

DOI: 10.7667/PSPC180381

基于时间分解技术的中远期逐时负荷预测模型

严通煜, 杨迪珊, 项康利, 柯圣舟, 林红阳

(国网福建省电力有限公司经济技术研究院, 福建 福州 350012)

摘要: 中远期电力负荷预测对于引导电网规划建设和提升电力系统资源优化配置具有重要意义。为解决当前中远期负荷预测时间尺度过大、预测精度有限的问题, 利用时间分解技术, 对电力负荷的长期趋势与短期特征分别进行建模分析, 从而提出一种新型预测方法, 将中远期负荷预测的时间尺度缩短至小时, 实现中远期逐时负荷预测。算例分析表明, 所建模型在中远期逐时负荷预测方面的性能优于现有的同类模型, 具有较高的全局精度和稳定性。同时, 能够有效呈现电力负荷的概率密度特征和极值特性, 有望为中远期电力规划提供参考。

关键词: 中远期负荷预测; 逐时负荷; 时间分解; 多元线性回归

Mid-long term hourly load forecasting model based on time decomposition

YAN Tongyu, YANG Dishan, XIANG Kangli, KE Shengzhou, LIN Hongyang

(Economic and Technological Research Institute, State Grid Fujian Electric Power Company Limited, Fuzhou 350012, China)

Abstract: Mid-long term load forecasting is of great significance for guiding the planning and construction of the power grid and improving the optimal allocation of power system resources. At present, the time scale of the medium and long term load forecasting is too large and the prediction precision is limited. To overcome this defect, this paper uses time decomposition technology to respectively analyze the long-term trend and short-term characteristics of power load, and then proposes a new prediction method, which will shorten the time scale of mid-long term load forecasting to hours and achieve mid-long term hourly load forecasting. The case study shows that the performance of the proposed model is better than the existing similar models. It has high global accuracy and stability. At the same time, it can effectively present the probability density characteristics and extreme value characteristics of power load. The proposed method is expected to provide reference for mid-long term planning of power grid.

This work is supported by R & D Consulting Program of Economic and Technological Research Institute of State Grid Fujian Electric Power Company Limited: Establishment of Residential Electricity Model (No. SGFJJY00GHJS1700062) and Science and Technology Project of State Grid Corporation of China: Electric Energy Alternative Planning Model, Policy Simulation and Application Research (No. XM2017020034670).

Key words: mid-long term load forecasting; hourly load; time decomposition; multiple linear regression

0 引言

中远期电力负荷预测对于电力系统的规划建设, 人力、物力与财力资源的投资, 电网运行的经济稳定与安全以及促进国民经济发展等方面具有重

要意义^[1-2]。然而, 由于电力负荷受宏观经济、气象条件等多种外部因素影响, 中远期负荷预测是复杂的多维问题。因此, 提出预测精度更高、时间尺度更小的中远期负荷预测模型是当前亟需解决的问题。目前国内外学者对于中远期负荷预测的研究已取得了一定的进展, 如提出了各种新型的预测方法与模型, 或对现有算法进行了改进等。传统的负荷预测方法包括回归分析、时间序列等。文献[3]基于用电量与国内生产总值、人均国内生产总值以及人口数量等经济变量的强相关关系构建模型, 实现了电力负荷

基金项目: 国网福建经研院研发咨询项目资助 (SGFJJY00GHJS1700062) “居民用电模型建立”; 国家电网公司科技项目资助 (XM2017020034670) “电能替代规划模型、政策模拟及应用研究”

的长期年度预测。文献[4]构建了多变量时间序列反演自记忆模型,对历史电力负荷数据进行拟合。而如灰色系统、遗传算法、神经网络等现代预测方法近年来也有不少国内外学者展开研究。文献[5]利用灰色理论对变权缓冲预处理后的负荷数据进行预测,通过算例验证,该方法可有效提高中远期年度负荷预测的预测精度。文献[6]提出了基于遗传算法的神经网络模型,充分发挥遗传算法的全局寻优能力与神经网络的局部搜索性能。文献[7-8]提出了利用支持向量机进行中远期负荷预测的新方法,运用模拟退火算法自动优化支持向量机模型的参数选择,有效降低支持向量机的建模误差,提高中远期负荷预测的精度。文献[9]提出了一种基于子空间旋转矢量不变技术(ESPRIT)的综合负荷预测方法。文献[10]为了对夏季城市的电力负荷进行预测,根据协整理论采用 SAS 软件建立起电力负荷序列与输入序列“温度”之间的 ARIMAX 模型,充分挖掘序列内部的自相关信息以及序列与序列之间的相关关系。文献[11]将灰色 Elman 神经网络模型应用于中远期负荷预测中,通过将灰色理论与人工神经网络方法结合,可有效降低原始数据突变对模型预测精度造成的影响。随着近年来大数据的不断发展,大数据分析技术也逐渐被学者们应用于中远期电力负荷预测。文献[12-13]采用综合分析的预测模式,在数据分析的基础上实现了负荷预测的建模,有效避免了负荷预测过于依赖模型、数据利用效率低等问题。文献[14]通过大数据聚类方法对大量负荷数据进行聚类划分,进而实现对不同增长特性的负荷区分预测。文献[15]通过对原始数据增长率的分析,采用马尔可夫链划分区间,从几种满足精度要求的模型中筛选出两种进行组合预测,通过方差-协方差方法分配权重。经过该种方法的筛选,不仅可以更准确地选择组合预测模型的类型,而且具有较高精度。综上所述,从目前国内外研究内容来看,虽然目前学者们对于负荷预测的模型与方法已取得了一定的研究成果,但这些方法对于中远期负荷的预测通常以年或月为时间尺度,对于中远期逐时、逐日时间尺度的负荷预测目前还鲜有研究成果。

