

负荷估计方法

摘要: 关键词:

Abstract:

Key words:

0 引言

需求响应(Demand Response, DR)利用价格和

基金项目: 激励信号实现电动汽车、分布式储能、弹性负荷等需求侧资源的海量聚合和快速响应^[1-3], 在提升系统供需平衡能力、削峰填谷、促进新能源消纳等方面发挥着重要作用^[4-5]。DR 分为价格型^[6]与激励型^[7]两种。与价格型 DR 相比, 激励型 DR 具有更快的响应速度和更强的容量扩展潜力, 在提升电力系统安全稳定性以及促进新能源消纳等方面有着更加广泛的应用^[8-9]。参与补偿金是激励型 DR 实施者向参与者支付的补偿, 等于负荷响应量与其补偿单价的乘积。响应量即“用户不参与 DR 时本应消耗的负荷与参与 DR 后实际消耗的负荷两者之差的绝对值”, 其中后者是实际测量数据, 前者就是用户基线负荷, 即用户不参与 DR 时本应消耗的负荷^[10], 如图 1 所示。图中红色虚线为用户执行 DR 后的负荷曲线, 蓝色实线为用户未执行 DR 的负荷曲线, 也即用户基线负荷。红色虚线与蓝色实线之间的面积为负荷响应量。用户一旦参与了激励型 DR, 其基线负荷真实值无法通过表计测量得到, 因此必须对其进行估计。

准确的基线负荷估计对于激励型 DR 的实施十分重要, 因为估计值过高或者过低都会影响 DR 效果的认定, 进而影响 DR 实施者和参与者双方的利益^[11]。

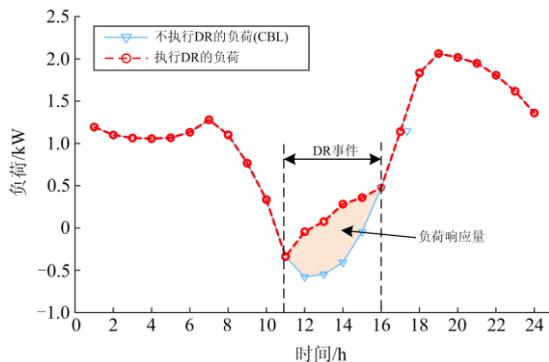


图 1 光伏用户的 CBL 示意图

Fig. 1 CBL schematic diagram of photovoltaic users

按照空间层级划分, 基线负荷可以分为 2 类: 个体和集群。个体基线负荷(Customer baseline load, CBL)指的是单个参与 DR 的用户基线负荷, 其估计结果主要服务于 DR 参与者补偿金的事后结算; 集群基线负荷(Aggregated baseline load, ABL)指的是负荷聚合商集成代理用户的基线负荷之和, 其估计结果主要服务于负荷聚合商和上级 DR 资源购买者(如系统运营商)之间的补偿结算^[12]。

对于 CBL 估计, 目前国内外主要有三类估计方法: 平均法、回归法和对照组法。平均法通过计算用户历史负荷数据的平均值来估计基线负荷。按照历史数据选取的不同, 平均法可以进一步分为 HighXofY^[13]、LowXofY^[14]、MidXofY^[15]等。国外多个独立系统运营商采用平均法来估计基线负荷。回归法通过拟合影响因素(历史负荷、气温、湿度等)与负荷间的关系来估计基线负荷。文献[16]利用分段线性回归法拟合温度和负荷之间的线性关系, 实现了商业建筑 CBL 的准确估计。对照组法将全部用户按是否参与 DR 分为: 参与的 DR 组和不参与的对照组, 其基本思想是利用与被估计对象具有相似负荷模式的对照组用户的负荷数据来估计 DR 用户的基线负荷^[17]。对于 ABL 估计, 尚无针对性研究^[18], 目前都采用个体用户估计值直接累加的方式得到。

近年来, 分布式光伏产业发展迅猛, 越来越多的居民用户开始安装分布式光伏发电系统。大多数小型分布式光伏系统(< 10 kWp)安装在表后(Behind-the-meter)^[19], 其没有单独表计量, 现有表计测得的是净负荷数据(即实际负荷减去光伏出力)。近年来, 表后分布式光伏辨识与出力估计问题越来越得到国际学术届的关注, 相关研究逐渐增多^[20-22]。在现实中, 即使系统运营商或者负荷聚合商可以在与用户合作之前要求用户提供表后光伏相关信息, 但是以下两种情况的存在使得用户提供的信息可能不准确: 1) 用户故意隐瞒。用户为了骗取补贴, 可能存在违法安装^[23]、私自扩容^[24]行为, 因此不愿透露安装光伏这一“隐私”信息, 这导致聚合商获得的信息与实际不符; 2) 用户无意误

报。大多数分布式光伏用户由于不具备专业知识, 因此对于安装的分布式光伏系统缺乏掌握, 如果光伏系统由于种种原因故障导致退出运行, 用户可能全然不知, 这也可能导致用户提供的信息与实际不符。

表后光伏的不可观性加上光伏本身具有的随机间歇特性大大增加了 ABL 估计的难度, 因为具有不确定性的光伏出力和同样具有不确定性的用户负荷相互耦合, 不确定性更强^[25-27]。针对该问题, 文献[28]提出了一种“光伏-负荷”解耦的 CBL 估计方法。该方法将每个用户的分布式光伏出力和用户负荷从净负荷中解耦, 然后分别进行估计, 最后再将两者合并得到 CBL 估计值。但该方法适用于单个用户的 CBL 估计, 在 ABL 估计方面尚无文献考虑分布式光伏的影响。事实上, 光伏出力水平在不同天气类型下存在明显差异^[29-30], 这会导致光伏用户的净负荷曲线形态在不同天气类型情况下存在较大差异, 而非光伏用户的净负荷曲线形态则较为相似, 即光伏用户净负荷相较于非光伏用户更容易受天气类型影响。为此, 应将两类用户分开考虑, 针对各自特点建立不同的估计模型。

为此, 本文提出了一种基于辨识解耦的含高渗透分布式光伏 ABL 估计方法。首先, 提出一种基于天气状态驱动特征的光伏用户辨识方法, 将用户集

群中的光伏用户与非光伏用户解耦分离; 其次, 对光伏用户集群建立基于天气类型分类的估计模型, 对非光伏用户集群建立分段线性回归估计模型; 最后将两类用户的估计值累加得到最终 ABL 的估计值, 以此提高 ABL 估计精度。

