P_g(Q_g, V_s)$ 的相关性,根据输入特征的贡献度大小筛选有效特征集 $M_1(M_2, M_3)$,以实现特征优化和降维,降低DNN的学习负担。

实际的OPF样本数据复杂度高、波动性大,而DNN较难拟合高波动性的样本,若直接采用原始样本数据进行DNN训练,则会增加学习难度。通过对DNN样本数据进行聚类可挖掘样本数据的内在规律,有效降低样本复杂度,在每个类中分别进行神经网络训练,可降低训练难度,提高预测精度。

为提高DNN输出特征的映射精度,在聚类分析时,要在充分挖掘输出特征规律性的前提下对样本进行分类。由于OPF求解过程中目标函数、潮流约束等的影响,负荷变量(输入特征)和决策变量(输出特征)的数值变化规律有所差别,例如,2个数值接近的负荷变量,可能因约束的限制而产生2个数值相差很大的决策变量,因此,若依据输入特征对样本进行聚类,则输入特征和输出特征的内在规律差异性会使输出特征的类别划分不合理,不能准确体现输出特征的内在数值规律。为此,本文以样本输出特征为导向,先对输出特征进行聚类,再根据输入特征与输出特征的对应关系得到输入特征所属类别信息,进而完成聚类。结果导向的聚类分析过程如附录A图A1所示。首先,将每个样本的输入特征和输出特征根据OPF中负荷变量和决策变量的对应关系进行匹配,同时利用 k -means算法对输出量进行聚类,并将聚类指标最大时的聚类数作为最佳聚类数,进而得到 z 个类别,即类1— z ;然后,基于负荷变量和决策变量的对应关系,在类1— z 中搜寻与输出特征对应的输入特征,完成类别提取,得到模块数据1— z ;最后,对每个模块数据的输入部分求均值,得到各模块中心值 $K_1—K_z$,并将其记为关联聚类中心。

2.3 结果导向的数据预处理

根据2.1、2.2节的分析过程,确定结果导向的数据预处理过程,如附录A图A2所示,具体如下。

1)针对数据特征维度过高的问题,基于典型相关性分析理论,在训练集输入量 M_x 中分别求得 P_g, Q_g, V_s 的最大关联集,记为负荷变量 M_1, M_2, M_3 (均为 M_x 的子集),进而构建关联变量组 $[M_1, P_g], [M_2, Q_g], [M_3, V_s]$,并生成表征DNN输入、输出匹配关系的关联变量组信息。

2)对 P_g, Q_g 进行数据预聚类处理。因发电机成本函数、网络结构等限制,发电机有功、无功出力会出现满发或零发的情况,且该出力数据对聚类无意义,需要对其进行剔除,将剔除后的有功、无功出力分别记为 P'_g, Q'_g 。

3)基于结果导向的聚类分析。分别对 $[M_1, P'_g], [M_2, Q'_g], [M_3, V_s]$ 生成3组模块,即模块1-1—1- l 、模块2-1—2- m 、模块3-1—3- n ,并记录各模块的关联聚类中心信息,其中 l, m, n 分别为3组模块的数量。

4)基于关联变量组信息,将测试集进行关联变量组匹配,并在各组内求得负荷变量到该组各关联聚类中心的距离,再按照距离最近原则进行模块划分,以完成对测试集所属模块的划分归类。

3 DNN辅助预测和潮流重构

3.1 DNN辅助预测

类似于OPF中负荷变量和决策变量之间的对应关系,DNN的输入、输出也为一一对应的变量对,且二者之间存在极为复杂的非线性关系。以图1所示的全连接DNN为例,DNN由 $r+1$ 层神经元堆叠而成, $Y_0, Y_1—Y_r, Z_{out}$ 分别为DNN输入、第1— r 层神经元输出、DNN输出, $f_1—f_{r+1}$ 分别为第1— $r+1$ 层神经元输入与输出的函数。由图可知, Y_0 和 Z_{out} 组成了一一对应的变量对 $[Y_0, Z_{out}]$,二者之间的函数关系可表示为:

$$Z_{out} = f_{r+1}(f_r(\dots f_1(Y_0))) \quad (7)$$

$$f_u(Y_{u-1}) = S_u(W_u Y_{u-1} + b_u) \quad u=1, 2, \dots, r+1 \quad (8)$$

式中: S_u 为第 u 层的激活函数; W_u, b_u 分别为第 $u-1$ 层输出与第 u 层输出之间的权重矩阵和阈值矩阵; Y_{u-1} 为第 $u-1$ 层输出(也为第 u 层输入)。由于激活函数引入了非线性关系,这使其具有无限接近非线性函数的能力。

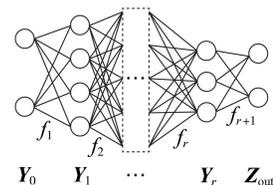


图1 全连接DNN

Fig.1 Fully connected DNN

利用训练集和测试集构成的 $l+m+n$ 组模块化数据进行DNN训练和预测,使 $M_1—M_3$ 替代 Y_0 作为

DNN的输入,相应地,使 P'_g, Q'_g, V_s 替代 Z_{out} 作为DNN的输出,最终得到满足潮流计算需求的部分决策变量 X 。

3.2 潮流重构

基于 X 和潮流模型对全网潮流进行重构。由于发电机出力已知,全网的PV节点将转化为PQ节点。此时将 X 中的平衡节点电压幅值 V_0 作为全网电压的参考,通过对式(2)、(3)进行潮流计算即可获得全网待求量。

