

DOI: 10.19783/j.cnki.pspc.181207

基于 EMD-QRF 的用户负荷概率密度预测

杨斌¹, 杨世海², 曹晓冬², 陈宇沁², 梁智³, 孙国强³

(1. 国网江苏省电力有限公司, 江苏 南京 210024; 2. 国网江苏省电力有限公司电力科学研究院, 江苏 南京 211103; 3. 河海大学能源与电气学院, 江苏 南京 211100)

摘要: 考虑用户负荷时间序列基数小、波动性与随机性强、难以取得较高预测精度的特点, 建立基于经验模式分解(Empirical Mode Decomposition, EMD)与分位数回归森林(Quantile Regression Forest, QRF)的用户负荷概率密度组合预测模型, 以提高用户负荷预测精度。首先, 采用 EMD 信号处理算法对用户负荷原始时间序列数据进行分解处理, 计算各模态函数样本熵值并根据样本熵大小对模态函数进行重构。其次, 对重构分量分别建立 QRF 用户负荷预测模型, 叠加不同分量预测结果从而获得预测值的条件分布。最后, 利用核密度估计输出任意时刻用户负荷概率密度预测结果。相对于确定性点预测方法, 概率密度预测具有描述用户负荷未来可能的波动范围及不确定性等优势, 且算例测试验证了模型的有效性。

关键词: 用户负荷; 经验模式分解; 分位数回归森林; 核密度估计; 概率密度预测

Short-term consumer load probability density forecasting based on EMD-QRF

YANG Bin¹, YANG Shihai², CAO Xiaodong², CHEN Yuqin², LIANG Zhi³, SUN Guoqiang³

(1. State Grid Jiangsu Electric Power Co., Ltd., Nanjing 210024, China; 2. State Grid Jiangsu Electric Power Company Research Institute, Nanjing 211103, China; 3. College of Energy and Electrical Engineering, Hohai University, Nanjing 211100, China)

Abstract: Considering the small base of consumer load time series, strong volatility and randomness, along with the difficulty of obtaining high forecasting accuracy, a hybrid model based on Empirical Mode Decomposition (EMD) and Quantile Regression Forest (QRF) is proposed for consumer load probability density forecasting, which is aimed at enhancing prediction precision. Firstly, the signal processing algorithm of EMD is applied to decompose the original consumer load time series, where the sample entropy of each decomposed mode function is calculated. Based on the values of sample entropy, the mode functions can be reconstructed. Then, each reconstructed component is modeled separately using QRF for consumer load forecasting, where the conditional distribution of predicted values can be obtained by superimposing prediction results of different components. Finally, the Kernel Density Estimation (KDE) is used to output the consumer load probability density forecasting results at any time. Compared with deterministic point prediction methods, the proposed probability density forecasting model has advantages of describing the possible fluctuation range and uncertainty of the consumer load in the future, where the case study has also verified its validity.

This work is supported by National Natural Science Foundation of China (No. 51507052).

Key words: consumer load; empirical mode decomposition; quantile regression forest; kernel density estimation; probability density forecasting

0 引言

电力市场化改革促进了售电侧产业链迅速发展, 现有供电、大型发电与节能服务企业相继成立并运营大量性质、规模不同的售电公司。售电公司

从发电厂以较低价格批量购买电能并以较高零售价格卖给工业用户, 依靠替用户承担风险获得相应的收益。然而, 受经济政治、气候条件、生产计划等因素的影响, 工业用户实际的用电量与合同电量往往存在不同程度的偏差^[1]。

售电公司必须细致把握用户负荷特性及用电规律, 从而准确预测用户未来用电量, 规避偏差考

基金项目: 国家自然科学基金项目资助(51507052)

核损失。然而,用户负荷基数小、随机性与波动性强烈、用户用电模式多样及受多种因素共同作用的特点增加了用户负荷预测的难度,也使得用户负荷预测精度低于系统级负荷预测^[2-4]。因此,亟需研究用户负荷预测技术与方法,以应对电力交易中心偏差考核压力。

目前,用户负荷预测研究较少且主要集中在对工业负荷、空调负荷及楼宇负荷进行预测^[5-8]。工业用户相对于普通的居民和商业用户,其特点是用电量较大,电费支出在所有成本中占较大比重,如钢铁、水泥、炼铝、冶金、矿产等高能耗工矿企业。这些用户的大容量设备电压等级较高,设备的频繁启停会对电网造成不同程度的冲击与谐波影响,其负荷曲线变化剧烈,随机性与波动性也较强^[9-10]。文献[11]采用质心相似度聚类负荷模式算法挖掘历史负荷数据中合群的典型负荷模式,并按相似性进行分组,然后对每类典型负荷模式建立径向基神经网络预测模型。文献[12]在分析负荷特性的基础上,结合自回归(Auto-regressive, AR)、支持向量机(Support Vector Machine, SVM)等模型的优势对地区典型大用户建立精细化负荷预测管理系统,且考虑了日期类型、气象因素及生产计划等影响因素对负荷的作用。文献[13]利用小波去噪与决策树算法建立用户负荷组合预测模型,充分挖掘用户历史用电模式,并对不同用电模式的用户分别建立了预测模型。文献[14]利用极限学习机建立用户负荷预测模型,取得较好的预测结果。