1 基于辨识解耦的 ABL 估计方法

本文提出了一种基于辨识解耦的用户集群基线负荷估计方法。该方法的步骤如下:

(1) 用户辨识解耦。基于不同天气类型下的用户净负荷特征, 识别出用户集群中的光伏用户与非光伏用户, 并将用户集群分为光伏用户集群和非光伏用户集群两个子集群。

(2) 非光伏用户 ABL 估计。非光伏用户负荷受天气类型影响较小, 只需考虑其与温度之间的关系, 建立分段线性回归模型估计非光伏用户的 ABL。

(3) 光伏用户 ABL 估计。光伏用户负荷受天气影响较大, 对光伏用户集群建立四种天气类型下的支持向量回归模型估计其基线负荷。

(4) 将步骤(2)和(3)两步的估计结果累加, 得到最终的 ABL 估计值。

本文提出的估计方法总体框架如图 2 所示。

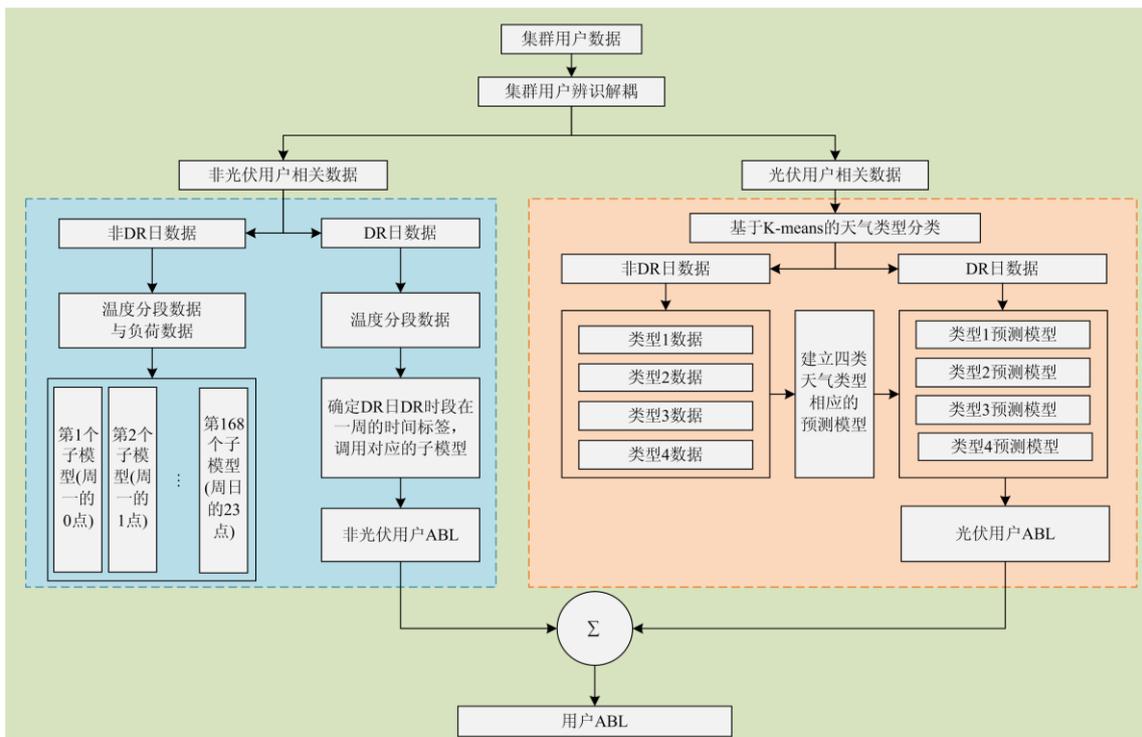


图 2 估计方法总体框架图

Fig. 2 Overall framework of the estimation method

2 集群用户辨识解耦方法

光伏出力易受天气因素影响，与天气类型紧密相关。因此，与非光伏用户相比，在不同天气类型下光伏用户净负荷区别明显，根据这一特点，本文提出了基于天气状态驱动特征的光伏用户辨识方法。该方法分为4个步骤如下述。

步骤 1: 根据可观测光伏出力，使用 K-means 聚类方法将所有日划分到几类天气类型下，在此之前需确定聚类数。晴天、多云、阵雨、大雨是四类具有典型代表意义的天气类型，这4种天气类型下的光伏出力各自具有鲜明的特点，且出现的几率较高，涵盖了绝大多数日期对应的天气状态^[30]。因此，本文利用 K-means 将所有天气类型归并成4种广义天气类型，并且得到每日的天气类型标签。

K-means 是一种保证簇内各样本间相似性的聚类算法。将样本分为 K 个簇类，并获取各簇均值向量——簇心，其目标是通过最小化簇内样本与簇心的距离，保证簇内相似性最高，簇间相似性最低，如式(1)所示。

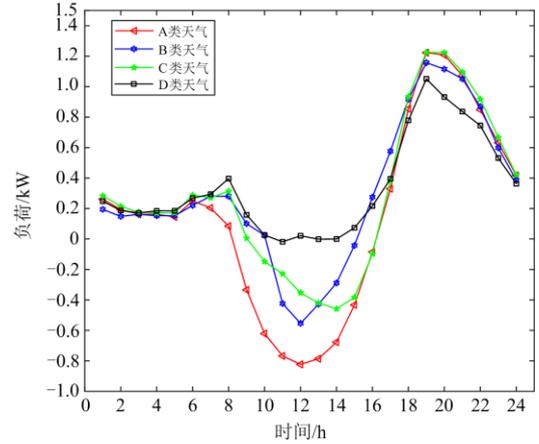
$$\min \sum_{k=1}^K \sum_{d \in D_k} \| p_{pv}^{k,d} - \bar{p}_{pv}^k \| \quad (1)$$

式中： $p_{pv}^{k,d}$ 表示第 k 类天气类型下第 d 日的光伏出力向量， $k \in \{A, B, C, D\}$ ； \bar{p}_{pv}^k 代表第 k 类天气类型下光伏出力的均值向量； D_k 表示第 k 类天气类型包含的所有天的集合。

