

## 基于混合自回归滑动平均潜周期模型的短期电价预测

曾勇红, 王锡凡, 冯宗建

(西安交通大学电气工程学院, 710049, 西安)

**摘要:** 应用混合自回归滑动平均潜周期模型对短期电价序列进行了预测. 对消除了趋势影响的电价序列, 经离散傅里叶变换转换为复值潜周期模型, 采用一种简单的周期图检测方法计算电价序列的周期特征参数. 为了计及历史信息对当前状态的影响, 采用自回归滑动平均模型拟合残差随机分量, 采用赤池信息准则确定模型的阶数, 参数则由矩估计得到. 该模型不要求预先假设电价序列的周期尺度, 周期的个数和大小由模型计算确定, 方法简单. 采用美国宾夕法尼亚、新泽西、马里兰电力市场的实际电价数据对模型进行了检验, 验证了模型的有效性.

**关键词:** 潜周期; 电价预测; 自回归滑动平均模型

**中图分类号:** TM73 **文献标志码:** A **文章编号:** 0253-987X(2008)02-0184-05

### Short Term Forecast of Electricity Price Based on Mixed Autoregressive Moving Average and Hidden Periodic Model

ZENG Yonghong, WANG Xifan, FENG Zongjian

(School of Electrical Engineering, Xi'an Jiaotong University, Xi'an 710049, China)

**Abstract:** The mixed autoregressive moving average (ARMA) and hidden periodicity model is chosen to predict a short term series of electricity price. Eliminating the trend influence, a complex hidden periodicity model of price sequences can be obtained with discrete Fourier transformation. The periodicity parameters of price sequences are calculated by a simple detection for periodogram. To take consider the impact of historical information on the present states into account, the ARMA model is used to fit the residual stochastic component. The Akaike's information criterion is employed to determine the number of order in autoregressive moving average model, whose parameters are estimated by moment approach, where the pre-assumption on periodicity scale gets unnecessary. The number and size of periodicities is only derived from a simple computation model. The model proposed is verified by actual data of electricity price in PJM power market in USA.

**Keywords:** hidden periodicity; electricity price forecast; autoregressive moving average model

电价的波动除受燃料价格、竞价机组可用容量、水力发电量、输电阻塞等电力系统特有约束的影响外, 还受到电力市场体制结构、发电商实施市场力等主客观因素影响. 因此, 电价预测相对负荷预测难度要大, 采用传统的时间序列分析或者神经网络方法

均存在一定的不足<sup>[1]</sup>.

相比于长期电价, 短期电价具有更大的波动性和随机性. 在分析短期电价时, 一般均假设给定的时间序列为一非平稳随机过程, 并且表现出以天、周、年为周期波动, 且大周期中嵌套小周期的特点, 从而

适合采用季节性求和自回归滑动平均(ARIMA)或其变型模型进行预测<sup>[2-3]</sup>。鉴于小波变换具有良好的时频局部化性质和变焦性质,能将原始电价序列分解为不同周期的块信号,再利用 ARIMA 对不同的子信号分别建模进行预测,能提高预测精度<sup>[4]</sup>。总体而言,考虑电价周期性通常会引起建模的复杂性,因此文献<sup>[5]</sup>考虑分时段电价预测,这样形成多个消除了周期性影响的电价子序列,并能计及周末和工作日电价变化的差异。以上这些研究的共同点是剔除非平稳序列中的趋势项和周期项,然后对提取的平稳系列进行相关分析。由于忽视了对电价真实周期的检验,直接根据经验设定,可能会导致预测的失效。

本文应用混合自回归滑动平均(ARMA)潜周期模型对短期电价序列进行了分析。该模型直接考虑非平稳电价序列的周期项和随机项,周期的个数和大小由模型计算确定,并用美国宾夕法尼亚、新泽西、马里兰(PJM)电力市场公布的电价数据对模型的有效性进行了验证。

## 1 混合 ARMA 潜周期模型

### 1.1 模型的建立

考虑短期电价序列 $\{x_t\}$ 由确定性分量和随机过程组成,其混合 ARMA 潜周期模型为

$$x_t = \sum_{j=1}^p a_j x_{t-j} + \sum_{j=1}^k A_j \cos(\omega_j t + \varphi_j) + \sum_{j=0}^q b_j \varepsilon_{t-j} \quad t \in N_+ \quad (1)$$