针对上述问题,本文提出了一种采用时间分解技术、结合长期和短期特征的新型预测方法,首先将影响电力负荷的因素分为高频分量与低频分量两类,再分别构建两类分量的预测模型,并以华东某省为案例进行算例分析。结论表明,本文所提出的

预测模型具有灵活性,预测精度较高,预测性能显著优于非线性自回归神经网络模型,在负荷的概率密度预测方面具有良好的全局性能。本文提出的预测方法可基于不同假设分析包括长期趋势与高频分量变化的不同情景,实现中远期逐时的负荷预测。

1 建模思路与模型框架

电力负荷的动态变化表明了建立预测模型时必须考虑外部因素的影响。本文将影响电力负荷的诸多外部因素按其作用对象和频率分为两类。其中:宏观经济环境和生产技术特性等长期趋势型变量是影响负荷的长期因素,作用于月度、年度的总体用电量,属于低频分量;节假日、气温、昼夜等短期波动型变量是影响负荷的短期因素,作用于逐日、逐时的用电量,属于高频分量^[6]。显然,不论是低频分量还是高频分量,对电力需求的影响都是不可替代的。

本文研究的目的是实现小时尺度上的中远期电力需求预测,因此需要实现低频分量和高频分量的兼顾和融合。一方面,将从长期需求模型中获得的年度电力需求预测结果转化为日频率。另一方面,基于日历效应、温度效应和白昼效应分析得到高频数据对日频率负荷预测结果进行修正,进而得到逐时负荷需求预测结果。具体而言,类似于文献[17]和文献[18],可通过逐时建模获得小时负荷曲线,如式(1)所示。

$$c_{i,t} \Rightarrow \begin{cases} c_{1,t} = \phi_1(LF_t, HF_t) + \varepsilon_{1,t} \\ c_{2,t} = \phi_2(LF_t, HF_t) + \varepsilon_{2,t} \\ \vdots \\ c_{24,t} = \phi_{24}(LF_t, HF_t) + \varepsilon_{24,t} \end{cases} \quad (1)$$

式中: $c_{i,t}$ 表示第 t 天第 i 小时的用电量; LF_t 是电力需求的低频分量(中长期趋势影响),而 HF_t 是高频分量(日历、温度和自昼效应); $\varepsilon_{i,t}$ 是误差项。对于具体特定的小时,低频分量和高频分量对其负荷的作用机理存在差异,式中用函数 ϕ_i 所示,本文假定该函数为线性。需要指出的是,由于日历效应和温度效应及其交互作用的复杂性,式(1)中的随机误差项 $\varepsilon_{i,t}$ 可能表现出一定程度的自相关和交叉相关。但是,由于模型的主要目的是进行预测,因此只要预测结果的精度满足要求,可不必对估计函数 ϕ_i 及模型的随机误差项进行特殊处理。

综上所述,本文构建的中远期逐时负荷预测模型框架如图 1 所示。

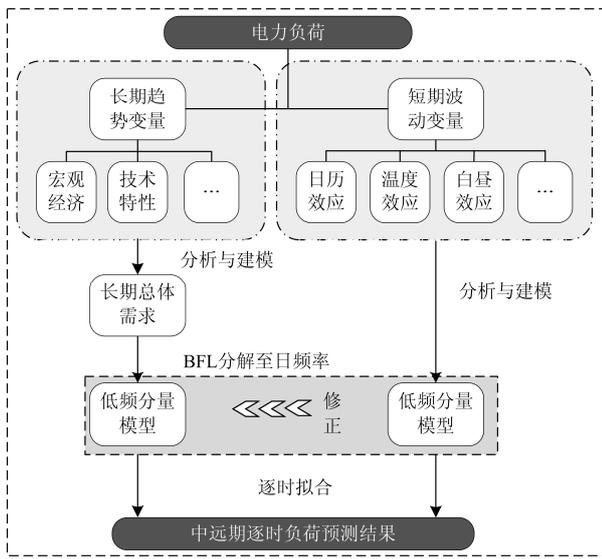


图1 中远期逐时负荷预测模型框架

Fig. 1 Framework of medium and long term load forecasting model

2 高低频分量预测模型

2.1 基于LDTM模型的负荷低频分量预测模型

如前文所述，本文中负荷的低频分量(LF)表征电力负荷变化的中长期趋势，该趋势反映了经济、技术因素驱动下电力需求稳步增长的本质。本文采用长期需求模型(Long Term Demand Model, LTDM)对其进行预测。文献[19]表明，我国的居民和非居民用电需求的形态特征和发展规律存在显著差异。本文分别对这两种类型的电力需求建模。为此，本文构建的长期预测模型如式(2)所示。

$$C_T = \sum_{i=1}^S C_{i,T} = \sum_{i=1}^{S-1} Y_{i,T} I_{i,T} + H_T I_{H,T} \quad (2)$$