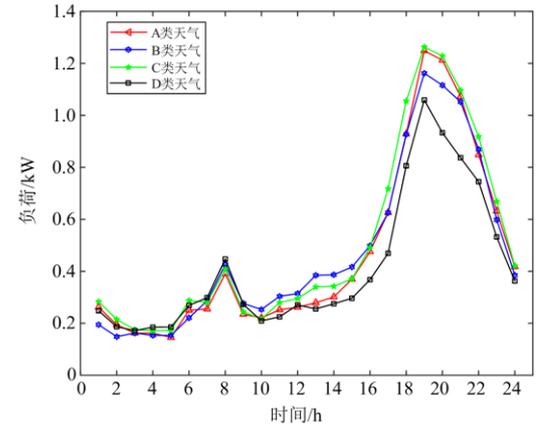
步骤 2: 为了降低居民负荷随机波动对辨识结果的不利影响，对每个用户提取不同广义天气类型下的典型净负荷曲线，如式(2)所示。

$$\tilde{p}_{net}^k(t) = \frac{1}{|D_k|} \sum_{d \in D_k} p_{net}^{k,d}(t), \quad k \in \{A, B, C, D\} \quad (2)$$

式中： $|D_k|$ 表示第 k 类天气类型包含的天数； $p_{net}^{k,d}(t)$ 表示第 k 类天气类型下第 $d \in D_k$ 天第 t 时段的净负荷功率； $\tilde{p}_{net}^k(t)$ 表示第 k 类天气类型下第 t 时段的典型净负荷功率。据此，绘制4种广义天气类型下的两类用户的净负荷曲线，如图3所示。



(a) 光伏用户在不同天气类型下的净负荷曲线



(b) 非光伏用户在不同天气类型下的净负荷曲线

图3 光伏用户和非光伏用户在不同天气类型下的净负荷曲线

Fig. 3 Net load curves of PV users and non-PV users under different weather types

步骤 3: 本文选择光伏出力不为0的一个时间窗口 $\delta = [t_s, t_e]$ ， t_s 表示起始时刻， t_e 表示结束时刻，并使用该时间窗内的典型净负荷曲线来提取能够描述光伏用户和非光伏用户净负荷形态差异的特征，具体如下：

(1) A类和D类天气类型下净负荷绝对值之比
光伏用户在不同天气类型下的典型净负荷曲线幅值存在差异，这种差异在A类和D类天气之间尤为明显。据此，本文将A类和D类天气类型下典型净负荷功率的绝对值之比作为第一个特征，记为 F_1 ，其计算公式如式(3)所示。

$$F_1 = \frac{\sum_{t=t_s}^{t=t_e} |\tilde{p}_{net}^D(t)|}{\sum_{t=t_s}^{t=t_e} |\tilde{p}_{net}^A(t)|} \quad (3)$$

式中， $|\tilde{p}_{net}^A(t)|$ 和 $|\tilde{p}_{net}^D(t)|$ 分别代表A类和D类天气类型下典型净负荷功率的绝对值。对于光伏用户来

说, 这个特征的值应该大于 1; 然而对于非光伏用户来说, 这个特征的值应该接近于 1。

(2) 凹凸特性

用户在安装光伏后, 其净负荷曲线在时间窗 δ 内将向下凹, 这一现象在 A 类天气类型下表现得更加明显, 因为与其他天气类型相比, 光伏发电量在 A 类天气类型下要大得多。据此, 基于 A 类广义天气类型下的典型净负荷曲线, 提取第二个特征, 称为“凹凸特性”, 记为 F_2 , 其计算过程如下。

一旦时间窗口确定后, 存在一条连接起始点和结束点的直线, 可用式(4)来表示。

$$y(t) = \tilde{p}_{\text{net}}^{\text{A}}(t_s) + (t - t_s) \cdot \frac{\tilde{p}_{\text{net}}^{\text{A}}(t_e) - \tilde{p}_{\text{net}}^{\text{A}}(t_s)}{t_e - t_s} \quad (4)$$

式中: $y(t)$ 表示第 t 时段直线上的采样点; 所有采样点构成一个集合, 记为 S , $S = \{y(t) | t \in [t_s, t_e]\}$ 。则位于直线下方的采样点所构成的集合 \underline{S} 可用式(5)表示。

$$\underline{S} = \{\tilde{p}_{\text{net}}^{\text{A}}(t) | t \in [t_s, t_e], \tilde{p}_{\text{net}}^{\text{A}}(t) < y(t)\} \quad (5)$$

据此, 第二个特征可由式(6)计算得到。

$$F_2 = \text{card}(\underline{S}) / \text{card}(S) \quad (6)$$

式中, $\text{card}(\cdot)$ 表示集合中元素的个数。这个特征描述了这条线下采样点的比例, 可以用来反映典型净负荷曲线在时间窗 δ 内的凹凸特性。

对所有用户, 提取上述两个特征, 每个用户的特征可构成一个向量 $F = [F_1, F_2]^T$, 其将作为步骤 4 的输入。

步骤 4: 由于用户集群中各个用户的类别(光伏用户或者非光伏用户)都是未知的, 也就是说没有已知标签的样本可供训练, 因此这是一个典型的无监督学习问题, 可采用聚类算法来进行辨识。将归一化后的特征向量作为输入, 采用 K-means 算法将所有用户聚成 2 类, 聚类中心特征值较大的那类对应光伏用户, 剩下的即为非光伏用户。

3 ABL 估计方法

3.1 非光伏用户 ABL 估计方法

影响非光伏用户负荷变化的主要因素为历史负荷和温度, 在构建估计模型时应合理考虑温度的处理方式。当温度高于一定值时, 温度越高, 用户制冷负荷越大; 当温度低于一定值时, 温度越低, 用户采暖负荷越大; 当温度处在“死区”时, 负荷几乎不随温度变化^[16]。温度与负荷之间的非线性关系可以由分段的线性关系来刻画。温度与负荷的关系如图 4 所示。