两阶段过程均可能引入不可行性:在第一阶段中,DNN会天然引入误差,生成的 X 可能会违反不等式约束;在第二阶段中,即使潮流计算收敛,但由于计算时未考虑线路功率限制,因此结果的可行性需进一步评估确定,若潮流计算不收敛,则需进行潮流修正,以恢复可行性。修正过程如附录A图A2所示。判断由神经网络预测出的部分决策变量 X 是否越限,若是,则进行潮流修正,否则进行潮流重构。若潮流重构不收敛或潮流重构收敛但求得的其他待求量越限,则进行潮流修正;若潮流重构收敛且求得的其他待求量不越限,则输出最优潮流全部待求量。

4 仿真分析

为验证本文模型的有效性,在IEEE 30、IEEE 118、IEEE 2383节点系统算例上进行仿真分析,仿真过程中的硬件为Intel(R) Core(TM) i7-10700F CPU @ 2.90 GHz 32 GB RAM。

4.1 算例基本设置

本文设置IEEE 30、IEEE 118、IEEE 2383节点系统的负荷在其额定值的80%至1.2倍之间的范围内随机波动,接入的新能源占负荷额定值的20%,风速、光照强度分别服从Weibull分布和Beta分布,风电、光电转换模型见文献[20-21],本文将接入的新能源设置在负荷位置。3个算例的数据集总样本数均为100 000,且均按照总样本数的70%、15%、15%的比例将数据集分别划分为训练集、验证集、测试集。

为充分验证本文所提模型的有效性,设计以下对比方案:方案 R_0 ,利用全连接DNN同时映射 P_g, Q_g, V_s ;方案 R_1 ,基于结果导向的数据预处理,利用全连接DNN分别映射 P_g, Q_g, V_s ;方案 R_2 ,利用全连接DNN直接映射OPF全部待求量;方案 R_3 ,利用长短期记忆神经网络直接映射OPF全部待求量;方案 R_4 ,利用卷积神经网络直接映射OPF全部待求量;方案 R_5 ,基于方案 R_1 进行两阶段OPF求解(组合1);方案 R_6 ,基于方案 R_1 进行两阶段OPF求解(组合2);方案 R_7 ,基于方案 R_1 进行两阶段OPF求解(组合3);方案 R_8 ,采用文献[18]中的方法求解OPF(基于DNN进行组合1的两阶段求解);方案 R_9 ,采用内点法求

解OPF。

4.2 模型有效性验证

为验证本文结果导向数据预处理的有效性,分别在IEEE 30、IEEE 118、IEEE 2383节点系统中对方案 R_0, R_1 中的神经网络进行训练分析,得到如附录B图B1—B3所示的DNN训练过程图。由图可知,结果导向的数据预处理有效提高了DNN的训练效果,具体分析如附录B所示。

为证明本文所提方法在OPF求解精度上的优势,分别基于IEEE 30、IEEE 118、IEEE 2383节点系统算例,利用方案 $R_2—R_7, R_9$ 进行求解,并在求解结果中随机挑选2台发电机的有功、无功出力,2条支路的有功、无功功率,以及2个节点的电压幅值、电压相角进行对比,如附录C表C1、C2所示。由表可知:本文方法的求解结果与内点法的求解结果最接近;在IEEE 30节点系统算例中,发电机有功出力、支路有功功率的预测误差均小于0.01 MW,发电机无功出力、支路无功功率的预测误差均小于0.01 Mvar,电压幅值、电压相角的预测误差分别小于0.0005 p.u.、0.001°;其他方案的求解结果与内点法的求解结果相差较大。对比方案 $R_2—R_4$ 可知:全连接DNN比长短期记忆神经网络、卷积神经网络的效果更佳,更适用于本文模型;由于循环神经网络具有记忆功能,不同样本会输出趋于一致的求解结果,这导致学习效果较差,因此长短期记忆神经网络并不适用于本文的非时序数据;卷积神经网络的局部提取特征能力割裂了系统负荷之间的全局特征,降低了其预测精度,而全连接DNN可提取系统中的全局负荷特征,捕捉全局输入量对神经网络输出量的影响,更适用于OPF的预测。对比方案 $R_5—R_7$ 可知:3种方案均为两阶段求解方法,其中方案 R_7 的效果最佳,由于潮流重构会造成电压预测误差的叠加放大,这使方案 R_5, R_6 的重构量精度较低,特别是线路功率流;只选取波动性较小的平衡节点电压作为电压预测输出,有利于DNN捕捉函数关系,减小潮流重构过程中因节点电压误差叠加而造成的误差放大。

4.3 动态最优潮流求解分析

为验证本文方法在动态最优潮流(dynamic optimal power flow, DOPF)中的适用性,本文基于IEEE 30、IEEE 118节点系统和广东某地区2016—2017年实际负荷构建DOPF样本,调度周期为24个时段,共收集100 000个样本,同样按照总样本数的70%、15%、15%将数据集分别划分为训练集、验证集、测试集,并设计以下对比方案:方案 R_{10} ,利用内点法进行DOPF的求解;方案 R_{11} ,利用DNN直接映射DOPF全部待求量;方案 R_{12} ,基于结果导向的数据预处理,利用DNN直接映射DOPF全部待求量;方案 R_{13} ,基于结果导向的数据预处理,利用全连接DNN