式中，第 T 年的总需求 C_T 可通过加总各部门具体用电量 $C_{i,T}$ 得到。对于非居民用户($i=1, \dots, S-1$)，用电量可以通过具体部门的用电强度 $I_{i,T}$ 乘以增长总值 $Y_{i,T}$ 计算得出。而居民用户的用电量，可通过户均用电强度 $I_{H,T}$ 与户数 H_T 相乘得到。

在模型中，各部门总增加值与家庭数量是外生变量，这些变量代表了影响未来电力需求的社会经济环境。同时，本文将各类用户的用电强度定义为有 3 个变量的函数，如式(3)所示。

$$I_{i,T} = f(\text{Num}_{i,T}, \eta_{i,T}, SE_{i,T}) \quad (3)$$

式中： $\text{Num}_{i,T}$ 表示生产部门或家庭中用电设备的数量； $\eta_{i,T}$ 表示效率与电气化系数，包括电气化设备的占比以及设备的整体能耗水平； $SE_{i,T}$ 表示电力与其替代能源(气、油等)价格相对变化影响导致的替代

效应。在实际计算中，生产部门的用电强度和居民用电强度可直接采用统计年鉴中的能源强度指标和户均用电量指标近似代替。

上述长期需求模型对长期电力需求，即负荷低频分量提供了基准估计，旨在捕捉电力需求的长期发展趋势，是本文所构建预测模型的基础。然而，长期需求模型得出的总需求时间尺度为一年。为实现逐时负荷预测，需将其分解至日频率，同时保证数据不失真。为此，本文采用文献[20]提出的布特-菲比斯和利斯曼分解法(Boot-Feibes and Lisman Disaggregation Method, BFL)，该方法的主要优点在于不需要额外指标，且尤其适用于从年度数据到逐日数据的分解。

BFL 方法通过线性优化得到高频序列 y_i ，使其实现最佳平滑性并满足序列纵向约束，如式(4)所示。

$$\begin{cases} \min = \sum_{i=2}^M (y_i - y_{i-1})^2 \\ \text{s.t.} \sum_{i=1}^{365/366} y_{i,T} = Y_T, \forall T \end{cases} \quad (4)$$

式中，目标函数表示高频序列相邻两项差值的平方和最小，即序列有最佳平滑性。其中： M 表示高频序列的样本数，即若要将连续两个平年的年度数据分解至日频率，则 $M=365 \times 2=730$ ； y_i 表示其中的第 i 个样本。式中，约束条件要求分解后每一年内的 365 或 366 个样本之和与分解前的年度数据相等，即序列纵向约束。其中： $y_{i,T}$ 表示第 T 年的第 i 个高频样本； Y_T 表示原始序列中第 T 年的年度电力需求。

该线性优化问题可用 CPLEX 直接进行求解，具体过程本文不再赘述。

至此，可得到负荷低频分量预测结果，该结果能够反映电力需求演化发展的长期趋势，且已被分解为逐日数据。

2.2 计及多元外生变量的负荷高频分量预测模型

对于负荷高频分量的预测，本文重点考虑以下三类影响：日历效应、温度效应和白昼时间。

2.2.1 日历效应^[21]

日历效应是指用电量与日期相关的波动，包括星期效应、假日效应和月份效应。星期效应是指工作日和周末的连续重复形成了一个潜在的 7 天周期性，这使得用电量在该周期内也呈现特定的循环规律。与此同时，8 h 工作制造成的日内负荷波动也属于星期效应的范畴。假日效应指节假日对用电量的影响，将会覆盖星期效应。月份效应，即温度变化不能解释的季节性，月份效应在暑假和春节期间较为显著，对其他月份也有不同程度的影响。

本文提出因果模型以处理日历效应, 该模型的主要优点是能够对样本中日历效应中难以用动态模型捕捉的非周期特征(农历假期、闰年等)进行有效处理。日历效应模型如式(5)所示。

$$f_i(D_t) = \sum_{w=1}^7 \beta_{wi} D_{wt} |D_{it} - 1| + \sum_{j=1}^{12} \beta_{ji} D_{jt} + \sum_{k=1}^{12} \beta_{ki} M_{k-f,t} \quad (5)$$

式中: $f_i(D_t)$ 表示日历效应对第 t 日第 i 时负荷的影响; 等号右侧第一项表示星期效应, D_{wt} 为虚拟变量, 当第 t 天是星期 w 时为 1; β_{ji} 表示星期 w 对第 i 小时电力负荷的影响; 等号右侧第二项表示假日效应, D_{jt} 为虚拟变量, 当第 t 天属于第 j 类假日时为 1; β_{ji} 表示假日类型 j 对 i 小时电力负荷的影响。其中, 当 $D_{jt}=1$ 时, $D_{j-1,t}=0$, 此时星期效应被假日效应覆盖。等号右侧第三项表示月份效应, $M_{k-f,t}$ 为虚拟变量, 当第 t 天属于第 k 月时为 1; β_{ki} 表示第 k 月对 i 小时负荷的影响。对节假日的分类如表 1 所示。

表 1 节假日分类

Table 1 Holiday categories

节假日类型	虚拟变量数量	全年天数
3 天节假日的第 a 天($a=1,2,3$)	3	$3 \times 5=15$
7 天节假日的第 b 天($b=1,2, \dots, 7$)	7	$2 \times 7=14$
节假日的前一天	1	$5+2=7$
节假日的后一天	1	$5+2=7$
合计	12	43

