

极值负荷及其出现时刻的概率化预测

陈新宇¹, 康重庆¹, 陈敏杰²

- (1. 电力系统及发电设备控制和仿真国家重点实验室(清华大学电机系), 北京市 海淀区 100084;
2. 麻省理工学院电气工程与计算机科学系, 美国 马萨诸塞州 剑桥市 02139)

Short Term Probabilistic Forecasting of the Magnitude and Timing of Extreme Load

CHEN Xinyu¹, KANG Chongqing¹, CHEN Minjie²

(1.State Key Lab of Control and Simulation of Power Systems and Generation Equipments(Dept. of Electrical Engineering, Tsinghua University), Haidian District, Beijing 100084, China; 2. Dept. of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, 02139, Cambridge, MA, USA)

ABSTRACT: As the foundation of system daily scheduling and operations, current deterministic forecasting algorithm of the magnitude and timing of extreme load is not satisfactory. Probabilistic forecasting is an effective way to reduce the risk of inaccurate forecasting of extreme load. This paper took peak load as an example, analyzed the multi sub-peaks characteristic of load curve, studied the statistical features of the peak load magnitude and timing, established the regression model between peak load occurrence time and sunset time; based on sorted statistics of peak load daily increments by weeks and seasons, the paper forecasted the probabilistic density functions (PDF) of sub-peak load magnitudes, calculated the PDF of the peak load magnitude via sequence operation theories, forecasted the timing PDF employing total probability formula. Method proposed in this paper has been applied to a city in North China and the results prove the effectiveness of this method.

KEY WORDS: power systems; short term forecasting; bus load forecasting; sequence operation theory; extreme load; occurrence time

摘要: 极值负荷的幅值与出现时刻是决定调度运行计划的重要依据, 而其预测精度往往不尽如人意。实现概率化的预测, 是规避极值负荷预测不准确所带来的风险的有效途径。以最高负荷为例, 深入剖析了极值负荷的多峰特性, 分析了幅值与出现时刻的统计规律, 建立了日落时刻与晚高峰出现时间之间的回归模型; 基于日间极值负荷增量的分类统计, 预测下一天的各个子高峰幅值的概率分布(probabilistic density function, PDF), 运用序列运算理论计算日最高负荷幅值的

概率分布, 最终根据全概率公式, 实现了日最高负荷出现时刻的概率性预测方法。中国北方某城市的实际预测表明, 所提出的概率化预测方法可以有效地解决极值负荷预测问题。

关键词: 电力系统; 短期负荷预测; 母线负荷预测; 序列运算理论; 极值负荷; 出现时刻

0 引言

无论是以全天负荷曲线为预测对象的短期负荷预测, 还是母线负荷预测^[1-3], 其难点往往在于如何准确地把握在负荷波动过程中的一些关键点, 如最高负荷、最低负荷。最高负荷的准确预测, 直接决定次日开机机组容量的大小。如果最高负荷预测偏低, 导致计划开机容量偏低, 则次日运行中将面临备用不足等严重的安全问题; 如果最高负荷预测偏高, 导致计划开机容量偏高, 则多数机组将在负荷率较低的状态运行, 导致全系统运行成本增加。最低负荷预测值与最高负荷预测值结合在一起, 将决定系统的峰谷差, 由此决策系统的调峰难度, 是决定开停机计划的直接依据。

不妨将日最高负荷、日最低负荷的预测统称为“极值负荷预测”。一方面, 可以利用全天负荷曲线预测技术^[4-7], 直接得到极值负荷的预测结果; 另一方面, 还需要研究直接预测极值负荷的特殊方法^[8-9]。

常规的日最高负荷预测结果一般都是确定性的, 仅给出一个确切的数值, 既无法估计该负荷值可能出现的概率, 又无法确定预测结果可能的波动范围, 忽视了预测结果本身的概率特性。实际上,

基金项目: 国家自然科学基金项目(51077077, 50777031)。

Project Supported by National Natural Science Foundation of China (51077077, 50777031).

由于预测问题的超前性，实现不确定性的预测更加符合客观需求，预测结果的概率特性有助于决策者在电网规划、风险分析、可靠性评估等方面更好地把握研究对象的客观规律，实现更为可靠和科学的分析与评估。因此，引入不确定性的分析思想，实现极值负荷概率性的预测，具有重要意义。

与一般的概率性负荷预测^[2]不同，极值负荷的概率密度具有非正态性——多个具有正态分布的随机变量最大值的概率密度是非高斯分布的，因此不能沿用一般的概率密度预测思路，而需要建立从多个随机变量的任意概率密度形成其最大值概率密度的数学方法。

在预测极值负荷概率密度的同时，工作人员往往还关心极值负荷出现的时刻，因此还需解决极值负荷出现时刻的概率性预测。文献[10]就同时研究了峰值负荷的幅值预测和峰值出现时刻的预测问题，只不过该文是提前一年左右的长期预测，仅预测年度最高负荷及其出现的日期，不属于本文所定位的短期(日内)预测范畴。

本文利用序列运算理论^[11-15]建立了极值负荷幅值及其出现时刻的概率性预测方法，根据极值负荷出现时间的时序和统计特性，运用序列运算理论形成了全天最高负荷幅值与出现时刻的概率分布。

为了表述方便，此处以日最高负荷为例进行分析，日最低负荷的预测方法与之类似，仅在关键点进行补充说明。

1 日最高负荷的基本特性分析

1.1 日最高负荷的多峰特性

日负荷曲线具有多个子高峰，各子高峰的幅值与出现时刻的变化决定了最高负荷的变化规律。以中国北方某重要城市(称为 A 市)为例，在图 1 的 96 点典型日负荷曲线中，负荷于第 48 点左右达到第 1 个子高峰(早高峰)，并于第 81 点左右达到第 2 个子高峰(晚高峰)，形成了典型的双峰特性。少数地区

