

基于分类波动性统计的短期负荷改进区间预测方法

吴 茵¹, 莫 东¹, 李秋文¹, 张德亮², 黄红伟², 毛文照²

(1. 广西电网电力调度控制中心, 广西 南宁 530000;
2. 北京清大科越股份有限公司, 北京 100084)

摘要:为支撑电力现货市场下实时交易与安全校核的需要,提出了一种基于分类波动性统计的短期负荷区间预测方法。首先,介绍了传统的负荷波动性统计分析及区间预测限值计算方法,通过将负荷历史数据标准化处理,绘制负荷波动性分布直方图,计算在给定精度下的区间限值;接着结合我国实际,从居民、商业、工业三类用户出发,分别讨论不同类型用户负荷的波动性特点,在此基础上汇总形成全网负荷区间预测的上、下限值,实现对全网负荷的区间预测。最后基于某省实际数据构造的算例表明,通过深入分析不同类型负荷的波动性,本文所提出的方法能实现预测准确性和区间宽度的整体最优,在保证相同的预测精度的前提下,减少区间宽度,提高负荷区间预测结果的实用性。

关键词:电力现货市场;短期负荷预测;区间预测;波动性统计;预测准确率

文章编号:2096-4633(2019)11-0014-07 **中图分类号:**TM714 **文献标志码:**B

短期负荷预测是电力系统调度运行、电力现货市场交易组织的重要基础数据。准确的短期负荷预测结果能够为电力系统运行优化提供可靠的数据支撑,对提升系统运行效益具有显著作用。

按照预测结果的呈现方式,短期负荷预测可分为确定型预测、区间型预测和概率型预测^[1-2]。确定型预测所提供的负荷预测结果在任意时刻均为一个具体数值,由于其预测结果简单直观,传统的调度运行模式下往往采用该模式^[3-4]。概率型预测所提供的是一种负荷取不同数值的分布概率,尽管该预测模式具有更科学和完备的理论基础^[5-6],然而由于概率型预测使用较为复杂,要求调度计划编制、系统出清等均相应改为概率化模式,因此实际使用中一般不采用^[7]。区间型预测与上述两种模式不同,给出的同一时刻负荷最可能出现的区间,在预测准确性上区间型预测往往高于确定型预测,而计算分析复杂性则明显低于概率型预测。随着电力市场改革的不断推进,市场成员对负荷预测准确率的要求不断提升,确定型预测方法已难以满足市场交易分析测算的实际要求。综合考虑使用便捷性和预测准确性,区间型预测成

为当前电力市场不断发展下负荷预测研究的重要方向^[8]。

从实现方法来看,区间负荷预测的实现方式一般为优先确定区间预测的中心位置,在此基础上设定一定的区间宽度,从而获得区间负荷预测结果^[9-12]。其中合理设定区间宽度是区间负荷预测的难点。区间范围越窄,对实际调度计划优化编制等指导性越强;而区间范围越宽,往往意味着实际负荷在区间范围内的概率越高。因此区间宽度和预测精度之间存在一定的耦合关系。为解决合理设定区间宽度的问题,文献[9-10]提出了基于高斯回归模型的区间负荷预测方法,通过高斯回归,确定区间范围最有效的分布范围。文献[11-12]引入博弈论方法,提出了一种最优区间范围的计算方法。文献[13-14]提出了基于人工智能算法的区间范围设定方法,通过动态调整区间范围,给定在满足调度优化准确率要求的最窄区间范围。

实际上不同类型负荷的波动性不同,采用分类统计的方法能更加准确的辨识不同时刻不同类型负荷的波动性。按照上述研究思路,本文介绍了基于

负荷波动性统计的短期负荷区间预测方法;在此基础上,按照等可信率原则,构建了多类型负荷的区间预测方法,实现了基于分类波动性统计的短期负荷区间预测方法。最后基于某省电网实际数据构造算例,验证了所提出方法的有效性。

1 基于波动性统计的区间边界计算方法

1.1 实施流程

所谓基于波动性统计的短期负荷预测方法,是指首先根据负荷历史数据,统计其在给定的预测时段间隔内的波动分布图,在此基础上考虑一定的可置信度,计算其波动区间。如图 1 所示,基于波动性统计的短期负荷预测方法的实施流程主要包括历史数据标幺化、绘制波动分布直方图和计算区间边界值三个主要步骤。