关于表 1 有以下说明: 第一, 我国的 3 天节假日包括元旦、清明、劳动、端午和中秋, 共计 5 个; 7 天节假日包括国庆和春节, 共计 2 个。第二, 文献[22]表明, 不同类型假期中每一天的用电特征存在差异, 为实现负荷的精准预测, 在基础数据较为充分的情况下, 本文将同类节假日的每一天都算作一类节假日并对其假日效应进行单独测算。第三, 节假日的前一天、后一天的负荷与仅受星期效应影响下的负荷也存在差异, 因此本文将此类日期也算作节假日进行处理。综上, 全年节假日共有 43 天, 分为 12 类。

2.2.2 温度效应

外部温度是决定电力需求的另一因素, 尤其是对中短时间尺度下的负荷预测结果具有显著影响, 具体有以下方面: 第一, 温度改变了电力消耗的变化方向或需求状态, 如当温度在较高时进一步升高, 空调的使用将增加电力消耗, 但当温度在较低时升高, 则会降低电力消耗。第二, 根据需求温度弹性, 温度变化影响电力需求。第三, 日历效应可能与温度效应相互影响。本文对温度效应建模沿两个维度进行, 第一个维度估计需求行为调整时的温度阈值,

首先, 构建如式(6)所示的分段线性辅助回归。

$$E_{i,t}^{HP} = \begin{cases} \beta_{i0} + \beta_{i1}T_t + \varepsilon'_{it} & T \leq \underline{\tau} \\ \varphi_{i0} + \varepsilon'_{it} & \underline{\tau} \leq T \leq \bar{\tau} \\ \theta_{i0} + \theta_{i1}T_t + \varepsilon'_{it} & T \geq \bar{\tau} \end{cases} \quad (6)$$

式中: $E_{i,t}^{HP}$ 为经 H-P 滤波器剔除长期趋势后, 第 t 天第 i 小时的负荷数据; $\underline{\tau}$ 和 $\bar{\tau}$ 分别为低温阈值和高温阈值; ε'_{it} 为随机误差项; T_t 为第 t 天的日平均气温; 其余变量均为回归参数。通过反复给定高低温阈值进行多次回归并进行 AIC 准则验证, 当 AIC 最小时, 对应的低温和高温阈值即为符合要求的估计值。

需要说明, 温度效应更多地与室内温度相关, 鉴于本文需对每小时负荷分别建模, 理想情况下最好使用每小时室内温度数据。然而, 获取每小时温度数据代价较高, 同时由于室内外温度的关系受建筑热力学性能的影响, 无法推导出由室外温度推算室内温度的普适公式, 因此本文在温度效应建模中只采用日频率室外温度。在数据受限的情况下, 为更好地捕捉真实的温度效应, 本文在建模中引入了最低、最高与日平均气温以及连续五天以上的移动平均温度; 同时, 为解决温度弹性中的非线性问题, 本文将温度的平方也作为解释变量。综上所述, 温度效应如式(7)所示。

$$g_i(T) = g_i \left(\begin{matrix} T \min_t, T \min_{t-1}, \dots, T \min_{t-4} \\ T \max_t, T \max_{t-1}, \dots, T \max_{t-4} \end{matrix} \right) = \beta_{0i} (T_t < \underline{\tau}) + \beta_{1i} (T_t \geq \bar{\tau}) + \beta_{2i} T_t (T_t < \underline{\tau}) + \beta_{3i} T_t (T_t \geq \bar{\tau}) + \sum_{l=1}^4 \theta_{ij} T_{t-j} (T_{t-j} < \underline{\tau}) + \varphi_{ij} T_{t-j} (T_{t-j} \geq \bar{\tau}) + \beta_{4i} R_t^5 (T_t < \underline{\tau}) + \beta_{5i} R_t^5 (T_t \geq \bar{\tau}) + \beta_{6i} A_t (T_t < \underline{\tau}) + \beta_{7i} A_t (T_t \geq \bar{\tau}) + \beta_{8i} T_t^2 (T_t \geq \bar{\tau}) + \beta_{8i} T_t^2 (T_t < \underline{\tau}) + \beta_{9i} (T \max_t - \bar{\tau}) (T_t > \bar{\tau}) + \beta_{10i} (\underline{\tau} - T \min_t) (T_t < \underline{\tau}) \quad (7)$$

式中: g_i 表示温度效应, 是温度 T 的函数; T_t 为日平均温度, 取日内最高温和最低温的均值; $\underline{\tau}$ 和 $\bar{\tau}$ 分别是低温和高温的估计阈值; R_t^5 为连续五天以上平均气温的移动平均值; A_t 为日内温差; 其余变量均为回归参数。

2.2.3 白昼效应

白昼效应是指白昼时间对用电量的影响, 通常白昼时间越短, 照明耗电越多, 用电量越大。我国经纬度跨度都较大, 导致不同地区、不同季节的白

昼时间差异显著,因此本文将白昼效应也纳入模型考量中。具体而言,在逐时分析预测的视角下,白昼效应主要体现在日出、日落时刻对临近时段用电量的影响,对稳定处于白昼或夜间的时段则无影响。通过查阅气象数据,在全年内我国日出时间均分布在北京时间 5:00—7:00;日落时间均分布在 17:00—19:00。因此,对于上述时间段的负荷预测,本文引入白昼效应作为高频分量以解释白昼时间对用电量的影响,如式(8)所示。

$$m_i(Sr_i, Ss_i) = \begin{cases} \theta_{sr1} + \theta_{sr2} Sr_i & i = 5, 6, 7 \\ \theta_{ss1} + \theta_{ss2} Ss_i & i = 17, 18, 19 \end{cases} \quad (8)$$