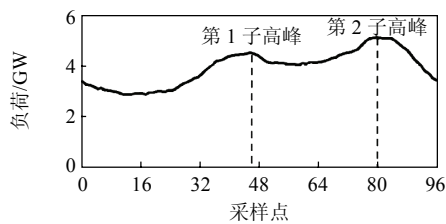


图 1 北方某城市典型日负荷曲线

Fig. 1 Typical load curve of a city in north China

的日负荷曲线呈现 3 峰特征(第 1 子高峰是早高峰，第 2 子高峰是午高峰，第 3 子高峰是晚高峰)。简起见，本文将以双峰的日负荷曲线为基础进行推导，所得结果容易推广至 3 个高峰甚至多个高峰的情形。

1.2 日最高负荷出现时刻的特性分析

日最高负荷的出现时刻具有分群特性。以 A 市为例分析，全年中日最高负荷出现时刻的频数统计如图 2 所示，在 96 点负荷曲线中，日最高负荷集中出现在 2 个时段：其一在第 40 至 50 点之间(午高峰时段)，呈现出一定的正态性；其二在第 60 至 85 点之间(晚高峰时段)，其余时段的频数接近于零。这一特性可由多峰特性解释。

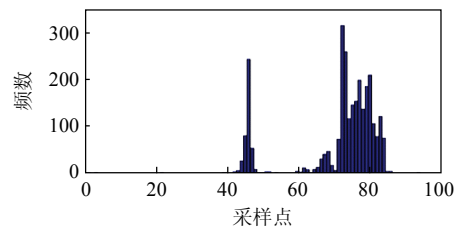


图 2 全年中日最高负荷出现时刻的频数统计图

Fig. 2 Frequency distribution of peak load occurrence time

同时，日最高负荷的出现时刻具有季节性。图 3 为 A 市最高负荷出现时刻的全年变化时序图，横坐标为全年日期序号，纵坐标为当日最高负荷出现时刻(96 点曲线)。图中夏秋季节最高负荷在午高峰和晚高峰时段之间跳变，春冬季节最高负荷出现在晚高峰时段。这表明各子高峰具有不同的季节变化趋势。

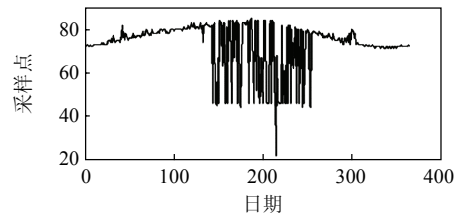


图 3 全年每天日最高负荷出现时刻的分布情况

Fig. 3 Peak load occurrence time during a year

另外，子高峰出现时刻具有规律性：一方面，由图 2 可知，第 1 子高峰出现时刻是平稳的，且具有正态性；另一方面，图 3 最高负荷出现时刻轨迹的上包络线为第 2 子高峰出现时刻轨迹，该轨迹的变化规律与日落时间具有明显的正相关关系。

1.3 日最高负荷幅值变化的特点

一般而言,日最高负荷序列不仅具有长期趋势性,还具有显著的季节特性。仍以A市为例,全年的日最高负荷序列如图4所示,日最高负荷分别在大约第20日(1月)与第220日(7月)达到局部极大值,在130日(4月)与280日(9月)达到局部极小值,由此形成了4个明显的增减阶段;同时,日最高负荷在工作日明显高于休息日,具有显著的周特性。

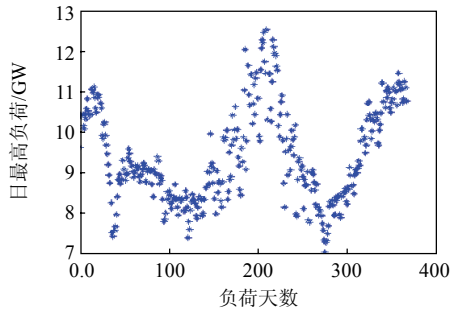


图4 全年的日最高负荷序列

Fig. 4 Altitude of peak load during a year

2 日最高负荷预测的总体思路

各子高峰幅值的变化规律不同,统计特性不同,笼统考虑最高负荷的预测过程难以保证预测结果的精确性。本文最高负荷的预测思路是首先分别预测各个子高峰的幅值及时刻的概率分布,最后依靠序列运算理论形成全天最高负荷的幅值和出现时刻的概率分布。

适用于子高峰的概率化预测较多,本文避免了直接利用负荷幅值的统计结果进行预测^[16-19],并将其解构为“前一日负荷”与“相邻两日负荷增量”之和。由于前一日负荷已知,因此待预测日负荷与负荷增量具有相同的分布特性;同时,负荷增量的分类统计结果可以自动涵盖负荷的周期性与趋势性,通过适当的变换可以形成平稳序列,故可将其分类统计结果用于对日最高负荷的预测中,是简洁有效的概率化预测方法。在时刻预测中,本文对第1子高峰和第2子高峰分别进行了统计和回归分析,得到了各自的概率分布。

全天日最高负荷的预测中,本文充分运用了序列运算理论,通过并积运算由各子高峰的幅值预测得到最高负荷的幅值运算,通过全概率公式以及卷积运算得到最高负荷出现时间的概率分布,从而最终形成最高负荷及其出现时刻的概率化预测。