式中: $a_j, b_j \in R$  为模型参数; $0 \leq \omega_1 < \omega_2 < \dots < \omega_k \leq \pi$  为角频率; $A_j$  为振幅; $\varphi_j \in [0, 2\pi)$  为初相位; $\varepsilon_t$  为白噪声过程<sup>[6-7]</sup>。显然,上述模型由自回归项、周期项和随机项组成,形成的混合 ARMA 潜周期模型不仅考虑了多个频率成分的叠加,还考虑了历史信息对当前状态的影响。当  $p=q=0$  时,模型称为潜周期模型,周期叠加项  $y_t = \sum_{j=1}^k A_j \cos(\omega_j t + \varphi_j)$ , 每一项的周期  $T_j = 2\pi/\omega_j$ , 当  $k=0$  时模型为 ARMA 模型。

上述混合 ARMA 潜周期模型为实值模型,其实值模型定义如下

$$x_t = \sum_{j=1}^p a_j x_{t-j} + \sum_{j=1}^r \alpha_j \exp(it\lambda_j) + \sum_{j=0}^q b_j \varepsilon_{t-j} \quad t \in N_+ \quad (2)$$

其中

$$r = 2k \quad (3)$$

$$\lambda_j = \begin{cases} \omega_j & j = 1, 2, \dots, k \\ -\omega_{j-k} & j = k+1, \dots, r \end{cases} \quad (4)$$

$$\alpha_j = \begin{cases} \frac{1}{2} A_j \exp(i\varphi_j) & j = 1, 2, \dots, k \\ \frac{1}{2} A_{j-k} \exp(-i\varphi_{j-k}) & j = k+1, \dots, r \end{cases} \quad (5)$$

此时周期叠加项  $y_t = \sum_{j=1}^r \alpha_j \exp(it\lambda_j)$ 。

在应用中,要求上述复值混合 ARMA 潜周期模型满足以下 2 个条件:①实向量  $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_r)^T$  满足  $-\pi < \lambda_1 < \lambda_2 < \dots < \lambda_r \leq \pi$ ;②复向量  $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_r)^T$  中没有 0 分量,即  $\prod_{j=1}^r \alpha_j \neq 0$ 。

对比实值与复值混合 ARMA 潜周期模型可以看出,当  $r \geq 1$  时, $\lambda$  和  $\alpha$  分别是 $\{x_t\}$ 的角频率和振幅向量,周期  $T_j = 2\pi/\lambda_j$ 。当  $0 \leq r < 1$  时,取  $r=0$ ,此时  $x_t$  为平稳随机过程,用 ARMA 模型描述。

对于非平稳时间序列,其确定性分量中通常还包含有反映社会经济状况等因素的趋势项,表现在历史数据上基本分为 3 种情况:①一段时间内保持相对稳定;②在某一均值附近发生交替波动;③从某一稳定值过渡到另一稳定值。因此,日平均电价的趋势项可以采用分时段平均得以消除。在建立时间序列的潜周期模型之前需要消除趋势项的影响,通常采用差分运算能满足精度要求。定义一阶差分算子  $\nabla^1 x_t = \nabla x_t = x_t - x_{t-1}$ , 则  $d$  阶差分算子为

$$\nabla^d x_t = \overbrace{\nabla \{ \nabla [ \dots ( \nabla x_t ) ] \}}^d \quad (6)$$

可以证明,一阶差分序列可以消除线性增长趋势,二阶差分可以消除二次增长趋势。一般增长趋势为  $d \leq 3$ , 当三阶差分仍不能消除增长趋势的影响时,通常需要采用分段趋势项或其他的方式消除趋势项的影响。假设从历史数据序列可以粗略估计出数据的周期为  $s$ , 样本容量为  $N$ , 则周期个数  $M$  的近似估计为

$$M = [N/s]$$

式中: $[\cdot]$  表示取整运算。

定义趋势项  $T$  为周期平均值,即

$$T_i = \frac{1}{s} \sum_{j=1}^s x_{(i-1)s+j} \quad i = 1, 2, \dots, M \quad (7)$$