式中: m_i 表示白昼效应对第 i 时负荷的影响,是日出时间 Sr_i 和日落时间 Ss_i 的函数;其余变量均为回归参数。

2.3 中远期逐时负荷预测结果

在前文分别对低频分量和高频分量建模预测的基础上,中远期逐时负荷预测结果是低频分量 LF_t 与高频分量 HF_t 的函数。每小时需求如式(9)所示。

$$c_{it} = \varphi_i(LF_t, HF_t) + \varepsilon_{i,t} = \begin{cases} h_i(C_t) + f_i(D_t) + g_i(T_t) + m_i(Sr_i, Ss_i) + \varepsilon_{i,t} & i = 5, 6, 7, 17, 18, 19 \\ h_i(C_t) + f_i(D_t) + g_i(T_t) + \varepsilon_{i,t} & \text{else} \end{cases} \quad (9)$$

式中: c_{it} 为第 t 天 i 小时的负荷; f_i 为日历效应,由式(5)给出; g_i 为温度效应,由式(7)给出; m_i 为白昼效应,由式(8)给出; h_i 为第 t 天 i 小时的低频分量,是长期预测模型预测结果的函数,如式(10)所示。

$$h_i(C_t) = \omega_i C_t \quad (10)$$

式中: C_t 为长期需求模型预测值应用 BFL 方法分解到日尺度上的预测结果; ω_i 为回归参数。

模型参数通过最小二乘估计,即最小化式 $\min \sum_{t=1}^T [c_{i,t} - \varphi_i(LF_t, HF_t)]^2$ 。所有时间都考虑同样的标准,本文可用数据范围为 2000 年 1 月 1 日至 2016 年 12 月 31 日,因此,每小时有超过 5 800 个电力负荷的日观测值,能够满足参数估计的计量经济学要求。

3 算例分析

本节以我国华东某省作为研究对象,以 2000 年 1 月 1 日至 2016 年 12 月 31 日的历史数据为基础,选择不同的样本和评价标准,从而对本文所建模型的性能进行验证评估。历史数据从该省电网公司调度部门和气象部门获得,同期宏观经济情况参考当地的统计年鉴。

3.1 模型预测误差分析

本节首先分析模型再现不同时间跨度数据的准确性,旨在验证模型的整体精度和稳定性。各预测样本如表 2 所示。

CASE1: 已知 2000/1/1—2015/12/31, 预测 2016/1/1—2016/12/31; CASE2: 已知 2000/1/1—2014/12/31, 预测 2015/1/1—2016/12/31; CASE3: 已知 2000/1/1—2013/12/31, 预测 2014/1/1—2016/12/31。对比真实数据,计算每小时预测的平均百分比误差(MPAE),相关结果如表 2 所示。

由表 2 可见,在上述案例中,预测结果的 MPAE 水平在 2%~4% 浮动,随着预测时间跨度的增长,预测误差小幅增大,但仍然控制在较低的水平内,说明模型的精度和稳定性都较理想。在上午、午后和晚间的用电高峰时段,模型预测精度很高;而在凌晨用电量相对较低时的预测效果相对较差。

表 2 各样本逐时预测误差汇总表

Table 2 Hourly prediction error for each sample

样本	CASE1		CASE2		CASE3	
	2016	2016	2015	2016	2015	2014
1:00	3.01	3.09	2.94	3.11	3.66	3.01
2:00	3.2	3.33	3.12	3.46	3.82	3.2
3:00	3.63	3.89	3.57	3.79	4.28	3.63
4:00	3.26	3.37	3.06	3.87	4.02	3.26
5:00	2.43	2.88	2.51	3.06	3.08	2.43
6:00	2.09	2.43	2.19	2.58	2.96	2.09
7:00	2.15	2.31	2.28	2.36	2.88	2.15
8:00	2.83	3.16	2.46	3.02	3.37	2.83
9:00	2.68	2.83	2.7	2.77	3.12	2.68
10:00	2.57	2.86	2.45	2.79	3.36	2.57
11:00	2.21	2.46	2.02	2.52	3	2.21
12:00	2.21	2.68	2.53	2.85	2.82	2.21
13:00	1.98	2.35	2.21	2.32	2.53	1.98
14:00	2.12	2.51	2.11	2.51	2.68	2.12
15:00	2.23	2.55	2.29	2.6	2.98	2.23
16:00	2.43	2.69	2.41	2.61	2.77	2.43
17:00	2.52	2.92	2.49	2.87	3.06	2.52
18:00	2.31	2.82	2.48	2.9	3.12	2.31
19:00	2.51	2.93	2.67	2.77	3.37	2.51
20:00	2.72	3.01	2.52	2.98	3.01	2.72
21:00	1.86	2.98	1.76	3.03	3.18	1.86
22:00	1.99	2.67	2.18	2.73	3.12	1.99
23:00	2.23	3.02	2.46	3.12	3.34	2.23
24:00	2.96	3.54	2.99	3.38	3.69	2.96