综上,本文基于各子高峰的分析与预测,提出如下的预测流程:首先在历史负荷曲线中找出每日负荷的各子高峰出现时间及幅值,分别进行统计分析;根据统计信息分别预测各个子高峰幅值的概率密度,并运用序列运算得到日最高负荷幅值的概率密度;根据各子高峰幅值的概率密度预测值,运用序列运算得到日最高负荷出现于各个子高峰的概率分布,并结合各个子高峰出现时刻的概率密度,运用全概率公式最终得到日最高负荷出现时刻的概率分布。预测流程如图5所示。

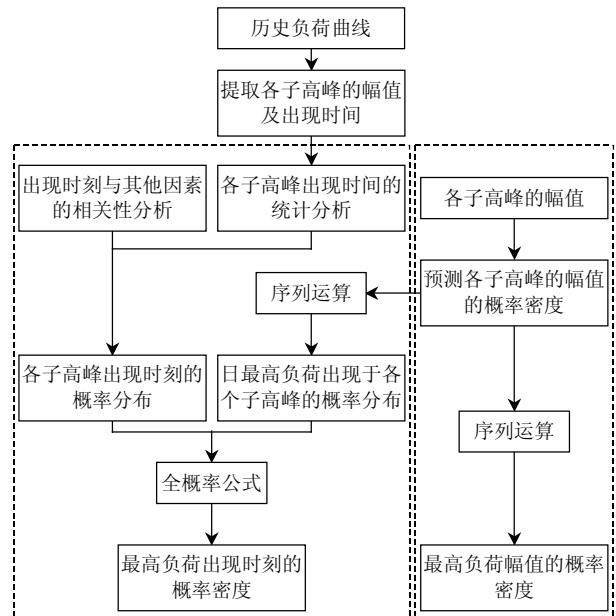


图5 日最高负荷概率化预测的整体流程

Fig. 5 Peak load altitude and timing forecasting procedure

3 日最高负荷幅值的概率性预测

3.1 负荷增量的分类统计

日最高负荷的波动性、趋势性与周期性是幅值预测的主要困难,依靠幅值本身的统计分析难以简洁而准确地反映最高负荷的以上变化规律。本文将待预测日负荷看作基准日负荷与两日间负荷增量之和。基准日的最高负荷已经发生,是确定的,因此最高负荷的分布特性等同于负荷增量的分布特性。基于增量的统计方法可以借助负荷的惯性消除趋势性、凸显周期性,形成简洁统一的概率表达式。假设前 n 天为历史日,第 $n+1$ 天为待预测日,历史负荷向量为

$$L = [l_1 \quad l_2 \quad \dots \quad l_n] \quad (1)$$

定义负荷增量序列为

$$\mathbf{A}=[\Delta_1 \ \Delta_2 \ \cdots \ \Delta_{n-1}] \quad (2)$$

$$\Delta_j=l_{j+1}-l_j, \quad j=1,2,\dots,n-1 \quad (3)$$

取 $n=365$ ，按照季节的不同将 \mathbf{A} 划分为 4 个子序列，剔除节假日因素后对每个子序列按照所属星期类型的不同重组为二维数组。图 6 将全年相邻两日的最高负荷增量进行了分类，4 幅图分别对应于春夏秋冬 4 季，每幅图横坐标表示增量星期类型，纵坐标表示幅值大小。由图 6 可知：夏季数据离散性较大，周五至周六增量为负，周六至周一增量为正。

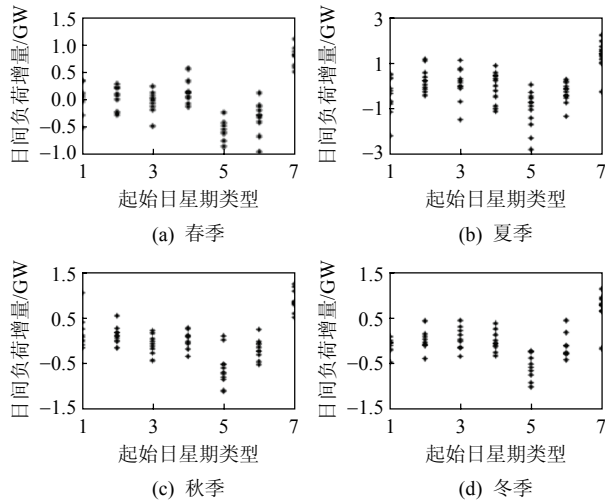


图 6 日间最高负荷增量的分类统计结果

Fig. 6 Sorted statistic of peak load increments between consecutive days

对不同季节、不同星期类型的数据(即以图 6 每一幅图的每一列为对象)分别进行统计分析可得到 7×4 的增量期望矩阵 \mathbf{E} 和增量方差矩阵 \mathbf{S} 。

增量期望矩阵为

$$\mathbf{E} = \begin{bmatrix} \mu_{11} & \mu_{12} & \mu_{13} & \mu_{14} \\ \mu_{21} & \mu_{22} & \mu_{23} & \mu_{24} \\ \vdots & \vdots & \vdots & \vdots \\ \mu_{71} & \mu_{72} & \mu_{73} & \mu_{74} \end{bmatrix} \quad (4)$$

增量方差矩阵为

$$\mathbf{S} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} & \sigma_{24} \\ \vdots & \vdots & \vdots & \vdots \\ \sigma_{71} & \sigma_{72} & \sigma_{73} & \sigma_{74} \end{bmatrix} \quad (5)$$

式中 μ_{ij} 、 σ_{ij} 分别为星期 i 到 $i+1$ 、季节为 j 的相邻 2 天负荷增量的统计均值和方差。增量期望矩阵 \mathbf{E}

和增量方差矩阵 \mathbf{S} 通过其中元素取值的不同体现负荷增量的不同星期类型与季节特性，自动蕴含峰值负荷幅值变化的周期性，简化和统一了概率密度的形成方法。