图 1 区间边界计算流程

Fig. 1 Calculation process of interval boundary

1.2 历史数据标幺化

历史数据归一化的目的在于消除负荷历史数据由于数据基准差异导致的统计偏差。规定负荷预测的时段间隔为 ΔT 。则对于该时段范围内,负荷历史数据序列其归一化公式可表示为:

$$\hat{P}_t^L = \frac{P_t^L}{\bar{P}^{L,\Delta T}} \quad (1)$$

式(1)中, P_t^L 、 \hat{P}_t^L 分别为在时段范围 ΔT 内第 t 时刻的负荷历史数据实际值和标幺值; $\bar{P}^{L,\Delta T}$ 为该时段范围内所有负荷历史数据的平均值,可表示为:

$$\bar{P}^{L,\Delta T} = \frac{\sum_{t=1}^{N_{\Delta T}} P_t^L}{N_{\Delta T}} \quad (2)$$

式(2)中, $N_{\Delta T}$ 为在时段范围 ΔT 内负荷历史数据数。需要特别说明的是一般电力系统短期负荷预测的时段间隔为 15 min,而一般采样时段为 5 s。因此通常情况下, ΔT 为取值为 15 min,而对应的时段范围历史数据数 $N_{\Delta T}$ 为 180。全天的时段范围共 96 个。

1.3 绘制波动分布直方图

如图 2 所示,波动分布直方图是反映指标数值波动性的图形,其横坐标为负荷历史数据标幺值,根据实际需要给定标幺值区间,将所有历史数据出现范围划分为 N 个区间;其纵坐标为各区间范围内统计负荷历史数据标幺值所得的出现次数。

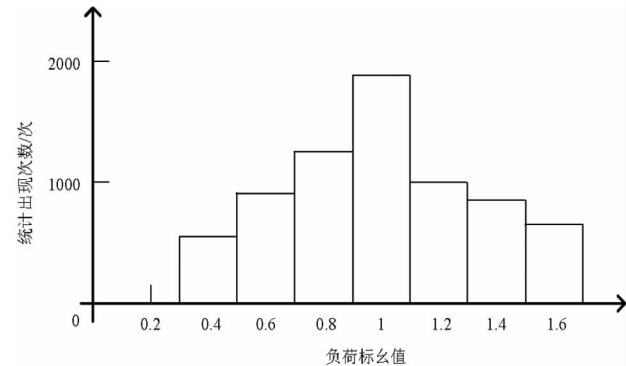


图 2 波动分布直方图示意图

Fig. 2 Straight diagram of wave distribution

由于已对待预测时段间隔内的负荷历史数据根据其在该时段范围内的平均值进行标幺化,因此负荷历史数据标幺化的波动分布直方图将以 1 为中心,呈近似正态分布的图形样式。

1.4 计算区间边界值

在负荷区间预测中,区间宽度和预测精度高度耦合。区间宽度就是预测结果中区间上边界和下边界的差;预测精度是指实际负荷出现在预测区间范围内的概率。一般来说,区间越宽,预测精度越高,然而对实际调度运行指导效果越弱;区间越窄,预测精度可能会因此下降,但是对实际调度运行的实用性增强。为了平衡区间宽度和预测精度两者之间的关系,一般可采用三种方法。

- (1) 给定区间宽度,对每个预测时刻计算其最优预测精度上下边界位置;
- (2) 给定预测精度,通过历史数据计算其最优区间范围;
- (3) 动态优化调整区间宽度和预测精度,在允

许的区间宽度和预测精度范围内, 获得最优的组合方式。

上述三种方法中, 尽管方法三效果最优, 然而由于计算过程往往需要通过采用博弈论方法, 降低了计算效率, 实际应用往往难以接受。方法一则存在可能预测精度不及预期的问题。为此本文采用方法二, 通过统计分析负荷历史分布特性, 掌握其分布规律, 来计算得到其最优区间范围。