由于本文的模型旨在实现中远期的逐时负荷预测,因此有必要确定模型最大有效预测年限。为此,

本文在上述 CASE1—CASE3 的基础上, 进一步扩大时间跨度, 逐步减少已知数据, 对 2016/1/1—2016/12/31 的逐时负荷进行多次预测并以各小时 MPAE 水平的均值作为观测指标。测算结果如图 2 所示。

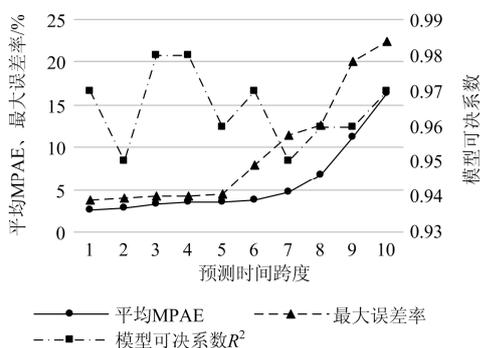


图 2 不同预测时间跨度下的误差和拟合效果

Fig. 2 Error and fitting effect under different forecasting time spans

由图 2 可见, 不同时间跨度下的模型可决系数 R^2 都在 0.95 以上, 说明回归拟合具有显著性。误差方面, 模型的平均 MPAE 和最大误差率均与预测时间跨度正相关。当预测时间跨度在 5 年内时, 误差水平较低, 平均 MPAE 和最大误差率都控制在 5% 以内; 但当预测时间跨度超过 5 年后, 模型的最大误差率开始增加, 超过 7 年后, 平均 MPAE 水平开始显著增加。这说明, 本文所建模型进行中远期逐时负荷预测时, 能够实现 5 年内逐时负荷的准确预测; 随后模型的预测性能开始恶化, 但在 5~7 年内的预测结果仍然有一定的参考价值; 当预测时间跨度超过 7 年后, 模型的预测性能较差。

表 3 本文所建模型与 NARX 模型的性能对比

Table 3 Performance comparison between the model proposed in this paper and NARX

	1 年跨度预测		3 年跨度预测		5 年跨度预测	
	平均 MPAE/%	最大误差率/%	平均 MPAE/%	最大误差率/%	平均 MPAE/%	最大误差率/%
本文模型	2.53	3.65	3.22	4.28	3.61	4.53
$ni=8; nl=11$	2.59	3.43	10.91	15.26	15.24	17.23
NARX $ni=3; nl=11$	2.83	3.77	12.72	17.11	18.28	22.36
$ni=1; nl=13$	3.01	5.11	10.34	13.35	14.52	19.14

由表 3 可见, 当预测时间跨度较小时, 本文所建模型和 NARX 模型均能实现逐时负荷的精准预测, 本文模型的优势不明显。但随着时间跨度增加, NARX 模型的性能下降较为明显, 而本文所建模型的精度仍然保持在高水平。可见, 在中远期逐时负荷预测方面, 本文所建的模型相较于 NARX 模型具有显著优势。

3.2 模型性能对比分析

本节对模型的预测性能进行比较, 将本文构建的模型与具有外源输入的非线性自回归神经网络模型(NARX)的结果进行比较。NARX 模型能够处理大量外生变量的信息得到预测结果, 与本文模型的功能和效果相似。NARX 模型的输入-输出函数如式 (11)所示。

$$c_{i,t} = \sum_j^{ni} w_j^i \left[f \left(\sum_{k=1}^{nx} w_{j,k}^i x_{k,t}^i + \sum_{l=1}^{nl} w_{j,l}^i c_{i,t-l} \right) \right] + e_{i,t} \quad (11)$$

式中: $c_{i,t}$ 表示第 t 天 i 小时的负荷值; $c_{i,t-l}$ 表示滞后 l 天 i 小时的负荷值; $x_{k,t}^i$ 表示用于解释第 t 天 i 小时负荷的 nx 个外生变量。隐含层有 ni 个神经元, 每个神经元有 nx 个输入或外生变量, 滞后 nl 期的输出变量权重分别为 $w_{j,k}^i$ 与 $w_{j,l}^i$ 。隐含层采用 Sigmoid 传递函数^[23], 在每一个输出层采用线性传递函数, 其中权值为 w_j^i 。 $e_{i,t}$ 为误差项, 即电力负荷实际值与 NARX 模型估计值之差。

同样由式(11)可见, NARX 模型中的外生变量由两部分构成: 第一部分与本文所建模型高频分量中的变量一致, 包括日历效应变量、温度效应变量和白昼时间变量; 第二部分是负荷曲线在小时 i 的滞后值, 可用于捕捉负荷变化的整体趋势。

为确定隐含层单元的数量 ni 和输出滞后期数 nl , 训练并检验 ni 和 nl 区间分别为 [1,15] 与 [1,14] 的 NARX 模型, 保留模型 MPAE 最优的 3 组 ni 与 nl , 并将所得到的参数进行多次训练并检验, 保留最优的 3 个 NARX 模型用于对比。分别用 3 个模型进行跨度为 1、3、5 年的逐时负荷预测, 预测结果的相关误差统计指标如表 3 所示。

3.3 负荷分布预测性能分析

由于进行中远期逐时负荷预测的主要目的之一是估计中长期峰谷特性以指导电网规划。因此, 除了对总体误差水平进行评估之外, 本节利用概率密度图对预测模型呈现中长期负荷整体分布和极值情况。利用 2000—2013 年的基础数据对 2016 年进行逐时负荷预测, 实际值和预测值的概率密度如图 3 所示。