进行组合3的两阶段DOPF求解。

分别基于IEEE 30、IEEE 118节点系统算例,采用方案 R_{10} — R_{13} 进行求解,并选取IEEE 30节点系统算例中5号节点发电机在24个时段的有功出力、IEEE 118节点系统算例中65号节点发电机在24个时段的有功出力进行对比分析,结果分别如图2和附录C图C1所示。

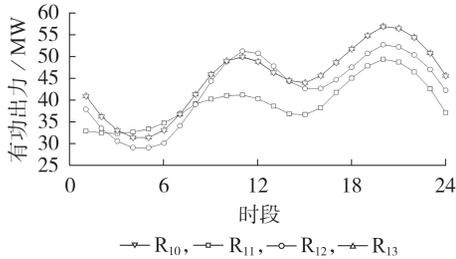


图2 IEEE 30节点系统算例结果对比

Fig.2 Comparison of example results of IEEE 30-bus system

由图可知:与方案 R_{11} 、 R_{12} 相比,本文方法所求得的24个时段的发电机有功出力曲线精度最高,该曲线和内点法得到的出力曲线接近重合;IEEE 30节点系统中各时段发电机出力的绝对误差均在0.04 MW内,相对误差均小于0.1%,IEEE 118节点系统中各时段发电机出力的绝对误差均在0.45 MW内,相对误差均小于0.15%。由此可知,本文方法同样适用于DOPF的求解且有较高的精度。此外,对比方案 R_{11} 、 R_{12} 可知,二者的求解效果均较差,但方案 R_{12} 得到的出力曲线更接近内点法得到的出力曲线,由此证明了结果导向数据预处理的有效性。

不同算例中采用方案 R_{10} — R_{12} 得到的指标对比如表1所示。表中: P_{grel} 、 Q_{grel} 、 P_{ijrel} 、 Q_{ijrel} 、 V_{rel} 、 θ_{rel} 、 E_{rel} 分别为发电机有功出力、发电机无功出力、支路有功功率、支路无功功率、电压幅值、电压相角,在各方案下本文方法对目标函数的求解值相较于内点法求解值的平均相对误差。由表可知,本文方法对DOPF所有待求量均有较好的求解效果。

此外,在取200个训练周期时,对于IEEE 30和IEEE 118节点系统算例,OPF问题的神经网络训练时间分别为308、366 s, DOPF问题的神经网络训练时间分别为869、1562 s。DOPF问题的神经网络训

练时间明显增加,这是由于相较于OPF, DOPF的特征维度增加,这使网络复杂度增大,训练时长增长,但离线训练的时间较为充足,并不影响在线校核与决策。

4.4 POPF求解分析

在POPF中,通常需要重复进行OPF计算,以满足计算精度要求。假设POPF需要进行15000(本文的测试集样本数)次重复计算,采用方案 R_2 — R_8 对POPF进行仿真求解,得到各变量的平均相对误差,如附录C表C3所示。由表可知:本文方法对POPF具有较高的求解精度,且相比文献[18]方法的求解精度有明显提升;IEEE 30节点系统中各指标的相对误差均小于 10^{-3} ,电压幅值和电压相角的相对误差更是小于 10^{-4} ;IEEE 2383节点系统中有功出力、无功出力、功率流的相对误差均小于 5×10^{-3} ,电压幅值和电压相角的相对误差均小于 3×10^{-4} 。在取200个训练周期时,对于IEEE 30、IEEE 118、IEEE 2383节点系统算例,文献[18]方法的训练时间分别为1236、1659、3029 s,本文方法的训练时间分别为308、366、1116 s,相较于文献[18]方法,本文方法的训练时间明显缩短。

关联性分析降低了DNN输入和输出的数据维度,从而可提取低维的最优特征集,弱化DNN的输入和输出之间复杂的函数关系。数据聚类分析降低了数据之间的复杂度,将具有相似性的样本归为同一类中。关联性分析和数据聚类分析的结合很好地减轻了DNN的学习负担并降低了映射难度,从而使得 X 的预测精度大幅提升,训练时间明显缩短。此外,重构方法的改进提高了重构量的映射精度,使POPF获得了较好的求解效果。

各算例中采用方案 R_7 、 R_9 的计算时间对比如表2所示。表中: T_{R7} 、 T_{R9} 分别为方案 R_7 、 R_9 的计算时间; p 为方案 R_7 相较于方案 R_9 的速度提升倍数。由表可

表2 各算例中采用方案 R_7 、 R_9 得到的计算时间对比

Table 2 Comparison of calculation time obtained by Scheme R_7 and R_9 in each example

算例	T_{R7} / min	T_{R9} / min	p
IEEE 30	26.8253	1.8256	14
IEEE 118	33.5528	1.9368	16
IEEE 2383	525.9635	9.5382	54

表1 不同算例中采用方案 R_{10} — R_{12} 得到的指标对比

Table 1 Comparison of indexes obtained by Scheme R_{10} — R_{12} in different examples