如图 2 所示的负荷波动分布直方图中, 当历史数据量足够大, 一般要求至少统计完整一年的数据量时, 负荷历史数据标幺值波动分布直方图中各区间的频数即可用于近似代表该区间范围内波动量的出现概率。本文近似认为每个区段范围负荷波动呈平均分布, 则对于给定的预测精度要求 $\delta^{set}\%$, 其上下边界需要满足如下条件:

$$1 - B^L = B^U - 1 \quad (3a)$$

$$\frac{\int_{B^L}^{B^U} 1 dn}{N^{ALL}} \geq \delta^{set}\% \quad (3b)$$

式(3a)及(3b)中, B^U 、 B^L 分别为待求解的标幺化的区间上、下限值, N^{ALL} 为波动分布直方图中统计的负荷历史数量总量, $\int_{B^L}^{B^U} 1 dn$ 为从区间下限到上限范围内的波动直方图范围负荷历史数据标幺值出现次数。则式(3a)表明区间上、下限值应对称分布在中间值 1 两侧; 式(3b)表明区间上下限值范围内的负荷历史数据标幺值出现次数在总次数的比例应不低于给定的预测精度要求。

2 分类波动性统计预测

传统区间预测方法中, 往往将待预测的全网负荷作为一个整体, 未细致的讨论不同类型用户的负荷波动特性。由于不同类型的用户负荷波动性差异较大, 因此细致的统计不同类型用户的负荷波动性, 在此基础上将其汇总, 将有助于提升负荷区间预测准确率。

按照本文所提出的分类波动性统计的实施思路, 结合传统区间预测中边界计算方法, 本文所提出的基于分类波动性统计的短期负荷区间预测方法实施流程如图 3 所示, 按照实施内容, 可将该方法划分为负荷分类及各分类波动性统计、分类汇总及整体区间边界计算、负荷区间预测三个步骤。

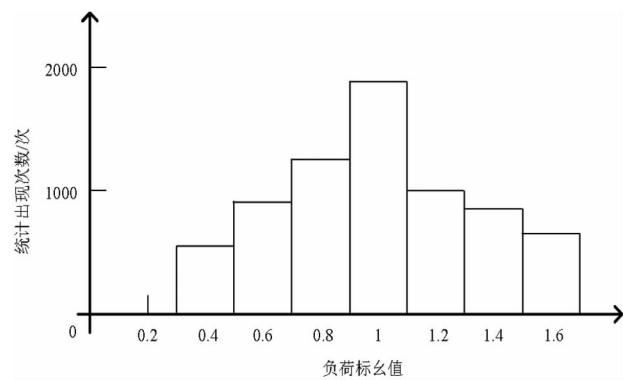


图 3 分类区间预测实施流程

Fig. 3 Implementation process of classification interval prediction

2.1 负荷分类及各分类波动性统计

负荷分类及各分类波动性统计是指将待预测的全网电力用户, 根据其类型划分为不同分类, 并按照本文第 1 节所介绍的负荷波动性统计分析方法对各分类负荷的波动性进行分析, 并计算其区间上、下限值。

结合我国负荷实际, 建议按照居民、商业、工业三大类对全网用户进行分类, 该分类方法不仅满足我国常用的电力用户分类类型, 也有助于测算售电收入。

根据负荷分类结果, 利用上一节所介绍的波动性分类及区间边界计算方法, 分别计算各分类负荷的波动直方图和区间上、下限值。规定居民、商业、工业三类负荷的区间上、下限值依次为 $B^{U,R}$ 、 $B^{L,R}$ 、 $B^{U,B}$ 、 $B^{L,B}$ 、 $B^{U,I}$ 、 $B^{L,I}$, 三类用户在波动性统计预测精度上均规定不低于给定值 $\delta^{set}\%$ 。

2.2 分类汇总及整体区间边界计算

分类汇总及整体区间边界计算是指将不同类型的负荷区间边界计算结果汇总, 考虑不同类型负荷的占比, 将其汇总计算得到全网负荷区间预测时的区间边界。

按照上述思路, 全网负荷区间预测后区间上下边界满足:

$$B^{U,A} = \alpha^R B^{U,R} + \alpha^B B^{U,B} + \alpha^I B^{U,I} \quad (4a)$$

$$B^{L,A} = \alpha^R B^{L,R} + \alpha^B B^{L,B} + \alpha^I B^{L,I} \quad (4b)$$

式(4a)及(4b)中, $B^{U,A}$ 、 $B^{L,A}$ 分别为全网负荷区间预测中的上、下限值, α^R 、 α^B 、 α^I 分别为居民、商业、工业负荷的占比, 应满足三类负荷的占比之和等于 1, 即:

$$\alpha^R + \alpha^B + \alpha^I = 1 \quad (5)$$

2.3 负荷区间预测

利用传统的确定型负荷预测方法,可以获得负荷区间预测的区间中间值。在此基础上考虑其区间边界上、下限标幺值,即可以得到负荷区间预测结果。

传统的确定型负荷预测方法较多,结合当前实际推荐采用基于神经网络的确定型负荷预测方法,常用的神经网络方法为 BP 神经网络,其优势在于能够充分考虑不同类型的相关因素,实现非线性数据拟合与分析^[15~17]。考虑到该预测方法属于较为成熟的预测方法,本文不赘述其实施细节。