概率密度预测能够描述用户负荷未来可能的波动范围、不确定性及面临的风险,可为供电公司购售电业务提供更多决策信息,但现有文献对用户负荷概率密度预测关注较少。因此,本文提出基于分位数回归森林(Quantile Regression Forest, QRF)与核密度估计算法的用户负荷概率密度预测模型,获得了任意时刻用户负荷概率密度预测结果。此外,为分析用户负荷时间序列的波动性与随机性,采用经验模式分解(Empirical Mode Decomposition, EMD)算法对用户负荷原始时间序列数据进行预处理,将其分解为若干特征互异的模态函数,使得原始时间序列中各特征尺度间信息解耦,弱化信息间干扰,从而降低预测模型误差。最后,通过算例说明所提EMD-QRF用户负荷概率密度预测模型的有效性。

1 EMD 实现步骤

EMD 目标是实现对原始信号的自适应分解,从而细致把握信号局部变化信息,在非线性非平稳时间序列处理中获得了广泛应用。EMD 假设任何复杂

信号皆由简单的固有模态函数(Intrinsic Mode Function, IMF)组成,且各 IMF 相互独立^[15]。

采用 EMD 技术对用户负荷原始时间序列数据进行分解处理,其步骤如下^[16]。

1) 找出用户负荷时间序列 $y(t)$ 中的所有极大值,通过插值法拟合其包络线为 $u(t)$;同理,识别 $y(t)$ 中所有极小值,通过插值法拟合其包络线为 $v(t)$ 。由此,计算上下包络线的平均值 $m_1(t)$ 为

$$m_1(t) = \frac{u(t) + v(t)}{2} \quad (1)$$

2) 将原始信号 $y(t)$ 与 $m_1(t)$ 相减得到 $h_1(t)$, 即

$$h_1(t) = y(t) - m_1(t) \quad (2)$$

判断 $h_1(t)$ 是否满足 IMF 分量条件,若不满足,则将 $h_1(t)$ 视为新的信号 $y(t)$, 重复步骤 1), 直到 $h_1(t)$ 满足 IMF 条件。若满足,则第 1 个 IMF 分量可表示为

$$c_1(t) = h_1(t) \quad (3)$$

3) 从原始信号 $y(t)$ 中分离出分量 $c_1(t)$, 相应的剩余分量可表示为

$$r_1(t) = y(t) - c_1(t) \quad (4)$$

4) 将剩余分量 $r_1(t)$ 作为新的原始信号,重复步骤 1)至步骤 3), 从而可得到其他 IMF 分量和 1 个余量。

5) 采用 EMD 方法将原始信号 $y(t)$ 分解为 n 个 IMF 分量 $c_i(t)$ 和一个剩余分量 $r_n(t)$ 。

因此原始信号 $y(t)$ 可由 IMF 分量和剩余分量表示为

$$y(t) = \sum_{i=1}^n c_i(t) + r_n(t) \quad (5)$$

2 基于 EMD-QRF 的概率密度预测模型

本文在采用 EMD 方法对用户负荷原始时间序列数据进行分解处理的基础上,建立 QRF 用户负荷短期预测模型,获得了不同分位点预测结果,进而采用核密度估计实现用户负荷的概率密度预测。

2.1 分位数回归森林基本原理

QRF 由分位数回归和随机森林发展得到,但保留了两种算法的优点,其输出值为不同分位点的回归预测结果。QRF 运算速度快,模型性能受参数影响小,且具有较强容噪性^[17-18]。因此, QRF 适用于对大量用户负荷进行短期预测。

QRF 可看作是适应性近邻分类与回归过程,对所有 $\mathbf{X} = \mathbf{x}$, 可以得到原始 n 个观察值一个权重集合 $w_i(\mathbf{x})$, $i = 1, 2, \dots, n$ 。QRF 将所有因变量观测值

的加权和作为因变量 Y 条件均值 $E(Y|X=x)$ 的估计。QRF 决策树是由标准随机森林算法产生, 条件分布是通过观测到的因变量加权估计得到, 其中每个观测值的权重等于随机森林算法权重。