算例	方案	$P_{\text{grel}} / \%$	$Q_{\text{grel}} / \%$	$P_{\text{ijrel}} / \%$	$Q_{\text{ijrel}} / \%$	$V_{\text{rel}} / \%$	$\theta_{\text{rel}} / \%$	$E_{\text{rel}} / \%$
IEEE 30	R_{10}	8.8370	18.6099	19.2693	20.3256	0.9683	4.2346	6.6526
	R_{11}	7.7522	13.2369	13.2364	13.2659	0.3269	3.6948	4.2361
	R_{12}	0.0886	0.0783	0.0936	0.0845	0.0062	0.0241	0.0211
IEEE 118	R_{10}	16.3259	15.2657	15.3322	26.3269	0.7757	5.3627	12.0323
	R_{11}	11.2558	14.2366	13.2384	20.6941	0.1025	4.2392	10.8376
	R_{12}	0.0903	0.1036	0.0995	0.1037	0.0068	0.0297	0.0321

知,本文方法对 OPF 的求解速度有大幅提升,对于 IEEE 2383 节点系统,速度提升倍数达到 54。

为验证潮流修正的有效性,将数据集中每个变量的最大值、最小值分别作为该变量的约束上、下限,以模拟电力系统的重载情况,并按照方案 R₁ 训练好的神经网络重新求解,得到表 3 所示的潮流修正指标。表中: P_{pre} 为修正样本占总样本的比例; v 为修正样本数; T_{post} 为潮流修正时间。

表 3 潮流修正指标对比

Table 3 Comparison of power flow correction indexes

算例	$P_{pre} / \%$	v	T_{post} / s
IEEE 30	0.067	10	1.073 2
IEEE 118	0.140	21	2.818 4
IEEE 2383	0.387	58	119.919 6

由表 3 可知:修正样本所占比例并不高,对于 IEEE 30、IEEE 118、IEEE 2383 节点系统算例,修正样本的比例分别为 0.067%、0.140%、0.387%,即 15 000 个验证集样本中,需要修正的样本数分别为 10、21、58,潮流修正时间分别为 1.073 2、2.818 4、119.919 6 s,虽然增加了总样本的求解时间,但有效保证了生成结果的可行性,降低了电力系统的运行风险。

5 结论

针对高渗透可再生能源接入带来的 POPF 计算负担,本文提出基于特征降维与分块的输电网 POPF 深度学习,算例分析证明该方法具有以下优势。

1) 在结果导向的数据预处理中,关联性分析对输入、输出特征的降维和提取,降低了输入、输出之间的复杂性以及 DNN 的训练成本。聚类分析充分挖掘了样本数据的潜在规律,有效降低了样本数据的复杂性、波动性,减轻了 DNN 的学习负担,提高了映射精度。

2) 在两阶段求解模型中,DNN 的引入使 POPF 计算速度得到大幅提升。潮流重构模型有效减少了重构过程中的误差叠加,使重构量获得了更高的精度,并保证了解的可行性。

DNN 的引入仅作为数据驱动和两阶段求解模型相结合的可行性证明,本文并没有依据电力系统的物理特性寻找合适的 DNN 类型,且不同算例对应的 DNN 不可共享,这增加了运算负担,因此如何将 DNN 和电力系统的物理特性相结合,更好地减轻训练负担,是笔者未来要探索的方向。

附录见本刊网络版(<http://www.epae.cn>)。

参考文献:

[1] 余娟,杨燕,杨知方,等. 基于深度学习的概率能量流快速计算方法[J]. 中国电机工程学报,2019,39(1):22-30,317.

YU Juan, YANG Yan, YANG Zhifang, et al. Fast probabilistic energy flow analysis based on deep learning[J]. Proceedings of the CSEE, 2019, 39(1): 22-30, 317.

[2] 李国庆,陆为华,李赫,等. 基于模糊 C-均值聚类的时序概率潮流快速计算方法[J]. 电力自动化设备,2021,41(4):116-122.

LI Guoqing, LU Weihua, LI He, et al. Fast calculation method of time sequence probabilistic power flow based on fuzzy C-means clustering[J]. Electric Power Automation Equipment, 2021, 41(4): 116-122.

[3] 杨燕,杨知方,余娟,等. 基于深度学习的含不确定性 N-1 安全校核方法[J]. 中国电机工程学报,2021,41(8):2716-2725.

YANG Yan, YANG Zhifang, YU Juan, et al. Fast analysis of N-1 contingency screening with uncertainty scenarios based on deep learning[J]. Proceedings of the CSEE, 2021, 41(8): 2716-2725.

[4] LIN W, YANG Z F, YU J, et al. Toward fast calculation of probabilistic optimal power flow[C]//2020 IEEE Power & Energy Society General Meeting (PESGM). Montreal, QC, Canada: IEEE, 2020:1.

[5] HASAN F, KARGARIAN A, MOHAMMADI J. Hybrid learning aided inactive constraints filtering algorithm to enhance AC OPF solution time[J]. IEEE Transactions on Industry Applications, 2021, 57(2): 1325-1334.

[6] YANG Z F, ZHONG H W, BOSE A, et al. A linearized OPF model with reactive power and voltage magnitude: a pathway to improve the MW-only DC OPF[J]. IEEE Transactions on Power Systems, 2018, 33(2): 1734-1745.