3.2 各个子高峰幅值的概率性预测

将负荷增量的分类统计方法用于预测各子高峰幅值，分别形成各子高峰增量期望矩阵与增量方差矩阵，估计子高峰概率分布的参数，形成各子高峰幅值的概率性预测。以双峰情况为例，设第 1、第 2 子高峰的历史序列分别为

$$\mathbf{L}'=[l'_1 \ l'_2 \ \cdots \ l'_n] \quad (6)$$

$$\mathbf{L}''=[l''_1 \ l''_2 \ \cdots \ l''_n] \quad (7)$$

$$\mathbf{A}'=[\Delta'_1 \ \Delta'_2 \ \cdots \ \Delta'_{n-1}] \quad (8)$$

$$\mathbf{A}''=[\Delta''_1 \ \Delta''_2 \ \cdots \ \Delta''_{n-1}] \quad (9)$$

分别对 \mathbf{A}' 和 \mathbf{A}'' 按季节与星期进行拆分与重组，得到对应的增量期望矩阵、增量方差矩阵分别为 \mathbf{E}' 、 \mathbf{S}' 、 \mathbf{E}'' 、 \mathbf{S}'' 。设待预测日对应的增量 Δ'_1 、 Δ''_1 满足正态分布，在增量期望矩阵与方差矩阵中提取相应的期望和方差分别为 μ'_{ij} 、 σ'_{ij} 、 μ''_{ij} 、 σ''_{ij} ，则待预测日的第 1、第 2 子高峰幅值的概率密度可以分别表示为

$$p'(x) = \frac{1}{\sqrt{2\pi}\sigma'_{ij}} e^{-\frac{(x-l'_n-\mu'_{ij})^2}{\sigma'^2_{ij}}} \quad (10)$$

$$p''(x) = \frac{1}{\sqrt{2\pi}\sigma''_{ij}} e^{-\frac{(x-l''_n-\mu''_{ij})^2}{\sigma''^2_{ij}}} \quad (11)$$

3.3 日最高负荷幅值的概率性预测

以下考虑日最高负荷的概率性预测。由于日最高负荷是各子高峰取值的最大值，故可借助序列运算理论，由各子高峰的概率密度直接计算得到最高负荷幅值的概率密度函数。首先对日最高负荷可能出现的取值空间进行离散化，令

$$\underline{l} = \min(l'_n + \mu'_{ij} - 3\sigma'_{ij}, l''_n + \mu''_{ij} - 3\sigma''_{ij}) \quad (12)$$

$$\bar{l} = \max(l'_n + \mu'_{ij} + 3\sigma'_{ij}, l''_n + \mu''_{ij} + 3\sigma''_{ij}) \quad (13)$$

则可认为 $[\underline{l} \ \bar{l}]$ 为所有可能的取值空间，将其等分成 k 段，若设第 i 段取值在子区间 $[a_i, b_i]$ 内，则有

$$\begin{cases} a_i = \underline{l} + \frac{\bar{l} - \underline{l}}{k}(i-1) \\ b_i = \underline{l} + \frac{\bar{l} - \underline{l}}{k}i \end{cases} \quad (14)$$

故第1、第2子高峰出现在各子区间的概率序列为

$$P' = [p'_1 \quad p'_2 \quad \dots \quad p'_k] \quad (15)$$

$$P'' = [p''_1 \quad p''_2 \quad \dots \quad p''_k] \quad (16)$$

式中

$$\begin{cases} p'_i = \int_{a_i}^{b_i} p'(x) dx \\ p''_i = \int_{a_i}^{b_i} p''(x) dx \end{cases} \quad (17)$$

日最高负荷在区间 $[a_i, b_i]$ 内, 当且仅当第1、第2子高峰的最大值在该区间内。由序列运算理论中并积运算的性质可知, 2个随机变量的分布序列的并积是这2个随机变量最大值的分布序列, 因此设日最高负荷落在 $[a_i, b_i]$ 的概率值为 $p_m(i)$, 对于多个子高峰的情形, 只需要对这多个子高峰的概率密度进行并积运算即可, 即

$$p_m(i) = P'(i) \odot P''(i) \quad (18)$$

并积运算是计算多个随机变量最值的概率分布的有效方法, 将复杂的非解析表达式依照全概率公式的思想划归成简洁的数学形式, 是解决最高负荷概率化预测的有力工具。

3.4 日最低负荷幅值的概率性预测

多个序列最大值的概率分布可以通过并积运算得出, 而多个序列最小值的概率分布则需借助交积运算得到。

根据交积运算的物理性质, 多个序列的交积, 表示这几个相互独立的随机变量的较小值的分布函数。假设 n 个可能出现最小负荷时段的概率序列为 P'_1, P'_2, \dots, P'_n , 则最小负荷的概率分布 p_l 为

$$p_l(i) = P'_1(i) \otimes P'_2(i) \otimes \dots \otimes P'_n(i) \quad (19)$$

4 日最高负荷出现时刻的概率性预测

4.1 各个子高峰出现时刻的分布特性

本节将以双峰情况为例, 分析假设已知最高负荷出现在某一子高峰的情况下, 最高负荷出现时刻的概率密度函数形式, 即某一子高峰出现时刻的概率分布。根据实际数据分析可知, 不同子高峰出现时刻与所分析的电网所在的地区关系密切。研究表明, 第1子高峰出现时刻与第2子高峰出现时刻的概率密度不同, 前者具有较好的正态性, 而后者与日落时间具有线性相关关系。