QRF 定义 $E(1_{\{Y \leq y\}} | X=x)$ 的估计为观测值 $1_{\{Y \leq y\}}$ 的加权平均, 即

$$\hat{F}(y|X=x) = \sum_{i=1}^n w_i(x) 1_{\{Y \leq y\}} \quad (6)$$

QRF 算法具体步骤^[19]:

1) 生成 k 棵决策树 $T(\theta_t)$, $t=1, 2, \dots, k$ 。对每棵决策树每个叶节点, 考察该叶节点所有观测值。

2) 给定 $X=x$, 遍历所有决策树。计算每棵决策树观测值的权重 $w_i(x, \theta_t)$, $i \in \{1, 2, \dots, n\}$ 。通过对决策树权重 $w_i(x, \theta_t)$, $t=1, 2, \dots, k$ 取平均得到每个观测值 $i \in \{1, 2, \dots, n\}$ 的权重 $w_i(x)$ 。

3) 对所有 $y \in R$, 利用步骤 2) 得出的权重, 通过式(6)计算分布函数的估计。

2.2 核密度估计算法

核密度估计通过同一未知分布函数的随机变量来估计其密度函数。设 X_1, X_2, \dots, X_M 是取

自一元连续总体的样本, 在任意点 x 处的总体密度函数 $f(x)$ 的核密度估计定义为^[20]

$$\hat{f}_h(x) = \frac{1}{Mh} \sum_{i=1}^M K\left(\frac{x-X_i}{h}\right) \quad (7)$$

式中: $K(x)$ 为核函数, 常用的高斯核函数形式为

$$K(x) = \frac{1}{\sqrt{2\pi}} \exp(-x^2/2);$$

h 为带宽系数, 取值为 1.8~2.0。

2.3 EMD-QRF 预测模型实现流程

本文所建立的 EMD-QRF 概率密度预测模型实现流程如图 1 所示。首先, 对用户负荷原始时间序列数据采用 EMD 算法进行分解处理, 得到若干特征互异的 IMF。其次, 由于 EMD 分解得到的 IMF 数量较多且 IMF 间存在一定的相关性, 因此有必要对 IMF 进行重构, 从而降低建立预测模型的重复工作量。本文采用样本熵衡量 IMF 复杂度并将样本熵值相近的 IMF 重构为趋势分量、细节分量和随机分量, 对各分量建立 QRF 预测模型可得到不同分位点回归预测结果, 将不同分量预测结果叠加即获得预测值的条件分布。最后, 采用核密度估计方法从预测值条件分布结果获得概率密度预测输出。

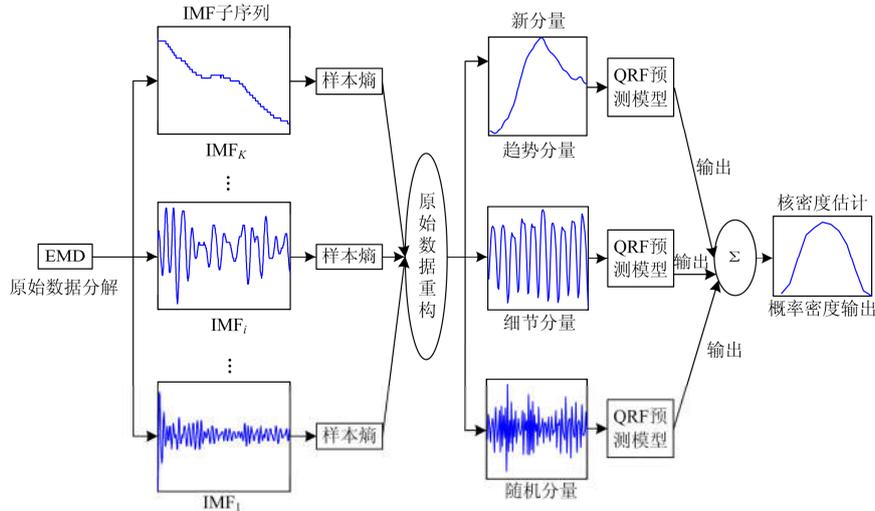


图 1 EMD-QRF 模型实现流程

Fig. 1 EMD-QRF model implementation process

2.4 评价指标

确定性点预测误差评价指标采用平均绝对百分比误差(Mean Absolute Percentage Error, MAPE)和均方根误差(Root Mean Square Error, RMSE)统计, 分别按以下方法计算。

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i} \times 100\% \quad (8)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (9)$$

式中: n 为预测点个数; y_i 为第 i 个预测点用户负荷真实值; \hat{y}_i 为第 i 个预测点模型预测值。

采用预测区间覆盖率 (Forecasting Interval Coverage Percentage, FICP) 和预测区间平均宽度 (Forecasting Interval Average Width, FIAW) 指标评