[7] REDDY S S. Optimal power flow using hybrid differential evolution and harmony search algorithm[J/OL]. International Journal of Machine Learning & Cybernetics. [2022-11-06]. <https://link.springer.com/content/pdf/10.1007/s13042-018-0786-9.pdf>. DOI: 10.1007/s13042-018-0786-9.

[8] GUHA N, WANG Z, WYTOCK M, et al. Machine learning for AC optimal power flow[EB/OL]. [2022-11-06]. https://www.researchgate.net/publication/338017912_Machine_learning_for_optimal_power_flow#read.

[9] RAHMAN J, FENG C, ZHANG J. A learning-augmented approach for AC optimal power flow[J]. International Journal of Electrical Power & Energy Systems, 2021, 130: 106908.

[10] 武新章,赵子巍,代伟,等. 基于改进的 Transformer 神经网络辅助的两阶段机组组合决策方法[J/OL]. 电力自动化设备. [2022-11-06]. <https://doi.org/10.16081/j.epae.202209014>.

[11] YANG Y, YANG Z F, YU J, et al. Fast calculation of probabilistic power flow: a model-based deep learning approach[J]. IEEE Transactions on Smart Grid, 2020, 11(3): 2235-2244.

[12] VELLOSO A, VAN HENTENRYCK P. Combining deep learning and optimization for preventive security-constrained DC optimal power flow[J]. IEEE Transactions on Power Systems, 2021, 36(4): 3618-3628.

[13] FIORETTO F, MAK T. Predicting AC optimal power flows: combining deep learning and Lagrangian dual methods[EB/OL]. [2022-11-06]. <https://arxiv.org/abs/1909.10461>.

[14] SINGH M K, KEKATOS V, GIANNAKIS G B. Learning to solve the AC-OPF using sensitivity-informed deep neural networks[J/OL]. IEEE Transactions on Power Systems. [2022-11-06]. <https://ieeexplore.ieee.org/document/9917161>. DOI: 10.48550/arXiv.2103.14779.

[15] LEI X Y, YANG Z F, YU J, et al. Data-driven optimal power flow: a physics-informed machine learning approach[J]. IEEE Transactions on Power Systems, 2021, 36(1): 346-354.

[16] YAN Z M, XU Y. Real-time optimal power flow: a Lagrangian based deep reinforcement learning approach[J]. IEEE

- Transactions on Power Systems, 2020, 35(4):3270-3273.
- [17] HUANG W J, PAN X, CHEN M H, et al. DeepOPF-V: solving AC-OPF problems efficiently[J]. IEEE Transactions on Power Systems, 2022, 37(1):800-803.
- [18] PAN X, CHEN M, ZHAO T, et al. DeepOPF: a feasibility-optimized deep neural network approach for AC optimal power flow problems[DB/OL]. [2022-11-06]. <https://arxiv.org/pdf/2007.01002.pdf>.
- [19] ZAMZAM A S, BAKER K. Learning optimal solutions for extremely fast AC optimal power flow[C]//2020 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids. Tempe, AZ, USA: IEEE, 2020: 1-6.
- [20] QIN Z L, LI W Y, XIONG X F. Incorporating multiple correlations among wind speeds, photovoltaic powers and bus loads in composite system reliability evaluation[J]. Applied Energy, 2013, 110:285-294.
- [21] LI Y M, LI W Y, YAN W, et al. Probabilistic optimal power

flow considering correlations of wind speeds following different distributions[J]. IEEE Transactions on Power Systems, 2014, 29(4):1847-1854.

作者简介:



武新章

武新章(1968—),男,教授,博士,主要研究方向为人工智能在电力系统的应用、机器人、电磁理论(**E-mail**:xwu@gxu.edu.cn);

郭苏杭(1997—),男,硕士研究生,主要研究方向为人工智能在电力系统优化方面的应用(**E-mail**:1363054525@qq.com);

代伟(1990—),男,助理教授,博士,通信作者,研究方向为人工智能在电力系统中的应用、综合能源系统优化调度等

(**E-mail**:weidai2019@163.com)。

(编辑 王锦秀)

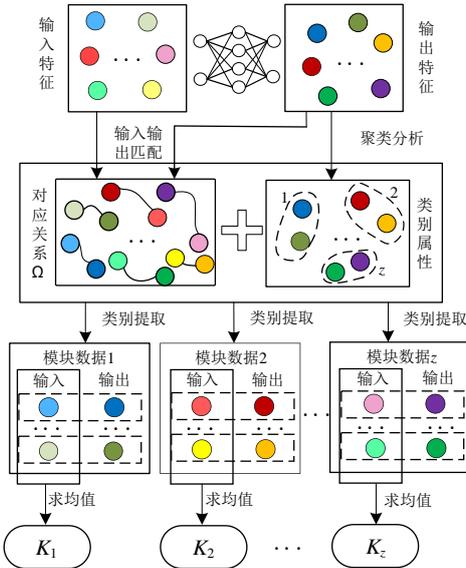
Feature dimension reduction and partitioning based deep learning method for probabilistic optimal power flow of transmission network