以A市的双峰特征为例分析, 第1子高峰出现

时刻序列相对于时间是平稳的, 故可根据历史统计信息得到未来概率密度的预测; 同时, 第1子高峰出现时刻的历史频度统计表现出良好的正态性, 故本节利用正态分布形成第1子高峰出现时刻的概率密度:

$$p'_i(t) = \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{(t-\mu_i)^2}{\sigma_i^2}} \quad (20)$$

式中 μ_i 、 σ_i 分别为历史上第1子高峰出现时刻的均值和统计方差。

同时, 由图3可以看出第2子高峰的出现时刻随日期的增加具有先增大后减小的趋势性, 将第2子高峰的出现时刻与日平均温度、日最高温度、以及日落时间等多个相关因素进行回归分析, 发现第2子高峰的出现时刻与日落时间具有极强的线性相关关系, 两者的散点图如图7所示。

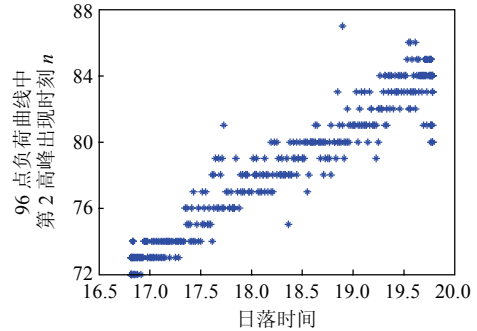


图7 日落时间与第2子高峰出现时刻的回归关系

Fig. 7 Relationship between sunset time and occurrence time of second sub-peak

第2子高峰的出现时刻与日落时间的线性相关关系可表示为

$$t_e = \hat{a} + \hat{b}t_s + e \quad (21)$$

式中: t_e 为第2子高峰的出现时刻; t_s 为当日日落时间; \hat{a} 、 \hat{b} 为拟合系数; e 为拟合误差, 服从正态分布。假设拟合残差的历史向量的均方差为 σ_s , 则第2子高峰出现时刻的概率分布可以表示为

$$p''_i(t) = \frac{1}{\sqrt{2\pi}\sigma_s} e^{-\frac{(t-\hat{a}-\hat{b}t_s)^2}{\sigma_s^2}} \quad (22)$$

由式(22)可得第2子高峰出现时刻的概率密度函数, 而这一概率等价于已知日最高负荷出现在第2子高峰时, 日最高负荷出现时刻的概率密度函数, 是日最高负荷出现时刻的条件概率。

4.2 日最高负荷出现在各个高峰的概率分析

假设最高负荷出现在第 1、第 2 子高峰的概率分别为 p_{11} 、 p_{12} ，则可知

$$p_{11} + p_{12} = 1 \quad (23)$$

同时，根据全概率公式，最高负荷出现在第 2 子高峰的概率等价于所有第 2 子高峰负荷值大于第 1 子高峰负荷值的发生概率之和，即

$$p_{12} = \sum_{i \leq j} p'(i)p''(j), \quad \forall i, j \quad (24)$$

根据序列运算理论中卷积运算的定义，令

$$\bar{P}(i) = P'(i) \Theta P''(i) \quad (25)$$

式(25)的卷积运算所得结果为第 1 与第 2 子高峰幅值状态数之差的概率分布，由卷积运算的物理意义可知，其 0 元素 $\bar{P}(0)$ 则表示第 2 子高峰幅值大于第 1 子高峰幅值的概率，与 p_{12} 意义一致，故有

$$\begin{cases} p_{12} = \bar{P}(0) \\ p_{11} = 1 - p_{12} \end{cases} \quad (26)$$

由此可知，在得知各个子高峰幅值的概率分布后，即可由序列运算理论直接计算出最高负荷出现在各个子高峰的概率分布。

4.3 日最高负荷出现时刻的概率性预测

基于上述分析，可以用全概率公式形成日最高负荷出现时刻的概率性预测。由于日最高负荷必出

现在各个子高峰中的某个高峰，例如：当已知日最高负荷出现在第 2 子高峰时，日最高负荷出现时刻的规律性转化为第 2 子高峰的出现时刻的规律性；反之，则表现出其他高峰出现时刻的规律性。因此，以双峰为例，日最高负荷出现时刻的概率密度表达式为

$$p_i(t) = P_i'(t)p_{11} + P_i''(t)p_{12} \quad (27)$$

即有

$$p_i(t) = \frac{1}{\sqrt{2\pi\sigma_t}} e^{-\frac{(t-\mu_t)^2}{\sigma_t^2}} p_{11} + \frac{1}{\sqrt{2\pi\sigma_s}} e^{-\frac{(t-\hat{a}-\hat{b}t_s)^2}{\sigma_s^2}} p_{12} \quad (28)$$

可见：日最高负荷出现时刻的概率密度首先决定于日最高负荷出现在各个子高峰的概率分布，因此日最高负荷出现时刻受到各子高峰幅值预测的影响；同时，日最高负荷出现时刻的概率分布还受到各子高峰出现时间的影响，体现出了出现时刻的双峰特性，并且随日落时间的缓慢变化而变化。这与前文中的特性是吻合的。