价区间预测模型有效性, 计算公式分别为

$$FICP = \frac{1}{n} \sum_{i=1}^n \xi_i^{(1-\alpha)} \quad (10)$$

$$FIW_{1-\alpha} = \frac{1}{n} \sum_{i=1}^n \frac{u_i - l_i}{y_i} \quad (11)$$

式中: 置信水平为 $1-\alpha$ 时, ξ 值为 0 或 1, 当实际值在预测区间范围内时, $\xi^{(1-\alpha)} = 1$, 否则 $\xi^{(1-\alpha)} = 0$; u_i 、 l_i 分别为第 i 个预测点置信区间上限和下限。

FICP 指标描述实际观测值落入预测区间的概率, 此值越大, 则预测结果更可靠; FIW 指标刻画预测区间宽度, 此值越小, 则更有利于做决策。

3 算例分析

以江苏省某用户 2016 年 7 月 1 日 01:00 时至 2016 年 7 月 29 日 24:00 时共 696 个(采样时间间隔为 1 h)负荷数据为研究对象, 建立 BP、SVM、QRF、EMD-BP、EMD-SVM 及 EMD-QRF 用户负荷预测模型, 并对比不同模型的预测精度, 以验证本文模型有效性。

3.1 用户负荷原始数据的分解与重构

图 2 为用户负荷时间序列经 EMD 算法分解得到的部分结果。可以看出, 用户负荷原始时间序列数据经 EMD 算法分解后得到 IMF1、IMF2、IMF3、IMF4、IMF5、IMF6、IMF7 共 7 个模态函数和 1 个余量序列 r , 且各模态函数具有一定的周期性, 反映出用户负荷的周周期和日周期用电特征与规律。

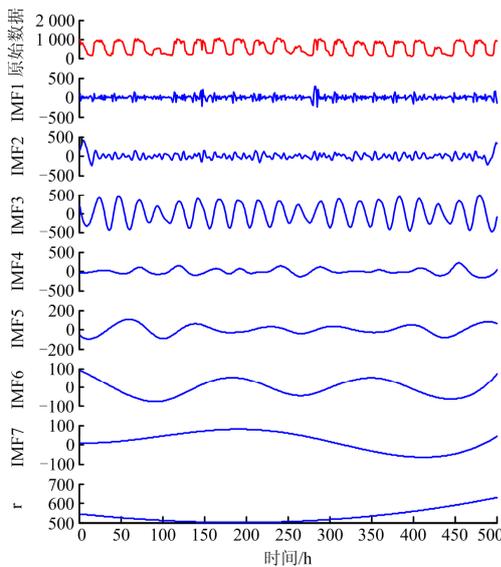


图 2 EMD 算法分解结果

Fig. 2 Decomposition results by EMD algorithm

逐一对各模态函数建立预测模型不仅增加了建模重复工作量, 且容易忽视各模态函数间的相关性。

因此, 本文计算各模态函数样本熵值并根据样本熵值大小对模态函数进行重构处理。图 3 为各 IMF 及余量序列的样本熵值分布。结合图 2 和图 3 可以看出, IMF1 样本熵值与原始数据样本熵值差别最大, 模态函数表现出强烈的随机性与波动性, 该序列是杂乱无章的、规律性差。IMF2—IMF5 各模态函数样本熵值与原始数据相近, 且都在一定阈值范围内, 这些序列表现出较好的规律性。IMF6、IMF7 以及 r 的样本熵值相对较小, 模态函数序列波动性小且变化平缓。因此, 依据样本熵值分布, 将 IMF1 重构为随机分量, IMF2—IMF5 重构为细节分量, IMF6、IMF7 以及 r 重构为趋势分量, 重构后的部分序列如图 4 所示。

趋势分量与用户负荷原始时间序列总体变化趋势一致, 刻画了用户负荷的周周期变化特征。细节分量与用户负荷原始数据变化规律相似, 反映用户负荷日周期与周周期特征, 描述了人类日常活动规律。随机分量规律性差、波动强烈, 一般难以预测或预测误差较大, 反映出气象因素等噪声或干扰对负荷造成的影响。

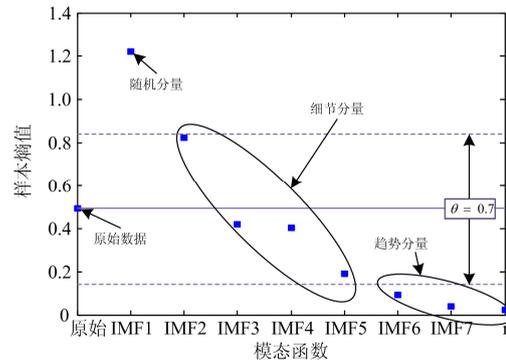


图 3 原始数据及 EMD 分解结果样本熵值分布

Fig. 3 Sample entropy value distribution of original data and EMD decomposition results