采用分段趋势消除的计算式为

$$x'_t = x_t - T_{[t/s]} \quad t = 1, 2, \dots, N \quad (8)$$

式中:  $x'_t$  为消除趋势项影响的变换序列。

## 1.2 混合 ARMA 模型参数估计

1.2.1 周期叠加项的参数估计 因为统计上的方便, 讨论复值潜周期模型比实值模型更加有利, 即讨论参数  $r, \lambda, \alpha$  的估计问题. 该估计分成两个阶段, 首先用格点法进行参数的粗略估计, 其次对粗略估计的值采用二次分析法进行精确计算。

(1) 计算  $r$  和  $\lambda$  的估计值. 令

$$S_N(\lambda) = \sum_{t=1}^N x_t \exp(-i\lambda t)$$

则  $|S_N(\lambda)|$  峰群的个数即为模型的潜周期个数  $r$ . 由于逐一检测  $|S_N(\lambda)|$  的峰群个数相当费时, 在实际计算中,  $r$  和  $\lambda$  可以采用区间划分法来粗估计, 具体算法如下。

取  $A > 0$ , 使得  $A = O(N^{0.75})$ ,  $N \rightarrow \infty$ . 定义  $\mu_j = \frac{\pi}{2N}j - \pi$ , 计算  $d_j = |S_N(\mu_j)|$ ,  $j = 1, 2, \dots, 4N$ , 即

$$d_j = \left\{ \left[ \sum_{t=1}^N x_t \cos t \left( \pi - \frac{\pi}{2N}j \right) \right]^2 + \left[ \sum_{t=1}^N x_t \sin t \left( \pi - \frac{\pi}{2N}j \right) \right]^2 \right\}^{\frac{1}{2}} \quad j = 1, 2, \dots, 4N \quad (9)$$

求  $d_{j_1} = \max\{d_j\}$ ,  $j = 1, 2, \dots, 4N$ . 若  $d_{j_1} \leq A$ , 则定义  $\hat{r} = 0$ , 停止计算, 否则定义

$$\Omega(1) = \{\lambda \in (-\pi, \pi] : N^{-1/2} \leq |\lambda - \mu_{j_1}| \leq 2\pi - N^{-1/2}\}$$

在  $\Omega(1)$  中计算  $d_j$  的最大值  $d_{j_2}$ , 当  $d_{j_2} \leq A$ , 定义  $\hat{r} = 1$ ,  $\hat{\lambda}_1 = \mu_{j_1}$ , 停止计算. 依此类推, 当  $d_j$  在  $\Omega(l) = \{\lambda \in (-\pi, \pi] : N^{-1/2} \leq |\lambda - \mu_{j_l}| \leq 2\pi - N^{-1/2}\}$  中的最大值  $d_{j_l} \leq A$  时, 定义  $r$  的估计为  $\hat{r} = p$ , 并将最大值点  $j_1, j_2, \dots, j_p$  按从小到大重新排列, 得到角频率向量  $\lambda$  的初估计  $\hat{\lambda} = (\hat{\lambda}_1, \hat{\lambda}_2, \dots, \hat{\lambda}_{r(N)})^T$ .

(2) 计算  $\alpha$  的估计值

$$\hat{\alpha}_j = \frac{1}{N} \sum_{t=1}^N x_t \exp(-i\hat{\lambda}_j t) \quad 1 \leq j \leq \hat{r} \quad (10)$$

这样估计的  $\hat{\alpha}_j$  在理论上不能保证有很好的估计精度, 故还需要进行精估计。

用  $\arg(c)$  表示复数  $c$  的幅角. 取正整数  $H > 1$ , 则存在正整数  $m$ , 满足

$$N = mH + H' \quad 0 \leq H' < H$$

对于  $s = 1, 2, \dots, H$ , 定义

$$T_j(s) = \sum_{t=(s-1)m+1}^{sm} x_t \exp(-i\hat{\lambda}_j t) \quad (11)$$

$$\theta_j = \arg \left[ \sum_{s=1}^H T_j(s) / |T_j(s)| \right] \quad (12)$$

$$z_j(s) = \arg[T_j(s) \exp(-i\theta_j)] \quad (13)$$

$$\beta_j = \sum_{s=1}^H z_j(s) \left( s - \frac{H+1}{2} \right) / \sum_{s=1}^H \left( s - \frac{H+1}{2} \right)^2 \quad (14)$$