5 算例分析

5.1 负荷增量分段统计结果

对 A 市 2008 年的负荷数据进行分析，形成负荷增量矩阵，以不同季节、不同星期对负荷增量进行重组，分别统计均值和方差，结果如表 1、2 所

表 1 不同星期类型、不同季节负荷增量的均值

Tab. 1 Expectations of peak load increments in different season and week type

MW

星期	早高峰(夏季)	早高峰(秋季)	早高峰(冬季)	晚高峰(春季)	晚高峰(夏季)	晚高峰(秋季)	晚高峰(冬季)	早高峰(春季)
周一	14.730	-373.070	150.370 0	-111.630	-46.201	-85.341	39.423	-117.920
周二	12.850	291.700	112.900 0	-21.300	10.320	297.000	23.950	45.900
周三	-39.955	84.543	-85.815 0	36.199	-41.634	39.171	-32.000	7.028
周四	151.360	-61.092	0.288 3	-29.981	41.190	-111.210	-36.535	-93.742
周五	-535.600	-1 073.600	-629.670 0	-565.160	-240.800	-477.900	-206.760	-298.370
周六	-331.140	-276.810	-197.070 0	-154.470	-132.520	92.281	-132.910	46.458
周日	756.540	1 352.400	857.190 0	719.950	364.120	243.990	477.560	293.370

表 2 不同星期类型、不同季节负荷增量的方差

Tab. 2 Variations of peak load increments in different season and week type

MW

星期	早高峰(春季)	早高峰(夏季)	早高峰(秋季)	早高峰(冬季)	晚高峰(春季)	晚高峰(夏季)	晚高峰(秋季)	晚高峰(冬季)
周一	264.38	798.16	335.22	159.27	214.13	745.08	209.95	159.95
周二	204.74	511.37	180.48	211.22	171.48	319.58	246.87	212.11
周三	201.94	693.31	190.07	236.01	225.81	894.76	159.28	309.33
周四	213.17	675.58	185.61	222.10	162.22	531.55	124.85	141.23
周五	203.07	850.93	391.70	288.99	83.04	461.10	250.91	195.95
周六	304.63	462.81	212.42	264.86	227.68	929.69	273.94	250.37
周日	188.71	626.50	227.97	347.21	188.71	626.50	227.97	347.21

示。从表1可以看出：周五至周六的负荷增量均值为负，表明有较大的下降，而周日至周一的负荷增量显著为正，第2子高峰周五、周日的增量的均值显著小于第1子高峰对应星期类型增量的均值，夏季高峰负荷的方差值要远大于其他季节的方差，表明夏季负荷的波动程度要高于其他季节。

5.2 日最高负荷的幅值预测

通过以上的增量均值、增量方差确定各个子高峰的概率密度函数的参数，形成相应的概率密度预测结果，并通过两者的并积运算，形成日最高负荷幅值的概率密度。图8为A市2008年第200天的极值预测过程，为了便于比较，用竖线画出了当天的实际最高负荷值。由图8可知，虽然2个子高峰概率密度的峰值均低于实际值，但是通过并积运算后，全天最高负荷的概率密度峰值右移，比单一序列的概率密度预测更加贴近实际值。

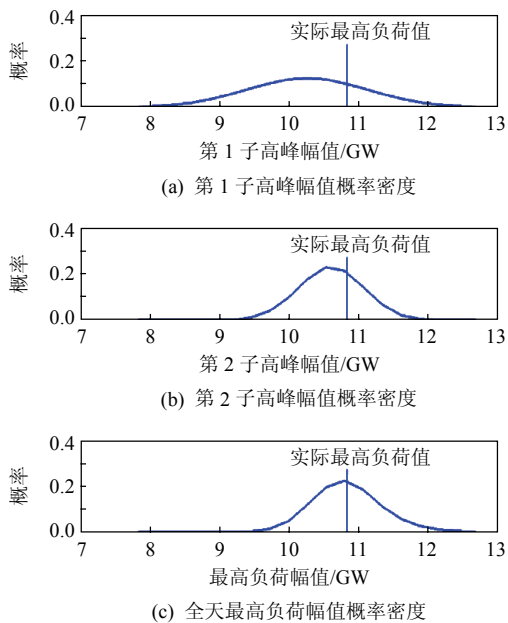


图8 各个高峰概率密度的并积过程

Fig. 8 Sequence operation process of different load peaks

以该年第281天至年末的预测结果为例，表3给出了不同置信区间下，理论置信度与实际预测命中率的对比关系。可以看到：随着区间宽度的减小，置信度逐渐减小，实际值出现在预测区间内的统计频率也逐渐减小。

图9为理论置信度与实际命中频率之间的关系，图中各点越接近于直线 $y=x$ ，则表明理论值与实际预测值吻合越好。由于实际预测过程中会受到干扰因素的影响，因此图中各点均位于直线下方。

表3 理论置信度与实际预测命中率的对比

Tab. 3 Probability and frequency comparison

区间宽度	区间宽度相对值	在区间内部的频率	置信度
1192.30	0.1041	0.9882	0.9943
1052.00	0.0918	0.9529	0.9868
911.77	0.0796	0.9058	0.9695
771.49	0.0673	0.8823	0.9345
631.22	0.0551	0.7647	0.8711
490.95	0.0428	0.6235	0.7670
350.68	0.0306	0.4941	0.6105
210.41	0.0183	0.2705	0.3974

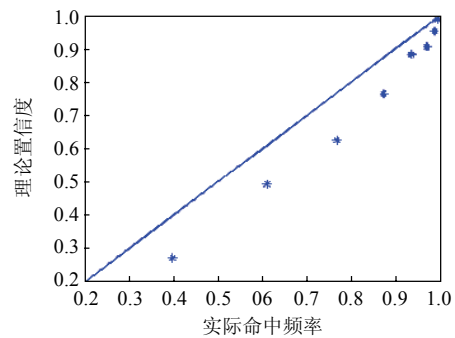


图9 理论置信度与实际预测准确频率关系图

Fig. 9 Probability and frequency comparison

由图9可知，预测结果在理论置信度接近于1，即预测区间较宽时，理论值与实际值更为接近，而当预测区间收紧时，实际值偏低。这是因为预测区间越窄，对随机干扰越灵敏，因而与理论值相差较大。