WU Xinzhang, GUO Suhang, DAI Wei, WANG Zeyu, ZHAO Ziwei, SHI Bochen, ZHANG Dongdong
(College of Electrical Engineering, Guangxi University, Nanning 530004, China)

Abstract: Probabilistic optimal power flow (POPF) requires repeated solving of the nonlinear optimal power flow (OPF) problem, thus its application is limited by large calculation amount. A two-stage solution method for OPF is proposed based on feature dimension reduction, partitioning and auxiliary forecasting of deep neural network (DNN). In the first stage, a DNN-based priority identification strategy for partial key decision variables of OPF is proposed, which solves the problem of numerical annihilation caused by too high feature dimension in the deep learning, further, guiding by the result characteristics of OPF, the correlation matching degree between input and output characteristics of OPF is extracted based on the correlation analysis and cluster analysis, and the block feature database of sample data is constructed to reduce the learning difficulty. In the second stage, DNN is used to complete the block mapping of partial key decision variables, the remaining state variables are recovered based on the power flow model, and the conditions that the calculation results do not converge or do not satisfy the constraints are corrected to restore the feasibility. A solution method of POPF is constructed according to the two-stage solution method of OPF. The simulative results show that the proposed method has good performance in the solving speed and accuracy of OPF and POPF.

Key words: optimal power flow; probabilistic optimal power flow; deep neural network; association analysis; cluster analysis; power flow correction

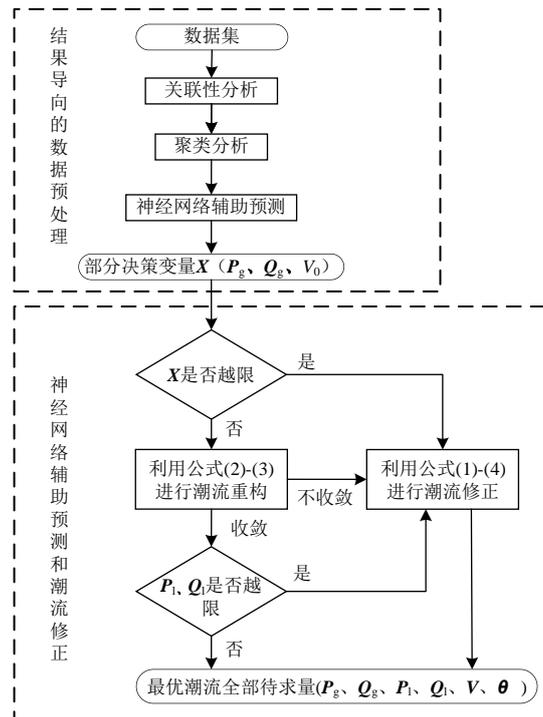
附录 A



注：单个样本中的输入特征、输出特征由色系相同但深浅不同的色块表示，分别对应于 OPF 求解过程中的负荷变量、决策变量。

图 A1 结果导向的聚类处理过程

Fig.A1 Dealing process of results-oriented clustering



注： P_1 、 Q_1 分别表示支路有功向量、支路无功向量。

图 A2 两阶段求解方法整体流程图

Fig.A2 Overall flowchart of two-stage solution method

附录 B

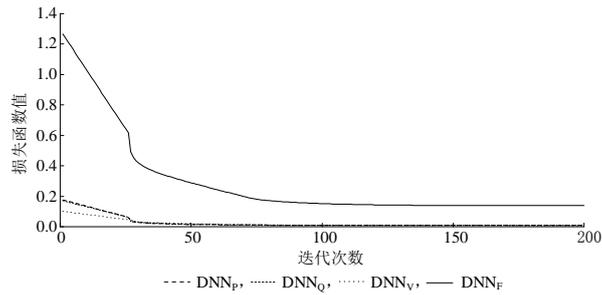


图 B1 IEEE 30 节点系统损失值与迭代次数的关系

Fig.B1 Relationship between loss value and iteration times of IEEE 30-bus system

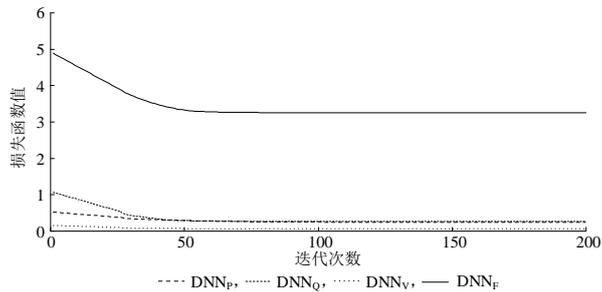


图 B2 IEEE118 节点系统损失值与迭代次数的关系

Fig.B2 Relationship between loss value and iteration times of IEEE 118-bus system

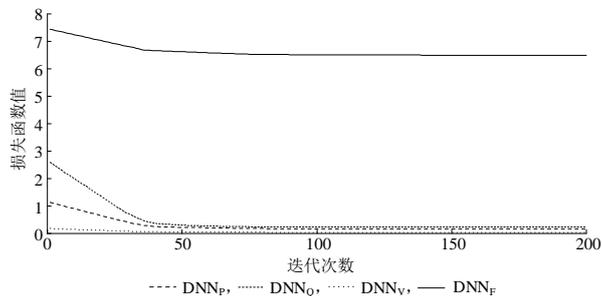


图 B3 IEEE 2383 节点系统损失值与迭代次数的关系

Fig.B3 Relationship between loss value and iteration times of IEEE 2383-bus system