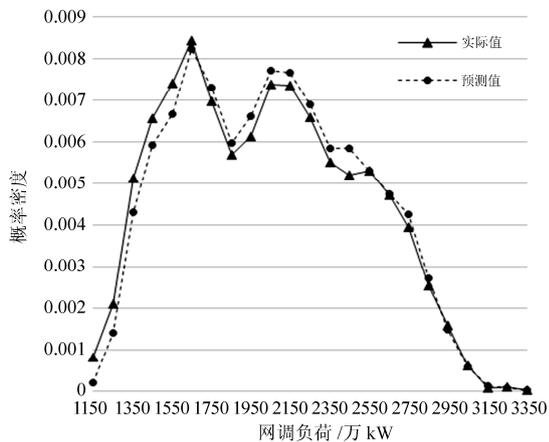


图3 实际负荷值与预测负荷值的概率密度

Fig. 3 Probability density of actual and predicted load value

由图3可见,模型的全局性能良好,重现了负荷值的双峰分布。当网调负荷较小时,预测值的概率密度与实际值存在一定误差;随着网调负荷增大,预测值和实际值的误差显著减小,在高负荷区两条曲线几乎重合。结合表2的数据可见,对低负荷区的预测是本文所建模型的弱点之一。但由于模型的主要目的是帮助进行电网容量规划,此弱点并非重要限制,负荷的最大值比最小值更为关键。

此外,图中的预测值和实际值都右偏,且右尾较大,这表明极端的外部条件(如极端气温)将导致高峰负荷的显著增加,与用户行为特征一致。模型能够对这一特性进行有效再现,使得在进行中长期规划时允许向不同的高峰负荷分配不同的概率,从而确定电网的安全裕度。

4 结论

本文提出了一种采用时间分解技术、结合长期和短期特征的新型预测方法,该方法具有灵活性,可基于不同假设分析包括长期趋势与高频分量变化的不同情景,实现中远期逐时负荷预测。通过算例分析对模型的性能进行综合评估,得到以下结论:

1) 本文所建模型对中远期逐时负荷预测结果的MPAE水平在2%~4%浮动,模型具有较高的精度。预测误差与时间跨度呈正相关,当预测时间跨度在5年内时,MPAE和最大误差率分别不超过3.61%和4.53%,预测结果具有较好的参考价值。

2) 通过和NARX模型进行对比表明,随着时间跨度增加,NARX模型的性能下降较为明显,而本文所建模型的精度仍然保持在高水平。可见,在中远期逐时负荷预测方面,本文所建的模型相较于NARX模型具有显著优势。

3) 本文所建模型在负荷的概率密度预测方面具有良好的全局性能,尤其在高负荷区域,概率密度的预测值和实际值高度相似。此外,模型还能够再现负荷概率密度的右偏特性,这对于指导中长期电网规划有积极意义。

4) 本文所建模型对谷时段负荷的预测精度相对较低,虽然这一弱点在现阶段对模型的应用不构成显著制约,但未来还应通过增加解释变量、调整回归函数形态等方式对模型进一步优化。

参考文献

- [1] 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.
- [2] 潘超, 鲁宝春, 孙颖, 等. 电力系统负荷预测方法综述[J]. 辽宁工业大学学报(自然科学版), 2016, 36(5): 296-299.
PAN Chao, LU Baochun, SUN Liying, et al. Review of load forecasting methods in power system[J]. Journal of Liaoning University of Technology (Natural Science Edition), 2016, 36(5): 296-299.
- [3] BIANCO O, MANCA S, NARDINI. Linear regression models to forecast electricity consumption in Italy[J]. Energy Sources, 2013(8): 86-93.
- [4] 邹晶晶, 姚建刚, 孔维辉, 等. 基于多变量时间序列反演记忆模型的中长期电力负荷预测[J]. 电力系统及其自动化学报, 2017, 29(10): 98-105.
ZOU Pinjing, YAO Jiangang, KONG Weihui, et al. Mid-long term power load forecasting based on multivariable time series inversion self-memory model[J]. Proceedings of the CSU-EPSCA, 2017, 29(10): 98-105.
- [5] 王大鹏, 汪秉文. 基于变权缓冲灰色模型的中长期负荷预测[J]. 电网技术, 2013, 37(1): 167-171.
WANG Dapeng, WANG Bingwen. Medium-and long-term load forecasting based on variable weights buffer grey model[J]. Power System Technology, 2013, 37(1): 167-171.
- [6] 李莉, 刘建勋, 刘崇新. 基于改进型遗传算法的混沌神经网络在电力负荷预测的应用[J]. 华中电力, 2010, 23(2): 13-17.
LI Li, LIU Jianxun, LIU Chongxin. Application of chaotic neural network based on improved genetic algorithm in load prediction[J]. Central China Electric Power, 2010, 23(2): 13-17.
- [7] 李瑾, 刘金朋, 王建军. 采用支持向量机和模拟退火算法的中长期负荷预测方法[J]. 中国电机工程学报, 2011, 31(16): 63-66.