5.3 日最高负荷时刻预测

对2008年A市第1子高峰出现时刻进行统计分析，得到均值为46.75(即在96点负荷曲线中，第1子高峰平均出现在第46点与47点之间)，方差为1.03，频数统计如图2所示。

通过式(21)对该地区日落时间与第2子高峰出现时间进行回归分析，得到回归方程：

$$t_e = 7.7096 + 0.0644t_s \quad (29)$$

式中 t_s 、 t_e 分别为日落时间和晚高峰出现时刻。回归得到决定系数为0.75，拟合方差为1.45，印证了4.1节中日落时间与第2子高峰出现时刻的强相关关系。

根据子高峰幅值及出现时刻的预测结果，通过式(25)~(28)形成全天最高负荷出现时刻的概率分布。该市1a的预测结果如图10所示。横坐标表示1d负荷的96点，纵坐标表示1a中的日期，Z坐标

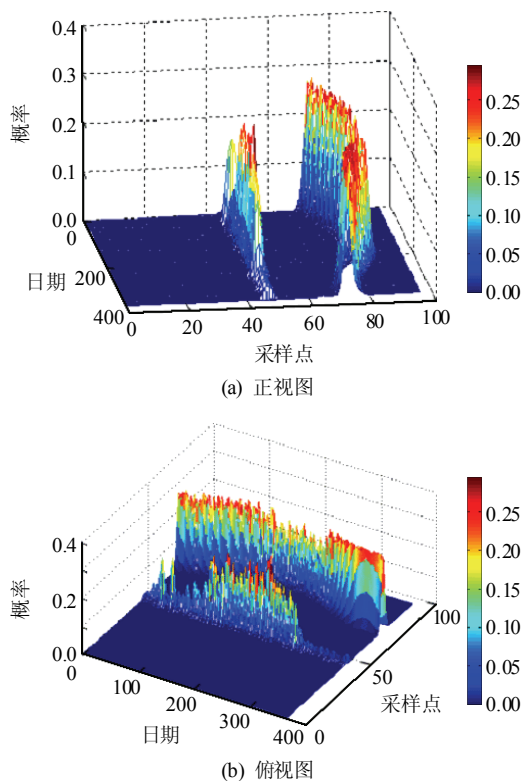


图 10 全年中日最高负荷出现时刻概率密度预测结果

Fig. 10 Peak load timing forecasting results of a year

表示该日最高负荷出现在该时刻的概率密度。由图 10 可知：固定日期坐标，即可得到某一天最高负荷出现时刻的概率密度，左侧图表明了概率密度集中形成了 2 个尖峰；同时，右侧图表明概率密度具有明显的季节特性，与 1.2 节的特性是吻合的。

6 结论

本文提出了基于序列运算理论的极值负荷及其出现时刻的概率性预测方法，并以日最高负荷为例进行了分析。

首先阐述了高峰负荷的分群特性，分别预测各个子高峰的概率密度，然后根据各个子高峰的概率密度序列，运用序列运算理论形成日最高负荷出现时刻与幅值概率密度。在幅值预测中，本文按季节与星期类别统计了相邻两日间极值负荷增量，并运用负荷的增量对下一天的各个子高峰进行概率性预测。在时刻预测中，本文揭示了晚高峰出现时刻与日落时间的强相关关系以及其他子高峰出现时间的正态性。最终根据全概率公式形成了日最高负荷出现时刻的概率性预测。

各个子高峰概率密度的准确预测是算法的基础，其预测精度决定了最终的概率性预测精度。本

文尚未考虑气象因素的影响，而气象因素的影响对极值负荷的影响是十分显著的，这还有待后续工作进行完善。

参考文献

- [1] 康重庆, 赵燃, 陈新宇, 等. 多级负荷预测的基础问题分析[J]. 电力系统保护与控制, 2009, 37(9): 1-7.
Kang Chongqing, Zhao Ran, Chen Xinyu, et al. Fundamental analysis of multilevel load forecasting[J]. Power System Protection and Control, 2009, 37(9): 1-7(in Chinese).
- [2] 陈新宇, 康重庆, 陈刚, 等. 规避坏数据影响的母线负荷预测新策略[J]. 中国电力, 2009, 42(9): 27-31.
Chen Xinyu, Kang Chongqing, Chen Gang, et al. Novel strategy for bus load forecast by preventing the effect of bad data[J]. Electric Power, 2009, 42(9): 27-31(in Chinese).
- [3] 赵燃, 陈新宇, 陈刚, 等. 母线负荷预测中的自适应预测技术及其实现[J]. 电网技术, 2009, 33(19): 55-59.
Zhao Ran, Chen Xinyu, Chen Gang, et al. Adaptive forecasting approach and its implementation in bus load forecasting[J]. Power System Technology, 2009, 33(19): 55-59(in Chinese).
- [4] Huang S J, Shih K R. Short-term load forecasting via ARMA model identification including non-Gaussian process considerations[J]. IEEE Trans. on Power Systems, 2003, 18(2): 673-679.
- [5] 陈志业, 牛东晓, 张英怀, 等. 电网短期电力负荷预测系统的研究[J]. 中国电机工程学报, 1995, 15(1): 30-35.
Chen Zhiye, Niu Dongxiao, Zhang Yinghuai, et al. Short-term load forecast duplication rate forecast grey forecast parameters rectifying[J]. Proceedings of the CSEE, 1995, 15(1): 30-35(in Chinese).
- [6] 杜杰, 陆金桂, 曹一家, 等. 短期电力负荷预报间隔采样混沌模型[J]. 中国电机工程学报, 2006, 26(10): 28-32.
Du Jie, Lu Jingui, Cao Yijia, et al. The interval sampling chaotic model of electric load forecasting[J]. Proceedings of the CSEE, 2006, 26(10): 28-32(in Chinese).
- [7] 雷绍兰, 孙才新, 周淙, 等. 电力短期负荷的多变量时间序列线性回归预测方法研究[J]. 中国电机工程学报, 2006, 26(2): 25-29.
Lei Shaolan, Sun Caixin, Zhou Quan, et al. The research of local linear model of short-term electrical load on multivariate time series[J]. Proceedings of the CSEE, 2006, 26(2): 25-29(in Chinese).
- [8] 卢建昌, 韩红领. 基于灰色神经网络组合模型的日最高负荷预测[J]. 华东电力, 2008, 36(2): 60-66.
Lu Jianchang, Han Hongling. Daily peak load forecast