基于本文结果导向的数据预处理方法，分别对 IEEE 30、IEEE 118、IEEE 2383 样本数据进行处理。将 R_0 中 DNN 记为 DNN_F ，将 R_1 中分别对变量 P_g 、 Q_g 、 V_s 进行预测的神经网络记为 DNN_P 、 DNN_Q 、 DNN_V 。对 DNN_P 、 DNN_Q 、 DNN_V 、 DNN_F 进行训练，得到如图 B1、图 B2、图 B3 所示的迭代次数和损失函数值之间的关系图，其中损失函数值为半均方误差。

由图 B1-B3 可知，本文所提方法在不同算例中均可达到更低的损失函数值，在 IEEE 30 节点系统中 DNN_P 、 DNN_Q 、 DNN_V 、 DNN_F 的损失函数值分别为 0.0076、0.0086、0.0072、0.1425，本文模型损失函数值较 DNN_F 分别下降 94%，94%，95%；在 IEEE 118 节点系统中 DNN_P 、 DNN_Q 、 DNN_V 、 DNN_F 的损失函数值分别为 0.2371、0.2619、0.0612、3.2654，本文模型损失函数值较 DNN_F 分别下降 93%，92%，98%；同样在 IEEE 2383 节点系统中 DNN_P 、 DNN_Q 、 DNN_V 、 DNN_F 的损失函数值分别为 0.2309、0.1680、0.0306、6.5213，本文模型损失函数值较 DNN_F 分别下降 96%，97%，99%。由此可知，本文 DNN 模型在学习效果上有很大提升，即数据处理部分有效降低了数据复杂度，削弱了输入输出之间函数的复杂性，进而降低了神经网络学习负担，提高了学习效果。

附录 C

表 C1 发电机出力和支路有功功率的求解结果
Table C1 Solution results of generator output and branch active power

算例	方案	P_{g1}/MW	P_{g2}/MW	Q_{g1}/Mvar	Q_{g2}/Mvar	P_{ij1}/MW	P_{ij2}/MW
IEEE 30	R ₂	32.8342	11.8634	24.2306	31.4526	6.9528	17.5038
	R ₃	32.6877	11.1472	24.9812	30.1762	6.8181	13.8934
	R ₄	31.3364	11.0359	26.9386	28.9408	5.2025	11.2954
	R ₅	33.4093	11.7930	23.8146	31.0528	7.2202	16.8168
	R ₆	32.3995	11.4103	21.8870	31.6204	7.2836	19.3709
	R₇	33.4732	11.7810	24.4836	30.9952	7.5784	17.0113
	R₉	33.4653	11.7864	24.4922	30.9896	7.5724	17.0190
	R ₂	29.5245	5.7796	3.9480	1.4856	9.5994	4.1757
	R ₃	24.9776	5.1665	3.5314	1.6383	5.7226	3.6182
IEEE 118	R ₄	28.8070	4.7175	3.9654	1.2708	5.1632	3.2650
	R ₅	26.8743	5.5588	3.6443	1.4848	10.019	4.4389
	R ₆	27.7478	5.9258	4.4243	1.3094	8.3710	4.0285
	R₇	27.2460	5.6357	3.8588	1.5839	8.0967	5.1191
	R₉	27.2577	5.6384	3.8928	1.5853	8.1025	5.1232
	R ₂	15.5446	34.1405	29.3764	37.9265	72.5349	11.3067
	R ₃	19.7193	25.6968	26.9265	29.1593	36.4489	27.6889
IEEE 2383	R ₄	12.4504	27.0366	19.3983	23.6406	46.1711	6.4324
	R ₅	17.6284	30.9330	28.8059	25.3487	52.9242	14.0678
	R ₆	17.6992	22.1726	12.1953	14.8795	78.5854	17.2565
	R₇	17.3509	30.2945	28.5659	34.8096	60.8574	16.1773
	R₉	17.3696	30.3268	28.5941	34.8470	60.9886	16.2122

表 C2 支路无功功率和电压幅值、电压相角的求解结果

Table C2 Solution results of branch reactive power and voltage amplitude and voltage angle