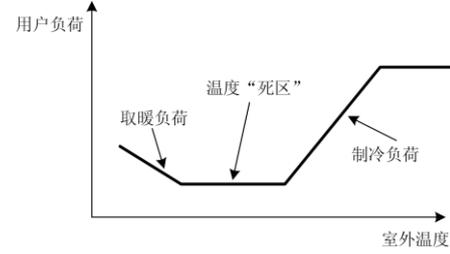


图 4 室外温度与用户负荷的关系曲线

Fig. 4 Relationship curve between outdoor temperature and user load

本文针对非光伏用户集群建立含有一周时间变量和分段温度的线性回归模型来估计 ABL。多元线性回归模型研究多个自变量和一个因变量之间的映射关系, 模型简单易懂, 相较于神经网络等机器学习模型, 具有因果清晰、易解释等优点。

本文把室外温度分成 6 段, 在进行分段之前, 先定义几个表达。 $T(t)$ 表示一周内时间 $t(t=1, 2, \dots, 168)$ 的温度; $T_j(t)$ 表示将 t 时间的温度分段后的第 j 段的温度, $j=1, 2, \dots, 6$ 。 $B_k (k=1, 2, \dots, 5)$ 代表温度间隔的第 k 个边界, 分段温度计算方法如下:

(1) 如果 $T(t) > B_1$, 那么 $T_1(t) = B_1$, 否则 $T_1(t) = T(t)$, $T_m(t) = 0 (m=2, 3, \dots, 6)$;

(2) 如果 $T(t) > B_n (n=2, 3, 4)$, 那么 $T_n(t) = B_n - B_{n-1}$, 否则 $T_n(t) = T(t) - B_{n-1}$ 且 $T_p(t) = 0, (p=n+1, \dots, 6)$;

(3) 若 $T(t) > B_5$, 则 $T_5(t) = B_5 - B_4$ 且 $T_6(t) = T(t) - B_5$ 。

温度分段方法示例(假设温度间隔的边界从 $B_1 - B_5$ 分别为 10, 20, \dots , 50)如表 1 所示。

根据上述内容, 构建回归模型如式(7)。

$$L(t) = \alpha_t + \sum_{j=1}^6 \beta_j \cdot T_j(t) \quad (7)$$

式中: α_t 是时间 $t(t=1, 2, \dots, 168)$ 对应的回归参数; β_j 是第 j 段温度对应的回归系数; $L(t)$ 是时间 t 所对应的负荷。

根据 DR 时段及相应温度, 即可得到该时段的负荷估计值。

表 1 温度分段计算示例

Table 1 Example of temperature section calculation

$T(t)$	$T_1(t)$	$T_2(t)$	$T_3(t)$	$T_4(t)$	$T_5(t)$	$T_6(t)$
5	5	0	0	0	0	0
13	10	3	0	0	0	0
25	10	10	5	0	0	0

36	10	10	10	6	0	0
59	10	10	10	10	10	9

3.2 光伏用户 ABL 估计方法

针对光伏用户在 4 类天气下净负荷差异明显的问题, 本文考虑将光伏用户每日数据根据四种典型天气类型分类, 按 4 种天气类型分别使用 SVR 建立估计模型。

SVR 的基本思想是通过非线性映射将输入空间的数据 x 映射到高维特征空间 G 中, 并在该空间中进行线性回归。然后通过损失函数 $f(x) = [\omega^T \varphi(x)] + b$ 解决回归问题。

设有 n 对训练样本 $\{(x_i, y_i) | i = 1, 2, \dots, n\}$, 则相应的回归函数为

$$\begin{cases} \min & c \sum_{i=1}^n (\eta_i + \eta_i^*) \\ \text{s.t.} & \begin{cases} \omega^T \varphi(x_i) + b - y_i \leq \varepsilon + \eta_i \\ y_i - \omega^T \varphi(x_i) - b \leq \varepsilon + \eta_i^* \\ \eta_i, \eta_i^* \geq 0, i = 1, 2, \dots, n \end{cases} \end{cases} \quad (8)$$

式中: c 为错误惩罚因子; η_i 和 η_i^* 表示两个非负稀疏变量; ε 代表损失函数。

引入稀疏数据点对决策函数予以表征:

$$L_\varepsilon(y) = \begin{cases} 0 & |f(x) - y| < \varepsilon \\ |f(x) - y| - \varepsilon & |f(x) - y| \geq \varepsilon \end{cases} \quad (9)$$

引入 Lagrange 乘子 a_i 和 b_i , 把凸优化问题转化为最大二次型形式, 如式(10)。

$$\begin{cases} \max W(a_i, b_i) = \sum_{i=1}^N y_i (a_i - b_i) - \varepsilon \sum_{i=1}^N (a_i + a'_i) - \frac{1}{2} \sum_{i,j=1}^N (a_i - a'_i)(a_j - a'_j)(x_i - x'_j) \\ \text{s.t.} \begin{cases} \sum_{i=1}^N a_i = \sum_{i=1}^N a'_i \\ 0 \leq a_i, a'_i \leq c \end{cases} \end{cases} \quad (10)$$

式中: c 用于控制模型的复杂度和逼近误差的折中, 因此 c 越大, 代表拟合度越好; ε 用于控制回归逼近误差和模型的泛化能力。

将每种天气类型下的用户数据做如下划分, 数据集(11)、数据集(12)代表非 DR 日与 DR 日的负荷与温度数据, 其中 I 代表总的时间点数, $L_i^{\text{non-DR}}$ 与 $T_i^{\text{non-DR}}$ 分别代表非 DR 日在 i 时刻的负荷与温度, L_i^{DR} 与 T_i^{DR} 分别代表 DR 日在 i 时刻的负荷与温度。估计模型在数据集(11)进行训练, 在数据集(12)检验估计效果。

$$\left. \begin{matrix} L_1^{\text{non-DR}}, \dots, L_i^{\text{non-DR}}, \dots, L_I^{\text{non-DR}} \\ T_1^{\text{non-DR}}, \dots, T_i^{\text{non-DR}}, \dots, T_I^{\text{non-DR}} \end{matrix} \right\} \quad (11)$$

$$\left. \begin{matrix} L_1^{\text{DR}}, \dots, L_i^{\text{DR}}, \dots, L_I^{\text{DR}} \\ T_1^{\text{DR}}, \dots, T_i^{\text{DR}}, \dots, T_I^{\text{DR}} \end{matrix} \right\} \quad (12)$$