- based on combined models of gray systems and neural networks[J]. East China Electric Power, 2008, 36(2): 60-66(in Chinese).
- [9] Ramanathan R, Engle R F, Granger C W J, et al. Short-run forecasts of electricity loads and peaks[J]. International Journal of Forecasting, 1997(13): 161-174.
- [10] McSharry P E, Bouwman S, Bloemhof G. Probabilistic forecast of the magnitude and timing of peak electricity demand[J]. IEEE Trans. on Power Systems, 2005, 20(2): 1166-1172.
- [11] 康重庆, 夏清, 相年德, 等. 序列运算理论及其应用[M]. 北京: 清华大学出版社, 2003: 1-61.
Kang Chongqing, Xia Qing, Xiang Niande, et al. Sequence operation theory and its application[M]. Beijing: Tsinghua University Press, 2003: 1-61(in Chinese).
- [12] 徐玮, 康重庆, 夏清. 序列运算的算法复杂性分析[J]. 中国电机工程学报, 2009, 29(28): 102-106.
Xu Wei, Kang Chongqing, Xia Qing, et al. Algorithm complexity analysis of sequence operation theory[J]. Proceedings of the CSEE, 2009, 29(28): 102-106(in Chinese).
- [13] 徐玮, 夏清, 康重庆. 基于序列运算理论的购售电主体市场风险评估[J]. 电力系统自动化, 2008, 32(4): 15-19, 97.
Xu Wei, Xia Qing, Kang Chongqing. Risk assessment of electricity buyers and sellers in power markets based on sequence operation theory[J]. Automation of Electric Power Systems, 2008, 32(4): 15-19, 97(in Chinese).
- [14] 徐玮, 康重庆, 夏清, 等. 序列运算离散化过程中的误差成因及补偿[J]. 清华大学学报: 自然科学版, 2009, 49(1): 17-20.
Xu Wei, Kang Chongqing, Xia Qing, et al. Sequence operation discretization error mechanism and compensation[J]. Journal of Tsinghua University: Sci. & Tech., 2009, 49(1): 17-20(in Chinese).
- [15] Kang Chongqing, Xia Qing, Xiang Niande. Sequence operation theory and its application in power system reliability evaluation[J]. Reliability Engineering & System Safety, 2002, 78(2): 101-109.
- [16] 杨文佳, 康重庆, 夏清, 等. 基于预测误差分布特性统计的概率性短期负荷预测[J]. 电力系统自动化, 2006, 30(19): 47-52.
Yang Wenjia, Kang Chongqing, Xia Qing, et al. Short term probabilistic load forecasting based on statistics of probability distribution of forecasting errors[J]. Automation of Electric Power Systems, 2006, 30(19): 47-52(in Chinese).
- [17] 周建中, 张亚超, 李清清, 等. 基于动态自适应径向基函数网络的概率性短期负荷预测[J]. 电网技术, 2010, 34(3): 37-41.
Zhou Jianzhong, Zhang Yachao, Li Qingqing, et al. Probabilistic short-term load forecasting based on dynamic self-adaptive radial basis function network[J]. Power System Technology, 2010, 34(3): 37-41(in Chinese).
- [18] 方仍存, 周建中, 张勇传, 等. 短期负荷概率性预测的混沌时间序列方法[J]. 华中科技大学学报: 自然科学版, 2009, 37(5): 125-128.
Fang Rengcun, Zhou Jianzhong, Zhang Yongchuan, et al. Short-term probabilistic load forecasting using chaotic time series[J]. Journal of Huazhong University of Science and Technology: Nature Science Edition, 2009, 37(5): 125-128(in Chinese).
- [19] 穆钢, 侯凯元, 杨右虹, 等. 负荷预报中负荷规律性评价方法的研究[J]. 中国电机工程学报, 2001, 21(10): 96-101.
Mu Gang, Hou Kaiyuan, Yang Youhong, et al. Studies on load regularity evaluating method for load forecasting[J]. Proceedings of the CSEE, 2001, 21(10): 96-101(in Chinese).



陈新宇

收稿日期: 2011-01-18。

作者简介:

陈新宇(1986), 男, 博士研究生, IEEE 学生会员, 研究方向为负荷预测、智能电网等, chenxy05@mails.tsinghua.edu.cn;

康重庆(1969), 男, 教授, IEEE、CSEE 高级会员, 研究方向为低碳电力、负荷预测、电力规划等, cqkang@tsinghua.edu.cn;

陈敏杰(1988), 男, 博士研究生, 研究方向为电力系统决策问题、电力电子电路、大功率电力电子器件以及可再生能源电力电子设备等, minjie@mit.edu。

(责任编辑 刘浩芳)