则由  $\tilde{\lambda}_j = \hat{\lambda}_j + \beta_j/m$  得改进的  $\lambda_j$ , 从而可得到改进的  $\alpha_j$  为

$$\tilde{\alpha}_j = \frac{1}{N} \sum_{t=1}^N x_t \exp(-i\tilde{\lambda}_j t) \quad 1 \leq j \leq \hat{r} \quad (15)$$

由复值模型的参数估计可以得到实值模型的参数估计为

$$\left. \begin{aligned} \hat{A}_k &= \hat{\alpha}_k; \hat{\varphi}_1 = 0; \hat{\omega}_1 = \pi & \hat{\lambda}_k &= \pi \\ \hat{A}_j &= 2|\hat{\alpha}_j|; \hat{\varphi}_j = \arg(\alpha_j) & \hat{\omega}_j &= \hat{\lambda}_j & \hat{\lambda}_k &\in (0, \pi) \end{aligned} \right\} \quad (16)$$

1.2.2 ARMA 模型阶数的确定 ARMA 模型中的阶数  $p$  和  $q$  采用信息量准则(AIC)来确定. 考虑 ARMA 模型, 根据电价数据特性能够判断  $p$  的上界  $P_0$  和  $q$  的上界  $Q_0$ , 则对于每一对  $(k, j)$ ,  $0 \leq k \leq P_0$ ,  $0 \leq j \leq Q_0$ , 其 AIC 函数定义为

$$A_{IC}(k, j) = \ln(\hat{\sigma}^2(k, j)) + 2(k + j)/N \quad (17)$$

式中:  $\hat{\sigma}^2$  表示白噪声方差  $\sigma^2$  的估计量. 计算  $A_{IC}(k, j)$  的最小值点  $(\hat{p}, \hat{q})$  (当不惟一时应先取  $k+j$  最小的, 然后取  $j$  最小的) 即为 ARMA 模型的阶数。

当确定了模型的阶数后, 可以采用矩法、最小二乘法或极大似然法等来确定模型中的未知参数  $a_j, b_j$ . 与其他方法相比, 矩法显得较为粗糙, 但具有简单易懂的优点, 且在许多情况下与其他精估计接近, 所以得到了广泛应用, 本文也采用了矩法估计模型的参数, 具体步骤可参考文献[8].

1.2.3 混合模型的耦合机制 混合 ARMA 潜周期模型采用了误差预估-校正来实现模型的耦合. 假设已知潜周期个数和角频率的估计  $\hat{r}$  和  $\hat{\lambda}$ , 令

$$B(j) = \frac{1}{N} \sum_{t=1}^N x_t \exp(-i\hat{\lambda}_j t) \quad j = 1, 2, \dots, \hat{r} \quad (18)$$

$$z_t = x_t - \sum_{j=1}^{\hat{r}} B(j) \exp(i\hat{\lambda}_j t) \quad t = 1, 2, \dots, N \quad (19)$$

混合模型对实值序列  $z_t$  建立 ARMA 模型. 当获得 ARMA 模型的参数  $a_j$  的估计参数为  $\hat{a}_j$  后, 对潜周期模型的振幅校正为

$$\hat{\beta}_j = \frac{1}{N - \hat{p}_{t=p+1}} \sum_{j=1}^{\hat{p}} (x_t - \sum_{j=1}^{\hat{p}} \hat{a}_j x_{t-j}) \exp(-i\hat{\lambda}_j t)$$

$$1 \leq j \leq \hat{r} \quad (20)$$

混合 ARMA 潜周期模型的计算流程如图 1 所示。

## 2 算例分析

为了检验本文方法的有效性,对美国 PJM 电力市场进行了日前电价预测. 该市场经过近 20 年的改革,已经成功地建立了一整套的电力市场规则,运营良好. 采用的历史电价数据从 PJM 的信息公布网站 (<http://www.pjm.com>) 获得,从中挑选 Pnode1 的日前市场出清价进行计算. 图 2 显示了该节点处 2006 年全年的日平均价格,从中可以看出,日平均电价是一个剧烈波动的随机过程,除了极少时间存在价格飙升外,具有明显的周期性. 注意到日电价的变化表现出前述趋势项的基本特征,因此日平均电价的趋势项可以采用分时段平均得以消除.