- LI Jin, LIU Jinpeng, WANG Jianjun. Mid-long term load forecasting based on simulated annealing and SVM algorithm[J]. Proceedings of the CSEE, 2011, 31(16): 63-66.
- [8] NIU Dongxiao, MA Tiannan, LIU Bingyi. Power load forecasting by wavelet least squares support vector machine with improved fruit fly optimization algorithm[J]. Journal of Combinatorial Optimization, 2017, 33(3): 1122-1143.
- [9] 马哲, 舒勤. 基于 ESPRIT 分解算法的短期电力负荷预测[J]. 电力系统保护与控制, 2015, 43(7): 90-96.
MA Zhe, SHU Qin. Short term load forecasting based on ESPRIT integrated algorithm[J]. Power System Protection and Control, 2015, 43(7): 90-96.
- [10] 崔和瑞, 彭旭. 基于 ARIMAX 模型的夏季短期电力负荷预测[J]. 电力系统保护与控制, 2015, 43(4): 108-114.
CUI Herui, PENG Xu. Summer short-term load forecasting based on ARIMAX model[J]. Power System Protection and Control, 2015, 43(4): 108-114.
- [11] 张健美, 周步祥, 林楠, 等. 灰色 Elman 神经网络的电网中长期负荷预测[J]. 电力系统及其自动化学报, 2013, 25(4): 145-149.
ZHANG Jianmei, ZHOU Buxiang, LIN Nan, et al. Prediction of mid-long term load based on gray Elman neural networks[J]. Proceedings of the CSU-EPSA, 2013, 25(4): 145-149.
- [12] 程潜善, 方华亮. 一种应用大数据技术的中长期负荷预测方法[J]. 武汉大学学报(工学版), 2017, 50(2): 239-244.
CHENG Qianshan, FANG Hualiang. A prediction method for mid-long term load forecasting using big data technology[J]. Engineering Journal of Wuhan University, 2017, 50(2): 239-244.
- [13] 袁铁江, 袁建党, 晁勤, 等. 电力系统中长期负荷预测综合模型研究[J]. 电力系统保护与控制, 2012, 40(14): 143-146, 151.
YUAN Tiejiang, YUAN Jiandang, CHAO Qin, et al. Study on the comprehensive model of mid-long term load forecasting[J]. Power System Protection and Control, 2012, 40(14): 143-146, 151.
- [14] 徐源, 程潜善, 李阳, 等. 基于大数据聚类的电力系统中长期负荷预测[J]. 电力系统及其自动化学报, 2017, 29(8): 43-48.
XU Yuan, CHENG Qianshan, LI Yang, et al. Mid-long term load forecasting of power system based on big data clustering[J]. Proceedings of the CSU-EPSA, 2017, 29(8): 43-48.
- [15] 张栋梁, 严健, 李晓波, 等. 基于马尔可夫链筛选组合预测模型的中长期负荷预测方法[J]. 电力系统保护与控制, 2016, 44(12): 63-67.
ZHANG Dongliang, YAN Jian, LI Xiaobo, et al. Mid-long term load forecasting based on Markov chain screening combination forecasting models[J]. Power System Protection and Control, 2016, 44(12): 63-67.
- [16] 熊军华, 牛珂, 张春歌, 等. 基于小波变异果蝇优化支持向量机短期负荷预测方法研究[J]. 电力系统保护与控制, 2017, 45(13): 71-77.
XIONG Junhua, NIU Ke, ZHANG Chungge, et al. LSSVM in short-term load forecasting based on wavelet transform and mutant fruit fly optimization algorithm[J]. Power System Protection and Control, 2017, 45(13): 71-77.
- [17] HYNDMAN R J, FAN S. Density forecasting for long-term peak electricity demand[J]. IEEE Transactions on Power Systems, 2010, 25: 1142-1453.
- [18] SOARES L J, MEDEIROS M C. Forecasting electricity demand using generalized long memory[J]. International Journal of Forecasting, 2006, 22: 17-28.
- [19] 陈伟, 乐丽琴, 崔凯, 等. 典型用户负荷特性及用电特点分析[J]. 能源技术经济, 2011, 23(9): 44-49.
CHEN Wei, YUE Liqin, CUI Kai, et al. Analysis on power load and consumption characteristics of representative consumers[J]. Energy Technology and Economics, 2011, 23(9): 44-49.
- [20] BOOT J, FEIBES W, LISMAN J. Further methods of derivation of quarterly figures from annual data[J]. Journal of the Royal Statistical Society, 1967, 16(1): 65-75.
- [21] 李啸聪, 李春涛, 从兰美, 等. 基于动态权值相似日选取算法的短期负荷预测[J]. 电力系统保护与控制, 2017, 45(6): 1-8.
LI Xiaocong, LI Chuntao, CONG Lanmei, et al. Short-term load forecasting based on dynamic weight similar day selection algorithm[J]. Power System Protection and Control, 2017, 45(6): 1-8.
- [22] MORAL-CARCEDO J, PEREZ-GARCIA J. Integrating long-term economic scenarios into peak load forecasting: an application to Spain[J]. Energy, 2017, 140: 682-695.
- [23] SAMARASINGHE S. Neural networks for applied sciences and engineering[M]. Auerbach Publications, 2006.

收稿日期: 2018-04-04; 修回日期: 2018-06-11

作者简介:

严通煜(1990—), 男, 通信作者, 工程师, 主要研究方向为电网规划与电力经济; E-mail: yantongyu1990@163.com

杨迪珊(1988—), 女, 工程师, 主要研究方向为土木工程与电力经济;

项康利(1990—), 女, 助理工程师, 主要研究方向为电力经济。

(编辑 魏小丽)