模型的输出为待估计点的负荷 L_i , 输入为历史负荷 $\{L_{i-11}, \dots, L_{i-2}, L_{i-1}\}$ 与温度数据 $\{T_{i-5}, \dots, T_{i-2}, T_{i-1}\}$, 如图 5 所示。分别为 4 种典型天气类型建立相应的估计模型, 再根据 DR 日天气类型对应的模型获得估计值。

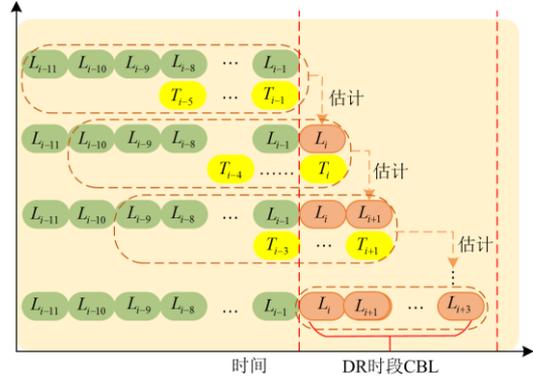


图 5 每种天气类型下估计模型的输入输出示意图

Fig. 5 Input and output diagrams of estimation models for each weather type

4 算例仿真

4.1 数据集

本文使用的数据集来自 Ausgrid, 该公司为澳大利亚悉尼的 170 万个家庭供电, 在其网站上提供了从 300 个随机选择的家庭收集的用电数据以及屋顶太阳能光伏出力数据。每个用户都有一个独立的仪表来记录分布式光伏出力数据。因此, 净负荷数据、光伏输出功率和实际负荷功率都是已知的。选择 2012 年 1 月 1 日至 12 月 31 日的全年数据, 采样间隔为 1 h。

考虑正午光伏出力造成的用户净负荷低谷, 本文聚焦于填谷需求响应, 选取了 10 个“类 DR 日”来验证基线负荷的估计精度, DR 时段为 11:00—14:00。“类 DR 日”是指与 DR 日具有相似的天气条件, 但实际没有执行 DR 的日子。对于这些“类 DR 日”, 其真实的基线负荷是已知的。利用提出的方法估计“类 DR 日”DR 时段的负荷, 而后将该值与真实的基线负荷进行比较, 以此评价基线负荷估计得准确与否。

4.2 评价指标

4.2.1 光伏用户辨识评价指标

混淆矩阵(Confusion Matrix)用来评估分类模型的精度, 它包含关于实际类和预测类的所有信息,

如式(13)所示。

$$\mathbf{M} = \begin{bmatrix} m_{11} & \cdots & m_{1n} \\ \vdots & & \vdots \\ m_{n1} & \cdots & m_{nn} \end{bmatrix} = [m_{ij}], \quad i, j = 1, \dots, n \quad (13)$$

式中: m_{ij} 是属于第 i 类却被分到第 j 类的对象数量; n 是总的对象数。

其具体评价指标有产品精度(Product's Accuracy, PA)、用户精度(User's Accuracy, UA)和整体精度(Overall Accuracy, OA), 分别如式(14)一式(16)所示。

$$PA_i = \frac{m_{ii}}{\sum_{j=1}^n m_{ij}}, \quad i = 1, \dots, n \quad (14)$$

$$UA_i = \frac{m_{ii}}{\sum_{j=1}^n m_{ji}}, \quad i = 1, \dots, n \quad (15)$$

$$OA = \frac{\sum_{i=1}^n m_{ii}}{\sum_{i=1}^n \sum_{j=1}^n m_{ij}} \quad (16)$$

产品精度 PA 指标是指对实际某类对象分类的正确率; 用户精度 UA 指标是指模型输出的某一类分类结果的正确率; 整体精度 OA 指标是指所有分类结果的正确率。

4.2.2 ABL 估计评价指标

本文对 ABL 估计的评价采用如下指标: 平均绝对百分比误差(Mean Absolute Percentage Error, MAPE)、标准平均绝对误差(Normalized Mean Absolute Error, NMAE)、标准均方根误差(Normalized Root Mean Square Error, NRMSE)。

$$MAPE = \frac{\sum_{d \in D} \sum_{t \in \alpha} |(p_{ABL}(d, t) - p_l(d, t))|}{|D| \cdot |\alpha| \cdot \sum_{d \in D} \sum_{t \in \alpha} |(p_l(d, t))|} \quad (17)$$

$$NMAE = \frac{\sum_{d \in D} \sum_{t \in \alpha} |(p_{ABL}(d, t) - p_l(d, t))|}{\sum_{d \in D} \sum_{t \in \alpha} |(p_l(d, t))|} \quad (18)$$

$$NRMSE = \sqrt{\frac{\sum_{d \in D} \sum_{t \in \alpha} (p_{ABL}(d, t) - p_l(d, t))^2}{\sum_{d \in D} \sum_{t \in \alpha} (p_l(d, t))^2}} \quad (19)$$

式中: D 为事先选定的 DR 日组成的集合; α 为事先选定的 DR 时段组成的集合; $|D|$ 和 $|\alpha|$ 分别为 DR 天数和 DR 时段小时数; $p_l(d, t)$ 和 $p_{ABL}(d, t)$ 分别为第 d 个 DR 日第 t 时段的基线负荷实际值和估计值。

4.3 仿真结果

4.3.1 辨理解耦仿真结果

对数据集中 300 个光伏用户进行可重复随机抽样得到 700 个非光伏用户的负荷数据, 并将这 1 000 个用户数据作为光伏用户辨识的仿真数据, 辨识结果的混淆矩阵如表 2 所示。

表 2 辨识结果的混淆矩阵

Table 2 Confusion matrix of identification results

实际类别	模型辨识结果		实际类别总和
	光伏用户	非光伏用户	
光伏用户	290	10	300
非光伏用户	147	553	700
模型辨识类别总和	437	563	1 000

由表 2 可以看到, 本文所提出的光伏用户辨识方法识别正确的用户总数为 843, 其中有光伏用户 290 个, 非光伏用户 553 个。错误识别的用户中, 有 10 个光伏用户可能由于安装的分布式光伏容量较小(经过数据分析发现, 当用户安装的分布式光伏年度最大出力与用户年度最大负荷之比小于 20% 时, 可认为光伏容量较小, 相应的光伏用户容易被误判为非光伏用户), 光伏出力对净负荷的影响不显著, 因而不同天气类型下的典型负荷区别不明显, 导致其被误识为非光伏用户; 147 个非光伏用户可能由于一些其他因素(用电习惯等)导致不同日期下负荷差异明显, 致使其被误识为光伏用户。