图1 混合 ARMA 潜周期模型的计算流程图

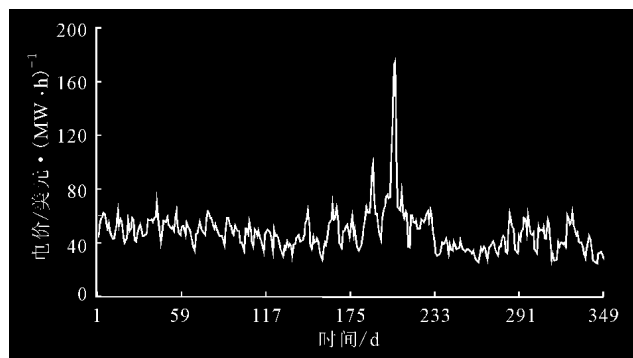


图2 PJM电力市场 Pnode1 的 2006 年日出清价

选取 2006 年 9 月 1 日到 12 月 31 日的历史日平均电价对模型进行了检验. 该时段的电价没有出

表1 潜周期模型的参数估计值

| $j$ | $\omega_j$ | $T_j$    | $\varphi_j$ | $A_j$   |
|-----|------------|----------|-------------|---------|
| 1   | 0.122 5    | 51.279 6 | 3.073 8     | 5.465 3 |
| 2   | 0.159 8    | 39.323 0 | 6.283 2     | 3.518 5 |
| 3   | 0.427 8    | 14.686 5 | 2.789 6     | 3.818 9 |
| 4   | 0.937 4    | 6.702 5  | 6.283 2     | 1.798 6 |
| 5   | 1.792 9    | 3.504 6  | 6.283 2     | 2.940 9 |
| 6   | 2.706 1    | 2.321 8  | 0.976 1     | 1.343 4 |

现价格飙升的情况,趋势性和周期性表现明显,具有较强的代表性. 从表 1 中的周期项  $T_j$  可以看出,该序列基本上保留了电价序列以星期和月为周期变动的信息,也揭示了序列中小的周期变化影响.

由于粗估计只能在格点上得到估计值,所以对不同的  $N$ ,估计值可能是一样的,这从统计上具有优势. 一般的统计模型需要大样本数据才能得到相应精度的估计,而潜周期模型的这一特性显示出该模型在小样本数据统计中的应用前景.

为了对比,分别给出了用潜周期模型和混合 ARMA 潜周期模型预测的结果,其中 ARMA 按照 AIC 准则确定的模型阶数分别为  $p=1$ 、 $q=2$ ,其真实值和预测值的比较如图 3 所示,误差情况则由图 4 和表 2 给出,其中绝对百分比误差  $E$  表示预测值偏离实际值的相对大小

$$E = \frac{|\text{预测值} - \text{实际值}|}{\text{实际值}} \times 100\% \quad (21)$$

表2 两种模型预测误差指标

| 模型          | $E$ 的平均值/% | $E$ 的最大值/% |
|-------------|------------|------------|
| 潜周期         | 10.79      | 65.14      |
| 混合 ARMA 潜周期 | 7.83       | 33.58      |

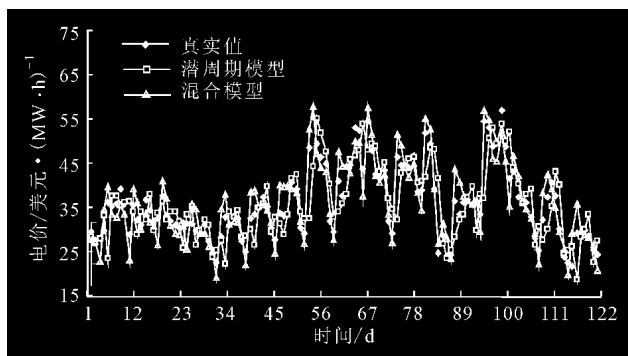


图3 实际值、潜周期模型和混合 ARMA 潜周期模型预测值

从图4和表2可以看出,混合ARMA潜周期模型的预测效果优于基本的潜周期模型.在电价波动平稳的时段,两种模型的预测效果都较好,而在波动较大的时段,潜周期模型的预测效果劣于混合ARMA潜周期模型,个别时段相对误差甚至超过了60%.这说明基本的潜周期模型对缺乏周期性的数据的拟合效果并不理想,而混合ARMA潜周期模型由于考虑了历史信息对现状的影响,提高了模型的预测精度.