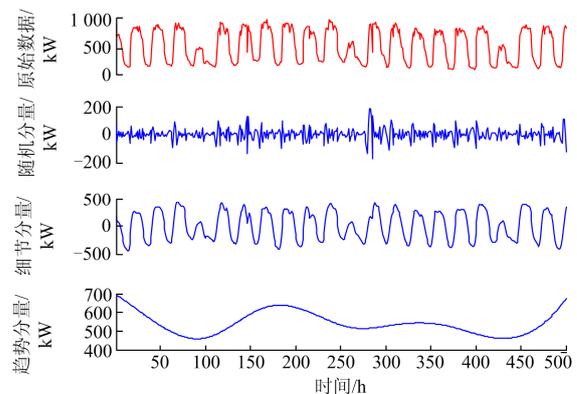


图 4 原始数据及重构分量

Fig. 4 Original data and reconstructed components

3.2 选取输入变量集合

对不同分量利用互信息理论从历史数据中选取输入变量。互信息描述多个线性或非线性变量间共同包含信息量的多少,对于随机变量 X 、 Y ,互信息值表示为^[21]

$$I(X, Y) = \sum_{i=1}^n \sum_{j=1}^m p(x_i, y_j) \log_2 \left(\frac{p(x_i, y_j)}{p(x_i)p(y_j)} \right) \quad (12)$$

式中: n 、 m 分别为随机变量 X 和 Y 的数据总个数。

若 X 与 Y 间共同信息量很大,则说明两变量间相关性较强;若两变量间共同信息量小,甚至为 0 时,表示变量间相关性较弱,甚至相互独立。

设待预测时刻用户负荷为 L_t ,基于用户负荷的周周期和日周期特性,设置原始输入特征集合为待预测时刻前 1 至前 168 个时刻负荷值,即待选输入变量集合为 $\{L_{t-1}, L_{t-2}, \dots, L_{t-168}\}$ 。经过测试,用户负荷原始时间序列及不同重构分量利用互信息选取输入变量集合结果如表 1 所示。其中, Day_{type} 表示日期类型,并用数字 1—7 代表周一至周日。

表 1 原始数据及重构分量输入变量集合

Table 1 Input variable set of original data and reconstructed components

时间序列	输入变量
原始数据	$L_{t-168}, L_{t-57}, L_{t-71}, L_{t-167}, L_{t-24}, L_{t-96}, L_{t-120}, L_{t-144}, L_{t-72}, L_{t-48}, Day_{type}$
随机分量	$L_{t-120}, L_{t-96}, L_{t-144}, L_{t-48}, L_{t-24}, L_{t-168}, L_{t-72}, L_{t-89}, L_{t-79}, L_{t-117}$
细节分量	$L_{t-168}, L_{t-144}, L_{t-167}, L_{t-137}, L_{t-24}, L_{t-72}, L_{t-132}, L_{t-84}, L_{t-107}, Day_{type}$
趋势分量	$L_{t-124}, L_{t-98}, L_{t-125}, L_{t-57}, L_{t-48}, L_{t-45}, Day_{type}$

3.3 预测结果比较

为验证模型有效性,对该用户 7 月 30 日负荷进行提前 24 h 预测。设置 QRF 模型分位点范围为 0.01~0.99,步长为 0.01,对每个预测点即可获得 99 个预测值,在此基础上获得不同置信水平下的用户负荷区间预测结果。进一步地,结合核密度估计及不同分位点预测值实现用户负荷的概率密度预测。

图 5 和表 2 分别为不同模型预测曲线及评价指标统计。可以看出用户负荷真实值曲线基值较小且波动较大,但不同预测模型都能够较好地跟踪用户负荷变化趋势。QRF 预测模型优于 BP 和 SVM 方法,体现为较小的 $MAPE$ 和 $RMSE$ 指标。此外, QRF 模型还能够提供区间预测结果,为售电公司购售电业务提供更多的决策依据。采用 EMD 算法对用户负荷原始时间序列数据进行预处理后,可以进一步提高预测精度,EMD-BP、EMD-SVM 及 EMD-QRF 模型的 $MAPE$ 指标相对于 BP、SVM、QRF 模型分别降低了 13.09%、7.25%和 7.72%, $RMSE$ 指标分

别降低了 5.46%、10.98%和 12.45%。EMD-QRF 模型 $MAPE$ 和 $RMSE$ 指标最小,验证所提模型的有效性。同时,EMD-QRF 模型区间覆盖率较高,预测结果更可靠,但以增加预测区间宽度为代价,体现了 $FICP$ 和 $FIAW$ 指标间的矛盾性。