上述光伏用户辨识结果的精度指标如表 3 所示。由表 3 可以看到, 本文所提出的光伏用户辨识方法精度较高, 能够实现光伏用户与非光伏用户的准确解耦分离。

表 3 辨识精度指标

Table 3 Metrics of identification accuracy

性能指标/%	用户类型	
	光伏用户	非光伏用户
PA	96.67	79.00
UA	66.36	98.22
OA	84.30	

4.3.2 ABL 估计结果

针对辨理解耦出的非光伏用户集群, 采用分段线性回归的方法进行 ABL 估计; 针对光伏用户集群, 利用温度和负荷数据建立 4 种不同天气类型下的估计模型。将两个子集群的估计结果相加, 得到最终的 ABL 估计值。

为了证明所提方法的有效性, 本文将该方法与其他六种方法进行了对比, 其中方法一~方法三为直接估计法(即不进行解耦), 方法四~方法六为其他分类估计方法。

方法一: 使用 HighXofY 直接进行集群基线负荷估计。

方法二：使用 MidXofY 直接进行集群基线负荷估计。

方法三：使用分段线性回归直接进行集群基线负荷估计。

方法四：对所有日使用本文的天气类型分类方法进行分类，然后为每种天气类型建立 SVR 估计模型，进行集群用户基线负荷估计。

方法五：对所有日使用本文的天气类型分类方法进行分类，然后为每种天气类型建立 LSTM 估计模型进行集群用户基线负荷估计。LSTM 的参数设置如表 4 所示。

表 4 LSTM 参数设置

Table 4 Configuration of LSTM parameters

参数	取值
输入层神经元数	16
输出层神经元数	1
隐藏层数	1
隐藏层神经元数	250
训练轮数	250
学习率	0.005
优化求解	Adam

方法六：利用 K-means 根据非 DR 时段负荷数据将日负荷分为 4 类，每类分别建立 SVR 估计模型进行集群基线负荷估计。

方法七：本文所提方法。

将七种方法得到的估计结果进行比较，评价指标如表 5 所示。

表 5 ABL 估计结果对比

Table 5 Comparison of ABL estimation results

方法	MAPE	NMAE	NRMSE
方法一	7.299 7	3.027 5	2.729 5
方法二	4.272 6	1.970 2	1.917 2
方法三	4.927 6	1.920 8	1.811 0
方法四	4.371 4	1.531 8	1.842 8
方法五	3.732 8	1.466 8	1.746 3
方法六	3.696 1	1.659 4	2.039 0
方法七	2.732 3	1.439 9	1.532 2

从表 5 中可以看出，本文所提方法估计结果的三种评价指标值均为最小，其估计性能优于三种直接估计方法(HighXofY、MidXofY 和回归法)以及三种分类估计方法。这是因为将具有显著特性差异的两类用户混合一起估计会增加 ABL 的估计难度，对用户类型不加区分且不考虑各自的负荷特性，仅对集群建立单一估计模型所得的估计结果并不理想。而本文将两类用户辨识分离并根据各自的负荷特性建立估计模型可有效提高估计的准确率。此外，与直接使用 K-means 根据负荷数据进行分类的方法相

比，本文的天气类型分类方法更能反映不同天气类型下的光伏出力特征，因而基于天气类型分类的估计方法所得结果的 NMAE 与 NRMSE 值更低，效果更好。

将某个 DR 日中的七种 ABL 估计结果与真实值进行对比，如图 6 所示。

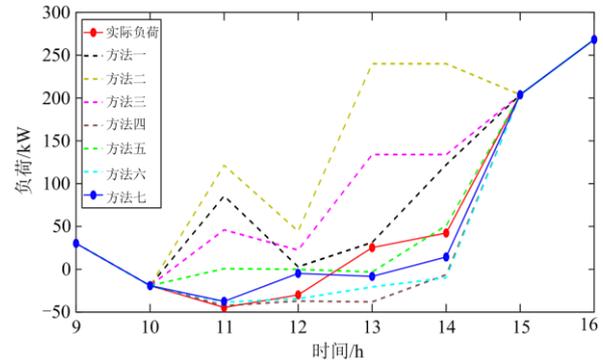


图 6 用户集群基线负荷估计结果

Fig. 6 Aggregated baseline load estimation results

图 6 中红色实线表示实际基线负荷，蓝色实线表示本文提出的基于辨识解耦的基线负荷估计方法的估计结果。其余六条虚线分别代表直接估计方法和其他分类估计方法的估计结果。从图中可以看出，本文所提方法的估计结果与实际值更加接近。

4.4 用户辨识精度的影响

为探究用户辨识精度对最终 ABL 估计精度的影响，本文设置不同光伏用户辨识精度(1-被误判为非光伏用户的光伏用户数/光伏用户总数)下的 ABL 估计场景，这里选取了 90%、80%、70%、60%、50% 五种辨识精度。具体步骤如下：

(1) 考虑到分布式光伏容量较小的光伏用户更易被误判为非光伏用户，本文从 300 个光伏用户中选择分布式光伏容量较小的 180 个用户。

(2) 从 180 个用户中按照随机抽样的方式，形成不同辨识精度下的数据。

(3) 对每种辨识精度下的数据，用本文方法进行 ABL 估计，计算相应的 MAPE、NRMSE、NMAE 值。

为减小随机因素的影响，确保结果的可靠性，对上述步骤重复 100 次，得到结果如图 7 所示。由图可知，随着辨识精度的降低，ABL 估计方法对应的三项指标均呈现递增趋势，其中，MAPE 值与 NMAE 值最为明显。根据上述实验现象，我们可以得出结论：随着用户辨识精度的降低，最终 ABL 估计精度呈下降趋势。