算例	方案	Q_{ij1}/Mvar	Q_{ij2}/Mvar	$V_1/\text{p.u.}$	$V_2/\text{p.u.}$	$\theta_1/(\text{ }^\circ)$	$\theta_2/(\text{ }^\circ)$
IEEE 30	R ₂	2.7724	5.7616	1.0427	1.0217	4.2007	6.2225
	R ₃	2.9664	5.4803	1.0380	1.0719	4.1754	6.1944
	R ₄	1.8308	4.8559	1.0376	1.0307	4.1742	6.1779
	R ₅	2.6268	8.4361	1.0399	1.0214	4.1956	6.2951
	R ₆	2.2481	5.9627	1.0386	1.0216	4.1971	6.2117
	R₇	2.5909	6.8700	1.0389	1.0219	4.1976	6.2124
	R₉	2.5921	6.8741	1.0394	1.0220	4.1982	6.2134
IEEE 118	R ₂	5.4175	8.0649	1.03764	1.0516	18.6443	31.6768
	R ₃	5.1657	7.6934	1.0373	1.0704	18.5778	30.9776
	R ₄	4.6845	6.9731	1.0369	1.0524	18.9371	30.6220
	R ₅	5.3447	7.9555	1.0356	1.0564	18.8707	31.1823
	R ₆	4.6054	6.8565	1.0362	1.0724	18.8192	31.2022
	R₇	6.5298	9.7207	1.0389	1.0509	18.7759	31.2948
	R₉	6.5363	9.7302	1.0384	1.0512	18.7779	31.2983
IEEE 2383	R ₂	10.2242	6.9886	1.0658	1.0496	3.0550	2.7692
	R ₃	3.9346	4.5510	1.0228	1.0302	3.0385	2.7552
	R ₄	11.6700	3.0606	0.9985	1.0495	3.0620	2.5404
	R ₅	5.4412	6.2367	0.9985	1.0546	3.1976	2.5013
	R ₆	3.9458	4.1793	0.9998	1.0364	2.9968	2.5200
	R₇	8.7791	9.6051	1.0109	1.0479	3.0933	2.6016
	R₉	8.4132	9.6444	1.0118	1.0480	3.0945	2.6027

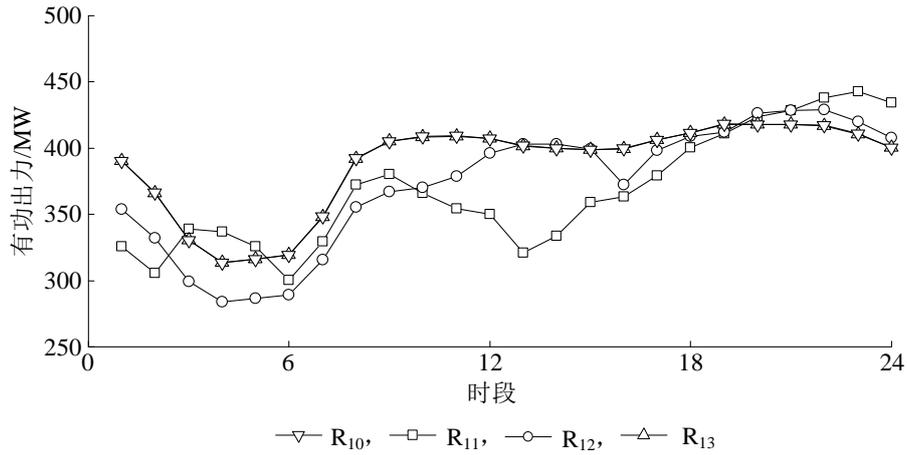


图 C1 IEEE 118 算例在 R_{10} — R_{13} 下的结果对比
 Fig.C1 Comparison of solution results of IEEE 118 examples under R_{10} to R_{13}

表 C3 各算例在方案 R_2 — R_8 下的指标对比结果
 Table C3 Comparison results of each calculation example under scheme R_2 to R_8

算例	方案	$P_{g\ rel}/\%$	$Q_{g\ rel}/\%$	$P_{ij\ rel}/\%$	$Q_{ij\ rel}/\%$	$V_{rel}/\%$	$\theta_{rel}/\%$	$E_{rel}/\%$
IEEE 30	R_2	1.8836	1.4751	14.7806	16.1847	0.0229	0.1729	0.5858
	R_3	2.0236	2.3659	18.3624	20.2659	0.0893	0.2986	0.9769
	R_4	6.3659	6.3659	33.6259	29.3648	0.1259	0.56478	1.8543
	R_5	0.1659	2.7658	4.8803	6.3659	0.0386	0.0563	0.0802
	R_6	3.1837	10.6358	3.8098	13.2658	0.0313	0.0204	1.2536
	R_7	0.0351	0.0411	0.0451	0.0575	0.0026	0.0091	0.0089
	R_8	1.1834	2.8326	5.2364	7.3269	0.0522	0.8361	0.8054
	IEEE 118	R_2	6.3559	6.2669	18.5495	17.1127	0.0353	0.7066
R_3		8.3624	9.2547	29.3647	20.9654	0.0586	1.0236	10.3695
R_4		16.3259	23.7549	36.2659	28.3267	0.1026	2.1596	13.6647
R_5		1.4038	6.3629	13.3626	18.2369	0.2236	0.3695	0.7659
R_6		5.5369	12.0239	21.3625	29.5367	0.1706	0.3058	2.3846
R_7		0.0403	0.0858	0.0745	0.0946	0.0036	0.0102	0.0094
R_8		1.8326	7.3695	14.3629	20.3695	0.3564	0.4071	1.3766
IEEE 2383		R_2	10.5036	11.2517	30.2568	37.9026	0.4365	1.2587
	R_3	15.2548	16.3218	40.2359	53.2316	0.8561	1.7965	13.3215
	R_4	30.6317	32.1586	60.3248	68.2634	1.2358	2.3618	26.2659
	R_5	4.2635	27.2569	13.2254	35.3248	1.3269	3.8669	2.4659
	R_6	26.8856	57.3721	36.7329	56.6639	1.0983	3.1496	14.9038
	R_7	0.1038	0.1069	0.2137	0.4026	0.0047	0.0204	0.0708
	R_8	8.6396	36.3619	18.3269	40.9658	1.4039	4.3398	4.3967