采用同样步骤对该用户 7 月 31 日负荷进行提前 24 h 预测以验证模型的适应性。图 6 为 QRF、EMD-QRF 模型预测曲线及 80%置信区间范围,表 3 为不同模型误差指标比较。可以看出,EMD-QRF 具有更好的预测性能,表现为较小的 $MAPE$ 和 $RMSE$ 指标。

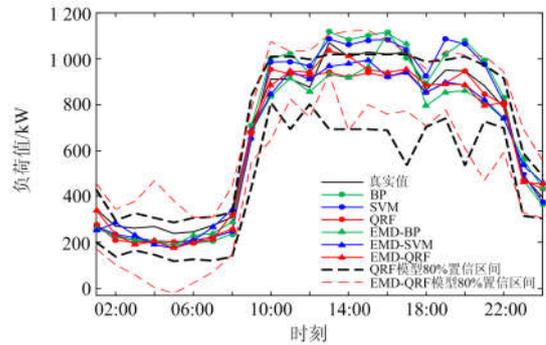


图 5 7 月 30 日用户负荷预测结果

Fig. 5 Consumer load forecasting results on July 30

表 2 7 月 30 日用户负荷预测误差

Table 2 Consumer load forecasting errors on July 30

预测模型	评价指标			
	$MAPE/\%$	$RMSE/kW$	$FICP/\%$	$FIAW/\%$
BP	12.07	69.58	—	—
SVM	10.48	64.3	—	—
QRF	10.23	59.43	83.33	0.457 8
EMD-BP	10.49	65.78	—	—
EMD-SVM	9.72	57.24	—	—
EMD-QRF	9.44	52.03	95.83	0.649 5

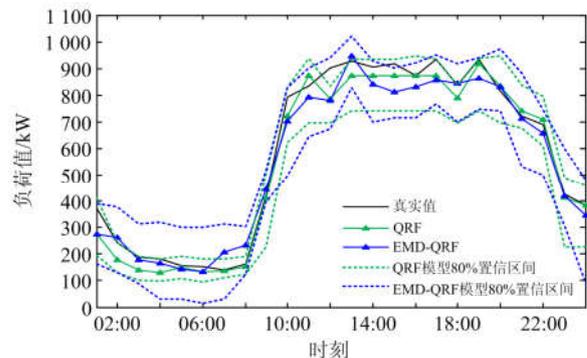


图 6 7 月 31 日用户负荷预测结果

Fig. 6 Consumer load forecasting results on July 31

表 3 7 月 31 日用户负荷预测误差

Table 3 Consumer load forecasting errors on July 31

预测模型	评价指标			
	MAPE/%	RMSE/kW	FICP/%	FIAW/%
QRF	13.48	74.44	79.17	0.39
EMD-QRF	9.23	71.75	95.83	0.70

对 QRF 及 EMD-QRF 模型获得的条件分布采用核密度估计, 可给出任意时刻用户负荷概率密度预测结果, 选取一天中不同时刻概率密度预测如图 7 所示。可以看出 QRF 和 EMD-QRF 模型以较高概率接近真实值, 并可给出用户负荷的波动范围及预测值出现的概率, 从而更有利于做出科学决策, 具有实际应用价值。