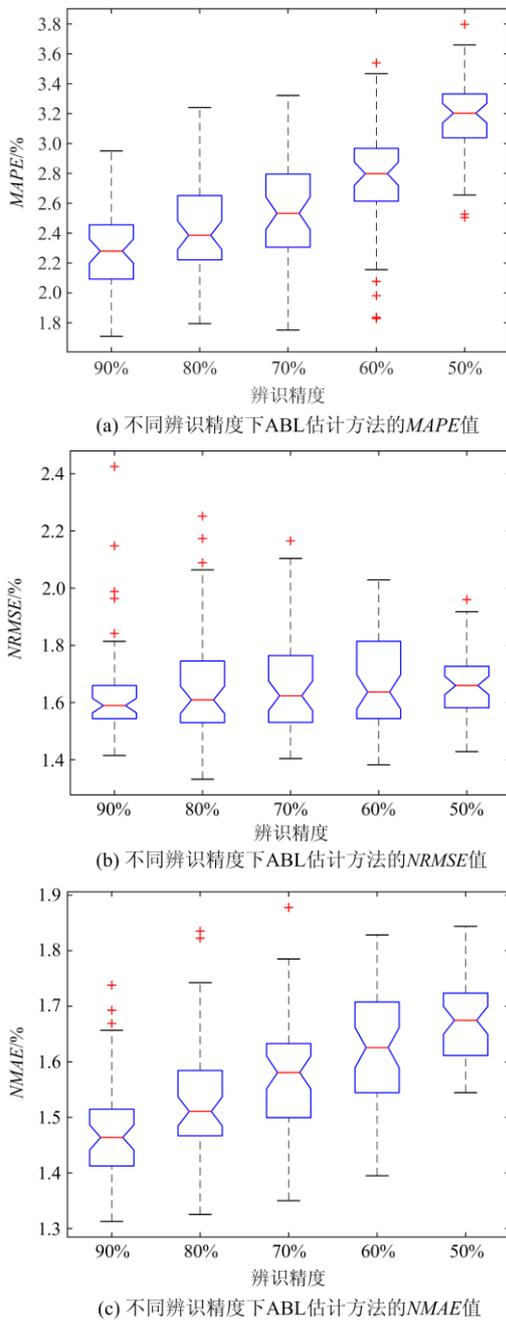


图7 不同辨识精度下ABL估计的相应指标

Fig. 7 Corresponding metrics of ABL estimation under different identification accuracy

5 结论

光伏用户和非光伏用户的负荷特性存在显著差异, 将两者混合在一起进行基线负荷估计会导致较大的估计误差。据此, 本文提出了一种基于辨识解耦的含高渗透分布式光伏 ABL 估计方法。将光伏用户与非光伏用户解耦分离, 分别考虑二者负荷特性

并建立相应的估计模型。在实际数据集上的仿真结果表明:

(1) 本文所提的光伏用户辨识方法可准确区分光伏用户与非光伏用户。

(2) 基于辨识解耦的 ABL 估计方法与直接估计法和其他分类估计方法相比, 估计精度更高。

参考文献

- [1] 何黎君, 程杉, 陈梓铭. 考虑交互功率控制和双边竞价交易的多微电网双层优化调度[J]. 电力系统保护与控制, 2020, 48(11): 10-17.
HE Lijun, CHENG Shan, CHEN Ziming. A scheduling model of a multi-microgrid system based on bi-layer optimization with consideration of PCC power control and bilateral bidding[J]. Power System Protection and Control, 2020, 48(11): 10-17.
- [2] MING Hao, XIA Bainian, LEE K Y, et al. Prediction and assessment of demand response potential with coupon incentives in highly renewable power systems[J]. Protection and Control of Modern Power Systems, 2020, 5(2): 124-137.
- [3] 刘坚, 熊英, 金亨美, 等. 电动汽车参与电力需求响应的成本效益分析—以上海市为例[J]. 全球能源互联网, 2021, 4(1): 86-94.
LIU Jian, XIONG Ying, JIN Hengmei, et al. Economic assessment of demand response delivered by electric vehicles in Shanghai[J]. Journal of Global Energy Interconnection, 2021, 4(1): 86-94.
- [4] 张颖, 容展鹏, 张宇雄, 等. 基于微电网的电网需求响应研究[J]. 电力系统保护与控制, 2015, 43(21): 20-26.
ZHANG Ying, RONG Zhanpeng, ZHANG Yuxiong, et al. Study of grid demand response based on micro grid[J]. Power System Protection and Control, 2015, 43(21): 20-26.
- [5] WANG Fei, XU Hanchen, XU Ti, et al. The values of market-based demand response on improving power system reliability under extreme circumstances[J]. Applied Energy, 2017, 193: 220-231.
- [6] 王文超, 庞丹, 成龙, 等. 考虑电价型需求响应的交直流混合配电网优化调度[J]. 电网技术, 2019, 43(5): 1675-1682.
WANG Wenchao, PANG Dan, CHENG Long, et al. Optimal dispatch approach for hybrid AC/DC distribution networks considering price-based demand response[J]. Power System Technology, 2019, 43(5): 1675-1682.
- [7] 许汉平, 李姚旺, 苗世洪, 等. 考虑可再生能源消纳效益的电力系统“源-荷-储”协调互动优化调度策略[J]. 电力系统保护与控制, 2017, 45(17): 18-25.
XU Hanping, LI Yaowang, MIAO Shihong, et al.