规定按照时段间隔 ΔT 时刻 t 的全网负荷预测结果为 PF_t^D ,则其区间预测结果需要进一步考虑区间上下限标幺值进行还原,还原公式表示为:

$$PF_t^U = PF_t^D B^{U,A} \quad (6a)$$

$$PF_t^L = PF_t^D B^{L,A} \quad (6b)$$

式(6a)及(6b)中, PF_t^U 、 PF_t^L 分别为负荷区间预测的上、下限实际预测值。

3 算例分析

3.1 基础数据

本文将在某省电网实际负荷预测数据基础上构造算例,以验证所提出方法的有效性。从该省区电网能量管理系统中获取其 2017 至 2018 年两年的全网负荷数据,能量管理系统中按照 5 s 时间间隔存储负荷数据,因此一共可以获得 1 261.44 万条数据。根据该省区负荷预测要求,一般要求以 15 min 为时间间隔预测未来一天的负荷。

在上述基础数据基础上,将 2017 年全年及 2018 年 1 月至 8 月数据作为波动性分类统计基础数据,并用该时段数据作为神经网络训练基础数据集。2018 年 9 月至 12 月数据则用于算例测试和验证对比。

3.2 预测结果

首先定义日区间负荷预测准确率指标,规定该指标为全天预测点中实际负荷位于预测区间范围内的点数与全天预测总点数的比值,可表示为:

$$\varphi\% = \frac{N^{F,R}}{N^{F,A}} \times 100\% \quad (7)$$

式(7)中, $\varphi\%$ 为日区间负荷预测准确率, $N^{F,A}$ 、 $N^{F,R}$ 分别为全天预测总点数和实际负荷处于该区间范围的点数。当预测时段间隔为 15 min 时,则 $N^{F,A}$ 应为 96。

按照本文所介绍的负荷波动性分类统计方法,居民、商业、工业三类负荷的波动性分布直方图分别如图 4(a)、4(b)、4(c)所示。直观可以看出,三类负荷的波动性差异很大。按照波动性从大到小依次为居民负荷、商业负荷、工业负荷。工业负荷的波动分布直方图波动范围为 0.55 至 1.35,其中心区间 [0.95~1.05] 频率高达 53.6%。而对于商业负荷,其波动区间为 0.55 至 1.45,中心区间 [0.95~1.05] 频率为 41.8%。而对于居民用户,其负荷波动区间为 0.45 至 1.55,而中心区间 [0.95~1.05] 频率仅为 32.5%。上述分布特性与上述负荷的生产生活自然规律也是一致的。工业生产的连续性较好,商业次之,而居民用户的随机性较大,导致其负荷的波动性也较高。