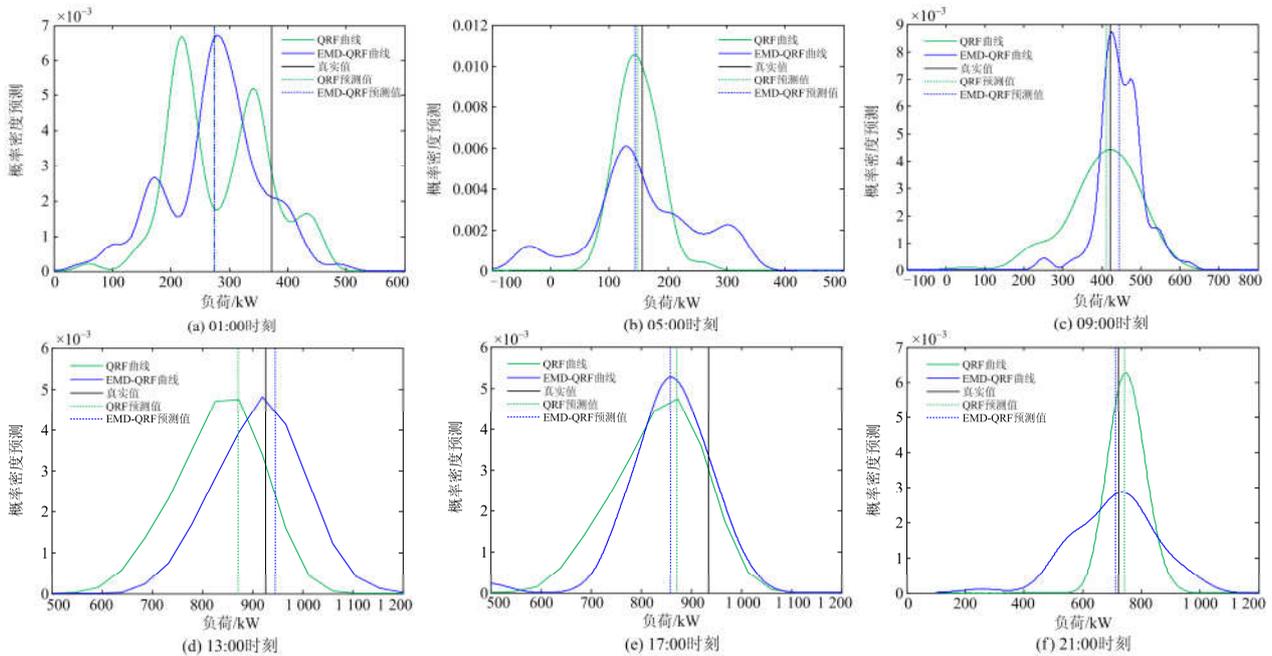


图 7 7 月 31 日不同时刻概率密度预测结果

Fig. 7 Consumer load probability density forecasting results at different times on July 31

4 结论

本文建立的 EMD-QRF 用户负荷概率密度预测模型相对传统的确定性点预测方法, 能够给出用户负荷未来波动范围等更多信息, 有利于为售电公司购售电业务提供更多决策支撑。同时, 利用 EMD 自适应分解算法对用户负荷原始时间序列进行预处理, 能够细致把握用户负荷局部变化特征, 提高预测精度。算例测试验证了本文模型的有效性。

参考文献

[1] 中国能源网. 偏差电量考核有助于形成售电市场寡头 [EB/OL]. [2017-03-22]. <http://www.china-nengyuan.com/news/106038.html>.
China Energy Network. Deviation electricity assessment helps to form an oligopoly in the electricity sales market [EB/OL]. [2017-03-22]. <http://www.china-nengyuan.com/news/106038.html>.

[2] 李若晨, 朱帆, 朱永利, 等. 结合受限玻尔兹曼机的递归神经网络电力系统短期负荷预测[J]. 电力系统保护

与控制, 2018, 46(17): 83-88.

LI Ruochen, ZHU Fan, ZHU Yongli, et al. Short-term power load forecasting using recurrent neural network with restricted Boltzmann machine[J]. Power System Protection and Control, 2018, 46(17): 83-88.

[3] QUILUMBA F L, LEE W J, HUANG H, et al. Using smart meter data to improve the accuracy of intraday load forecasting considering customer behavior similarities[J]. IEEE Transactions on Smart Grid, 2015, 6(2): 911-918.

[4] 杨慧霞, 邓迎君, 刘志斌, 等. 含有历史不良数据的电力负荷预测研究[J]. 电力系统保护与控制, 2017, 45(15): 62-68.
YANG Huixia, DENG Yingjun, LIU Zhibin, et al. Study on electric load forecasting with historical bad data[J]. Power System Protection and Control, 2017, 45(15): 62-68.

[5] 刘云, 张杭, 张爱民. 需求侧响应下基于负荷特性的改进短期负荷预测方法[J]. 电力系统保护与控制, 2018, 46(13): 126-133.
LIU Yun, ZHANG Hang, ZHANG Aimin. Improved load forecasting method based on load characteristics under