- Optimization dispatch strategy considering renewable energy consumptive benefits based on "source-load-energy" coordination in power system[J]. *Power System Protection and Control*, 2017, 45(17): 18-25.
- [8] 陈冉, 杨超, 沈冰, 等. 基于微电网的需求响应优化策略[J]. *电力系统保护与控制*, 2018, 46(11): 124-130.
CHEN Ran, YANG Chao, SHEN Bing, et al. Micro-grid based demand response optimization policy[J]. *Power System Protection and Control*, 2018, 46(11): 124-130.
- [9] 孙毅, 刘迪, 崔晓昱, 等. 面向居民用户精细化需求响应的等梯度迭代学习激励策略[J]. *电网技术*, 2019, 43(10): 3597-3605.
SUN Yi, LIU Di, CUI Xiaoyu, et al. Equal gradient iterative learning incentive strategy for accurate demand response of resident users[J]. *Power System Technology*, 2019, 43(10): 3597-3605.
- [10] 牛文娟, 王磊, 李扬. 需求响应项目中用户基线负荷计算方法及其应用[J]. *东南大学学报(自然科学版)*, 2014, 44(3): 556-560.
NIU Wenjuan, WANG Lei, LI Yang. Calculation method and application of customer baseline load in demand response project[J]. *Journal of Southeast University (Natural Science Edition)*, 2014, 44(3): 556-560.
- [11] 王均, 黄琦. 基于优惠券激励的需求响应双层优化机制[J]. *电力系统保护与控制*, 2019, 47(1): 108-114.
WANG Jun, HUANG Qi. Coupon incentives based customers voluntary demand response program via bilevel optimization mechanism[J]. *Power System Protection and Control*, 2019, 47(1): 108-114.
- [12] 吴宛璐, 韩帅, 孙乐平, 等. 负荷聚合商多类型需求侧资源激励价格制定一般模型及应用[J]. *电力建设*, 2021, 42(1): 1-9.
WU Wanlu, HAN Shuai, SUN Leping, et al. A general incentive pricing model and its application for multi-type demand-side resources of load aggregators[J]. *Electric Power Construction*, 2021, 42(1): 1-9.
- [13] COUGHLIN K, PIETTE M A, GOLDMAN C, et al. Estimating demand response load impacts: evaluation of baseline load models for non-residential buildings in California[R]. Berkeley: Lawrence Berkeley National Laboratory, 2008.
- [14] WIJAYA T K, VASIRANI M, ABERER K. When bias matters: an economic assessment of demand response baselines for residential customers[J]. *IEEE Transactions on Smart Grid*, 2014, 5(4): 1755-1763.
- [15] MOHAJERYAMI S, DOOSTAN M, ASADINEJAD A. An investigation of the relationship between accuracy of customer baseline calculation and efficiency of peak time rebate program[C] // 2016 IEEE Power and Energy Conference at Illinois (PECI), February 19-20, Urbana, IL, USA: 1-8.
- [16] MATHIEU J L, PRICE P N, KILICCOTE S, et al. Quantifying changes in building electricity use, with application to demand response[J]. *IEEE Transactions on Smart Grid*, 2011, 2(3): 507-518.
- [17] WANG Fei, LI Kangping, LIU Chun, et al. Synchronous pattern matching principle-based residential demand response baseline estimation: mechanism analysis and approach description[J]. *IEEE Transactions on Smart Grid*, 2018, 9(6): 6972-6985.
- [18] 孙毅, 刘昌利, 刘迪, 等. 计及滚动评价的居民群需求响应策略[J]. *中国电力*, 2019, 52(2): 34-45.
SUN Yi, LIU Changli, LIU Di, et al. Residents demand response strategy considering rolling evaluation[J]. *Electric Power*, 2019, 52(2): 34-45.
- [19] LI Kangping, WANG Fei, MI Zengqiang, et al. Capacity and output power estimation approach of individual behind-the-meter distributed photovoltaic system for demand response baseline estimation[J]. *Applied Energy*, 2019, 253.
- [20] ZHANG Xiaochen, GRIJALVA S. A data-driven approach for detection and estimation of residential PV installations[J]. *IEEE Transactions on Smart Grid*, 2016, 7(5): 2477-2485.
- [21] BU F, DEGHANPOUR K, YUAN Yuxuan, et al. A data-driven game-theoretic approach for behind-the-meter PV generation disaggregation[J]. *IEEE Transactions on Smart Grid*, 2020, 35(4): 3133-3144.
- [22] KABIR F, YU Nanpeng, YAO Weixin, et al. Joint estimation of behind-the-meter solar generation in a community[J]. *IEEE Transactions on Sustainable Energy*, 2021, 12(1): 682-694.
- [23] Hawai'i Free Press. HECO prepares to disconnect unapproved solar systems[EB/OL]. [2021-05-07]. <http://www.hawaiifreepress.com/Articles-Main/ID/13463/HECO-Prepares-to-Disconnect-Unapproved-Solar-Systems>.
- [24] 光伏盒子. 私自增容被罚款 19 万! 扩容光伏电站多赚钱, 正规渠道看这里![EB/OL]. [2021-05-07]. <https://mp.weixin.qq.com/s/3vNlUoK9rvApFBsT9uoEiA>.
- [25] WANG Chong, LEI Shunbo, JU Ping, et al. MDP-based distribution network reconfiguration with renewable distributed generation: approximate dynamic programming approach[J]. *IEEE Transactions on Smart Grid*, 2020, 11(4): 3620-3631.
- [26] 左远龙, 黄玉水, 杨晓辉, 等. 基于 PFA-MBAS-BP 神经网络模型的光伏发电短期预测[J]. *电力系统保护与控制*, 2020, 48(15): 84-91.
ZUO Yuanlong, HUANG Yushui, YANG Xiaohui, et al. Short-term prediction of photovoltaic power generation

- based on a PFA-MBAS-BP neural network model[J]. *Power System Protection and Control*, 2020, 48(15): 84-91.
- [27] 方保民, 李红志, 孔祥鹏, 等. 含高比例光伏出力的长期分布式储能配置研究[J]. *电力系统保护与控制*, 2021, 49(2): 121-129.
- FANG Baomin, LI Hongzhi, KONG Xiangpeng, et al. Research on long-term distributed energy storage configuration with a high proportion of photovoltaic output[J]. *Power System Protection and Control*, 2021, 49(2): 121-129.
- [28] XUAN Zhiming, GAO Xue, LI Kangping, et al. PV-load decoupling based demand response baseline load estimation approach for residential customer with distributed PV system[J]. *IEEE Transactions on Industry Applications*, 2020, 56(6): 6128-6137.
- [29] 王晶, 黄越辉, 李驰, 等. 考虑空间相关性和天气类型划分的多光伏电站时间序列建模方法[J]. *电网技术*, 2020, 44(4): 1376-1384.
- WANG Jing, HUANG Yuehui, LI Chi, et al. Time series modeling method for multi-photovoltaic power stations considering spatial correlation and weather type classification[J]. *Power System Technology*, 2020, 44(4): 1376-1384.
- [30] 王飞, 米增强, 甄钊, 等. 基于天气状态模式识别的光伏电站发电功率分类预测方法[J]. *中国电机工程学报*, 2013, 33(34): 75-82.
- WANG Fei, MI Zengqiang, ZHEN Zhao, et al. A classified forecasting approach of power generation for photovoltaic plants based on weather condition pattern recognition[J]. *Proceedings of the CSEE*, 2013, 33(34): 75-82.

收稿日期: 2021-05-08; 修回日期: 2021-08-27

作者简介:

(编辑 魏小丽)