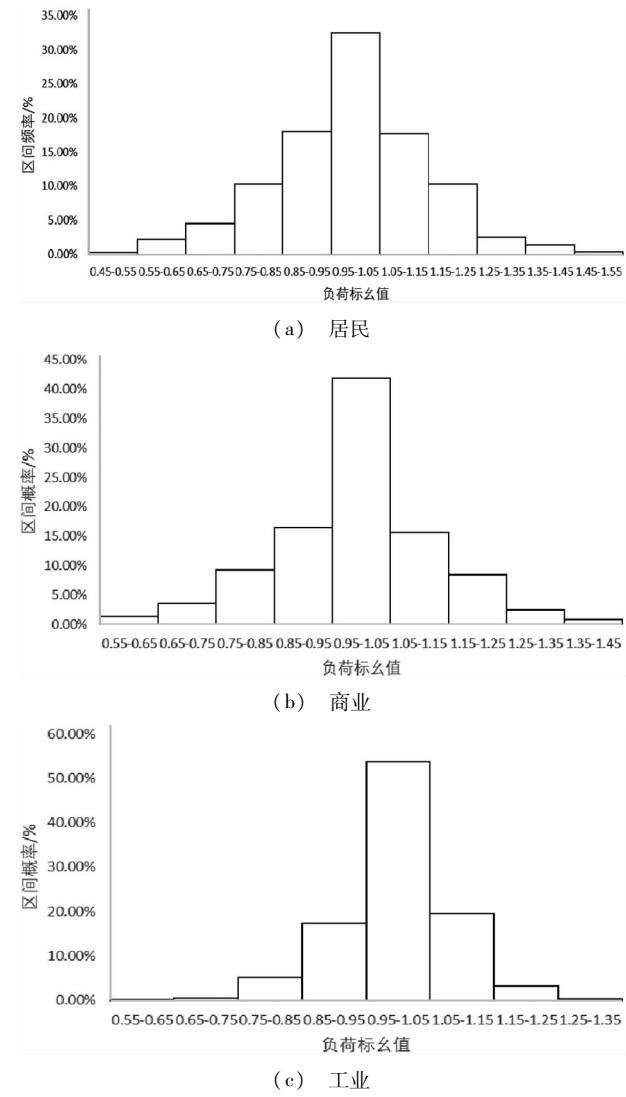


图 4 负荷波动直方图

Fig. 4 Load fluctuation histogram

按照上述波动性分布直方图,给定预测准确率要求为 95%,则居民、商业、工业三类负荷的区间上下限值如表 1 所示。居民、商业、工业三类负荷的区间宽度依次为 0.60、0.53、0.46。进一步考虑不同类型用户的负荷比例,即可以计算全网负荷的区间波动上下限值。该省区电网居民、商业和工业负荷占比约为:0.2:0.4:0.4,则已知该省区负荷全网负荷区间预测的区间上限为 1.28,下限为 0.76。可以看出尽管居民用户的负荷波动性较大,然而由于其占比较低,因此全网负荷区间趋近于工商业负荷。

表 1 分类用户区间限值

Tab. 1 Classified user interval limits

类型	区间上限	区间下限
居民	1.30	0.70
商业	1.27	0.74
工业	1.28	0.82
全网	1.28	0.76

利用神经网络训练基础数据集对神经网络预测智能体进行预测训练。将训练成熟的神经网络预测智能体用于后续测试验证集的预测,进一步考虑全网负荷区间分布,即可以实现对全网负荷的区间预测。以 2018 年 11 月 1 日为例,当前预测结果如图 5 所示。全天共有 5 个点实际负荷超过区间限值,区间预测准确率为 94.79%。

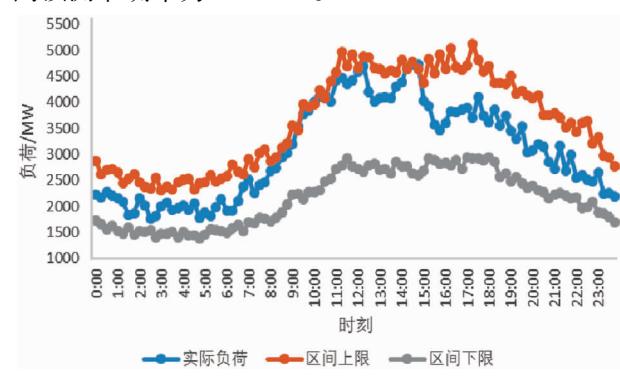


Fig. 5 单日区间预测

Fig. 5 One-day interval forecast

3.3 结果对比

为了分析本文所提出方法与传统区间预测方法的差异,将对比本文预测结果与不进行分类波动性统计而直接统计全网负荷波动性区间上下限的方式(以下简称“不分类区间预测”)预测

结果差异。

重复上面的波动性统计分析流程,全网负荷波动性分析后所得的区间上、下限值分别为 1.30、0.74。利用相同的神经网络预测后,两种预测方法的逐日预测准确率对比如图 6 所示。其中“分类”对应曲线为采用本文所提出的方法所得的逐日区间负荷预测准确率;“未分类”对应采用不分类区间预测方法所得到的逐日负荷预测准确率。在验证集对应的 122 天时间范围内,两种方法的预测准确率分别为 95.47% 和 95.60%,不分类区间预测的平均预测准确率提高仅为 0.13%;而不分类区间预测的区间宽度较本文所提出方法超出了 7%。需要特别注意的是,由于一天预测点数为 96 点,因此全天区间预测准确率取值为离散化数值,导致图中大部分天数的预测准确率数值重合。