- demand-side response[J]. *Power System Protection and Control*, 2018, 46(13): 126-133.
- [6] 高相铭, 杨世凤, 潘三博. 基于 EMD 和 ABC-SVM 的光伏并网系统输出功率预测研究[J]. *电力系统保护与控制*, 2015, 43(21): 86-92.
GAO Xiangming, YANG Shifeng, PAN Sanbo. A forecasting model for output power of grid-connected photovoltaic generation system based on EMD and ABC-SVM[J]. *Power System Protection and Control*, 2015, 43(21): 86-92.
- [7] JAIN R K, SMITH K M, CULLIGAN P J, et al. Forecasting energy consumption of multi-family residential buildings using support vector regression: investigating the impact of temporal and spatial monitoring granularity on performance accuracy[J]. *Applied Energy*, 2014, 123(3): 168-178.
- [8] AHMAD A S, HASSAN M Y, ABDULLAH M P, et al. A review on applications of ANN and SVM for building electrical energy consumption forecasting[J]. *Renewable & Sustainable Energy Reviews*, 2014, 33(2): 102-109.
- [9] 甘雯. 小负荷大波动地区短期负荷预测模型研究[D]. 杭州: 浙江大学, 2007.
GAN Wen. Short-term load forecasting of small load heavy wave area model research[D]. Hangzhou: Zhejiang University, 2007.
- [10] HUANG L, YANG Y, ZHAO H, et al. Time series modeling and filtering method of electric power load stochastic noise[J]. *Protection and Control of Modern Power Systems*, 2017, 2(2): 269-275. DOI: 10.1186/s41601-017-0059-8.
- [11] 谷云东, 张素杰, 冯君淑. 大用户电力负荷的多模型模糊综合预测[J]. *电工技术学报*, 2015, 30(23): 110-115.
GU Yundong, ZHANG Sujie, FENG Junshu. Multi-model fuzzy synthesis for larger consumers[J]. *Transactions of China Electrotechnical Society*, 2015, 30(23): 110-115.
- [12] 蔡剑彪, 罗滇生, 周小宝, 等. 大用户负荷预测管理系统[J]. *电力需求侧管理*, 2012, 14(4): 7-10, 24.
CAI Jianbiao, LUO Diansheng, ZHOU Xiaobao, et al. Load forecasting management system about large consumer[J]. *Power Demand Side Management*, 2012, 14(4): 7-10, 24.
- [13] 罗敏, 程将南, 王毅, 等. 基于小波去噪和决策树的个性化大用户负荷预测[J]. *南方电网技术*, 2016, 10(10): 37-42.
LUO Min, CHENG Jiangnan, WANG Yi, et al. Wavelet de-noising and decision tree based load forecasting of large consumers[J]. *Southern Power System Technology*, 2016, 10(10): 37-42.
- [14] LI Y, GUO P, LI X. Short-term load forecasting based on the analysis of user electricity behavior[J]. *Algorithms*, 2016, 9(4): 80.
- [15] ZHANG W, LIU F, ZHENG X, et al. A hybrid EMD-SVM based short-term wind power forecasting model[C] // *Power and Energy Engineering Conference*, November 15-18, 2015, Brisbane, QLD, Australia: 1-5.
- [16] 王守相, 王亚旻, 刘岩, 等. 基于经验模态分解和 ELM 神经网络的逐时太阳能辐照量预测[J]. *电力自动化设备*, 2014, 34(8): 7-12.
WANG Shouxiang, WANG Yamin, LIU Yan, et al. Hourly solar radiation forecasting based on EMD and ELM neural network[J]. *Electric Power Automation Equipment*, 2014, 34(8): 7-12.
- [17] BREIMAN L. Random forests[J]. *Machine Learning*, 2001, 45(1): 5-32.
- [18] NICOLAI M. Quantile regression forests[J]. *Journal of Machine Learning Research*, 2006, 7(2): 983-999.
- [19] 方匡南. 随机森林理论及其在经济金融中的应用研究[D]. 厦门: 厦门大学, 2010.
FANG Kuangnan. Random forest theory and its application in economic and financial research[D]. Xiamen: Xiamen University, 2010.
- [20] 周松林, 茆美琴, 苏建徽. 风电功率短期预测及非参数区间估计[J]. *中国电机工程学报*, 2011, 31(25): 10-16.
ZHOU Songlin, MAO Meiqin, SU Jianhui. Short-term forecasting of wind power and non-parametric confidence Interval estimation[J]. *Proceedings of the CSEE*, 2011, 31(25): 10-16.
- [21] 赵腾, 王林童, 张焰, 等. 采用互信息与随机森林算法的用户用电关联因素辨识及用电量预测方法[J]. *中国电机工程学报*, 2016, 36(3): 604-614.
ZHAO Teng, WANG Lintong, ZHANG Yan, et al. Relation factor identification of electricity consumption behavior of users and electricity demand forecasting based on mutual information and random forests[J]. *Proceedings of the CSEE*, 2016, 36(3): 604-614.

收稿日期: 2018-09-25; 修回日期: 2018-11-27

作者简介:

杨斌(1977—), 男, 学士, 高级工程师, 研究方向为电力需求侧管理、电能替代、综合能源服务等; E-mail: liuzhiren1835@163.com

杨世海(1976—), 男, 博士研究生, 研究员级高级工程师, 研究方向为电力需求侧管理、智能用电技术; E-mail: ysh.young@163.com

曹晓冬(1990—), 男, 博士, 中级工程师, 研究方向为电力需求侧管理、智能用电技术研究工作。E-mail: cumtcdx@126.com

(编辑 许威)