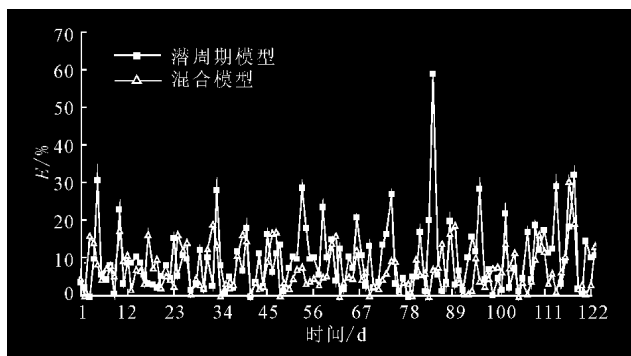


图4 潜周期模型以及混合模型的预测误差

值得注意的是,按照文献[2]对西班牙和加利福尼亚两个电力市场的电价预测的总结来看,如果采用推荐的10%~11%的绝对百分比误差来考察预测的效果,潜周期模型和混合潜周期模型均可以认为是好的短期电价预测模型.由此可见,直接综合考虑随机序列中的周期分量和随机分量可能比仅仅考虑随机分量的其他模型更加合理.

### 3 结 论

由于受众多因素的影响,电价序列本质上是一个非平稳随机过程,表现出很强的周期性特征.在提取了趋势项和周期项分量后,基于随机过程分析的模型直接处理电价序列中的随机变动分量,不可避免地忽视了周期项和随机项的相互影响作用.对消除了趋势影响的电价序列,直接采用潜周期模型可以同时考虑电价序列的周期分量和随机分量,方法简单,但忽略了历史信息对现状的影响.混合ARMA潜周期模型则结合了两者的优点.采用美国PJM电力市场的实际电价数据对模型进行了检验,

并与基本潜周期模型的预测效果进行了对比,验证了模型的有效性.

从原理上看,本文提出的方法能够方便地推广到小时或年电价预测,而且由于负荷通常比电价呈现出更好的周期特性,将本文的方法用于负荷预测可能会有更好的效果.

### 参考文献:

- [1] 陈思杰,周浩. 电力市场电价预测方法综述[J]. 继电器, 2006, 34(11): 54-60.  
CHEN Sijie, ZHOU Hao. Electricity price forecast methods of electricity market [J]. Relay, 2006, 34 (11): 54-60.
- [2] CONTRERAS J, ESPINOLA R, NOGALES F J, et al. ARIMA models to predict next-day electricity prices [J]. IEEE Trans Power Systems, 2003, 18(3): 1014-1020.
- [3] 周明,严正,倪以信,等. 含误差预测校正的ARIMA电价预测新方法[J]. 中国电机工程学报, 2004, 24(12): 63-68.  
ZHOU Ming, YAN Zheng, NI Yixin, et al. A novel ARIMA approach on electricity price forecasting with the improvement of predicted error [J]. Proceedings of the CSEE, 2004, 24 (12): 63-68.
- [4] CONEJO A J, PLAZSA M A, ESPINOLA R, et al. Day-ahead electricity price forecasting using the wavelet transform and ARIMA models [J]. IEEE Trans on Power Systems, 2005, 20 (2): 1035-1042.
- [5] 张显,王建学,王锡凡,等. 考虑多重周期性的短期电价预测[J]. 电力系统自动化, 2007, 31(3): 4-8.  
ZHANG Xian, WANG Jianxue, WANG Xifan, et al. Short-term electricity price forecasting based on price subsequences [J]. Automation of Electric Power Systems, 2007, 31 (3): 4-8.
- [6] 何书元. 应用时间序列分析[M]. 北京: 北京大学出版社, 2003: 230-252.
- [7] HE Shuyuan. Estimation of the mixed AR and hidden periodic model [J]. Acta Mathematicae Applicatae Sinica, 1997, 13(2): 196-208.
- [8] 申鼎焯. 随机过程[M]. 武汉: 华中理工大学出版社, 1995: 207-211.

(编辑 杜秀杰)