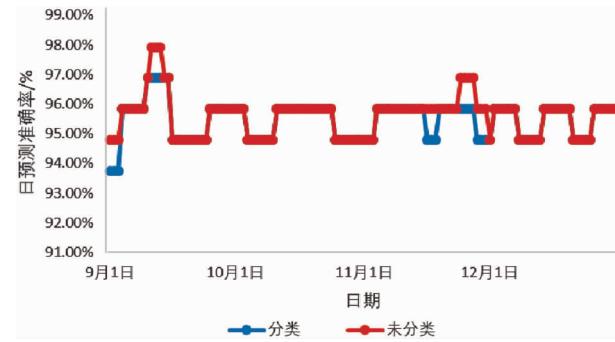


图 6 预测准确率对比

Fig. 6 Comparison of prediction accuracy

对比上面的预测准确率结果,可以发现:由于不分类区间预测的区间涵盖了本文所提出方法的区间范围,因此其预测准确率高于本文所提出方法。然而从提升效率来看,上述准确率提升不足 1%,而区间宽度增加超过 7%,表明本文所提出方法具有更高的使用效益。究其原因在于尽管不分类区间预测增加了区间宽度,但是由于未细致考虑不同类型用户负荷的波动性差异,造成所增加的区间宽度范围并不能有效反映负荷的实际波动特性,对预测准确率提升无显著作用。

4 结论

本文在传统区间负荷预测基础上,引入负荷分类波动性统计方法,通过分析居民、商业、工业三类电力用户的负荷波动性更细致的分析全网负荷波动性,在此基础上实现基于分类波动性统计的负荷区

间预测方法。

实际上负荷波动性不仅与用户类型有关,还与时段有关,考虑到不同时段不同类型用户的比例也是动态波动的,因此后续可从提升区间负荷预测适应性角度出发,采用人工智能算法,研究更加灵活的区间负荷预测方法。

参考文献:

- [1] 夏博,杨超,李冲. 电力系统短期负荷预测方法研究综述[J]. 电力大数据,2018,21(07):22–28.
XIA Bo, YANG Chao, LI Chong. Review of the short-term load forecasting methods of electric power system [J]. Power Systems and Big Data,2018,21(07):22–28.
- [2] 张凌云,肖惠仁,吴俊豪,等. 电力系统负荷预测综述[J]. 电力大数据,2018,21(01):52–56.
ZHANG Lingyun, XIAO Huiyan, WU Junhao, et al. Review of power system load forecasting [J]. Power Systems and Big Data, 2018,21(01):52–56.
- [3] 冯磊,杜佩仁,李超,等. 计及行业间同时率的分区最大负荷预测方法及应用[J]. 电力大数据,2018,21(03):7–14.
FENG Lei, DU Peiren, LI Chao, et al. Method and application of the Subarea maximum load forecasting with inter-industry simultaneous [J]. Power Systems and Big Data,2018,21(03):7–14.
- [4] 张场,汪洋,祝宇翔,等. 基于PAM和ELM的电力短期负荷预测相似日选取算法[J]. 电力大数据,2017,20(12):84–87.
ZHANG Yang, WANG Yang, ZHU Yuxiang, et al. Similarity day selection algorithm for power short-term load forecasting [J]. Power Systems and Big Data,2017,20(12):84–87.
- [5] BIDONG LIU, JAKUB NOWOTARSKI, TAO HONG, et al. Probabilistic load forecasting via quantile regression averaging on sister forecasts [J]. IEEE Transactions on Smart Grid,2017,8(02):730–737.
- [6] MEHDI RAFIEI, TAHER NIKNAM, JAMSHID AGHAEI, et al. Probabilistic load forecasting using an improved wavelet neural network trained by generalized extreme learning machine [J]. IEEE Transactions on Smart Grid,2018,9(06):6961–6971.
- [7] 黄南天,齐斌,刘座铭,等. 采用面积灰关联决策的高斯过程回归概率短期负荷预测[J]. 电力系统自动化,2018,42(23):64–75.
HUANG Nantian, QI Bin, LIU Zuoming, et al. Probabilistic short-term load forecasting using Gaussian process regression with area grey incidence decision making [J]. Automation of Electric Power Systems,2018,42(23):64–75.
- [8] 包伟涔,韦钢,刘佳,等. 基于区间模型的多能互补系统优化运行[J]. 电力系统自动化,2019,43(11):8–20.
BAO Weicen, WEI Gang, LIU Jia, et al. Interval model based optimal operation of multi-energy complementary system [J]. Automation of Electric Power Systems,2019,43(11):8–20.
- [9] 宗文婷,卫志农,孙国强,等. 基于改进高斯过程回归模型的短期负荷区间预测[J]. 电力系统及其自动化学报,2017,29(08):22–28.
ZONG Wenting, WEI Zhinong, SUN Guoqiang, et al. Short-term load interval prediction based on improved Gaussian process regression model [J]. Proceedings of the CSU – EPSA, 2017, 29 (08):22–28.
- [10] 余鹏,唐权,张文涛,等. 基于高斯模糊信息粒化和改进小波神经网络的短期负荷区间预测研究[J]. 机电工程,2017,34(02):167–172.
YU Peng, TANG Quan, ZHANG Wentao, et al. Short-term load interval forecasting based on gaussian fuzzy information granulation and improved wavelet neural network [J]. Mechanical & Electrical Engineering Magazine,2017,34(02):167–172.
- [11] YAN J, LI K, BAI E, et al. Analytical iterative multistep interval forecasts of wind generation based on TLGP [J]. IEEE Transactions on Sustainable Energy,2019,10(02):625–636.
- [12] CAN W, ZHAO X, PIERRE P, et al. Optimal prediction intervals of wind power generation [J]. IEEE Transactions on Power Systems,2014,29(03):1166–1174.
- [13] DORIS SÁEZ, FEMAND ÁVILA, DANIEL OLIVARES, et al. Fuzzy prediction interval models for forecasting renewable resources and loads in microgrids [J]. IEEE Transactions on Smart Grid,2015,6(02):548–556.
- [14] HAO QUAN, DIPTI SRINIVASAN, ABBAS KHOSRAVI. Short-term load and wind power forecasting using neural network-based prediction intervals [J]. IEEE Transactions on Neural Networks and Learning Systems,2014,25(02):303–315.
- [15] A. GARULLI, S. PAOLETTI AND A. VICINO. Models and techniques for electric load forecasting in the presence of demand response [J]. IEEE Transactions on Control Systems Technology, 2015, 23 (03): 1087–1097.
- [16] TAT KEI CHAU, SAMSON SHENGLONG YU, Tyrone Fernando, et al. A load-forecasting-based adaptive parameter optimization strategy of STATCOM using ANNs for enhancement of LFOD in power systems [J]. IEEE Transactions on Industrial Informatics, 2018, 14 (06): 2463–2472.
- [17] 樊唯钦,张伟,李勇钢,等. 基于改进人体舒适指数的微电网短期负荷预测[J]. 广东电力,2017,30(04):137–142.
FAN Weiqin, ZHANG Wei, LI Yonggang, et al. Ultra short-term load forecasting for micro-grid based on improved human comfort index [J]. Guangdong Electric Power,2017,30(04):137–142.

收稿日期:2019–09–03

作者简介:



吴茵(1979),女,硕士,高级工程师,主要从事电力系统运行与控制等方面的研究工作。

(本文责任编辑:范斌)

Improvement interval prediction methodof short-term load based on classified volatility statistics

WU Yin¹, MO Dong¹, LI Qiuwen¹, ZHANG Deliang², HUANG Hongwei², MAO Wenzhao²

(1. Guangxi Power Grid Power Dispatching Center, Nanning 530000 Guangxi, China;

2. Beijing Qingda Keyue Co., Ltd., Beijing 100084, China)

Abstract: In order to support the demand of real-time transaction and safety check in the spot electricity market, a short-term load interval prediction method based on classified volatility statistics is proposed. Firstly, the traditional statistical analysis of load volatility and the calculation method of interval prediction limit are introduced. By standardizing the historical data of load, the distribution histogram of load volatility is drawn, and the interval limit under the given precision is calculated. Then, based on the reality of China, this paper discusses the fluctuation characteristics of load of different types of users from the perspectives of residents, commercial and industrial users, and summarizes the upper and lower limits of load interval prediction of the whole network, so as to realize the interval prediction of the whole network load. Finally, a case study constructed based on the actual data of a certain province shows that through in-depth analysis of the fluctuation of different types of loads, the method proposed in this paper can achieve the overall optimal prediction accuracy and interval width, reduce the interval width and improve the practicability of load interval prediction results on the premise of ensuring the same prediction accuracy.

Key words: power spot market; short-term load forecasting; interval prediction; volatility statistics; prediction accuracy