加权马尔可夫链在负荷预测中的应用

董继征[1],王 桓[1],何怡刚[1],陈洪云[1],王 薇[2]
(1. 湖南大学电气工程学院,湖南 长沙 410082; 2. 湖南大学机械与汽车工程学院,湖南 长沙 410082)

摘要: 利用小波将月售电量序列进行分解,对不受气象因素影响的趋势项和周期项,分别采用趋势外推和周期图法进行预测,对受气象因素影响的随机项,应用模糊聚类的方法确定分级标准,将观测值之间的相关系数作为权值,采用加权马尔可夫模型进行预测,然后进行综合。该法避免了精度较低的月气象资料对负荷预测精度的影响,不仅得到了未来月售电量的具体值,而且得到了其所属的区间,因此更加符合实际。最后给出了预测实例,验证了所提方法的有效性。

关键词: 负荷预测; 马尔可夫链; 售电量; 气象资料; 区间

中图分类号: TM715 **文献标识码**: A **文章编号**: 1003-4897(2006)03-0032-05

0 引言

电力负荷预测是电力部门的一项重要工作,季度型负荷预测对于节约用煤、机组检修、制定发购电计划、制定购煤计划、实行经济调度、实现电网商业化运营、保障社会生产和生活用电具有重要意义。电力系统负荷的变化无论从宏观与微观的角度分析都受太多因素的影响,这些影响因素的随机性一直制约着负荷预测精度的提高[1~4]。小波分析对局部时域和频域的分析具有明显的优势,通过选择合适的小波,对不同性质的负荷进行分类,从而可以针对某种性质的负荷,根据其规律采用相应的预测方法,提高预测精度[5~8]。文献[7]利用小波分析方法对非平稳时间序列进行分析,将趋势项、周期项和随机项分离,分别进行分析与预测,最后再合成得到原时间序列的预测值。文献[8]采用小波变换对日负荷数据进行分解处理,使得数据信息相对集中,在此基础上将小波分量分解为受气象因素影响的部分与不受气象因素影响的部分,对受气象因素影响的部分采用回归方法建立气象因素影响模型,对不受气象因素影响的部分,幅值大的分量建立回归神经网络预测模型,进行重点预测,而对幅值小的分量建立时间序列模型。上述方法均使预测精度得到一定程度的提高。

对于月售电量序列,在利用小波分析的方法进行分解之后,不受气象因素影响的部分为趋势项和周期项,这两项采用多项式(或趋势外推)模型和周期图法进行预测即可达到足够的精度[9],但对于受气象因素影响的随机项,序列本身波动较大,并且月度气象预报不准确甚至难以得到,所以要象短期负荷预测那样建立随机项与气象因素的关联模型比较困难,因此随机项的预测精度成为问题。实际上,由于随机项序列所表现出来的随机性主要是由气象因素的干扰造成的,而各月份气象之间具有一定的相关性,这些相关信息已经包含在时间序列之中。马尔可夫链可以描绘一个随机变化的动态系统,它根据状态之间的转移概率来推测一个系统未来的发展变化,而转移概率反映了各随机因素的影响程度,反映了各状态之间转移的内在规律性,因此,适合描述随机波动性较大的预测问题[10]。

通过对负荷的物理成因以及大量负荷序列资料的分析得知,负荷分解后的随机项为一相依随机变量,其相依关系的强弱,可以采用自相关系数作为其定量的测度。所以可以考虑应用模糊聚类的方法划分出反映负荷大小的变化区间,使每个区间内的数据包含相似的气象因素,然后以各阶自相关系数为权,用加权的马尔可夫链预测未来的负荷状况。

1 马尔可夫链

马尔科夫过程是具有无后效性的随机过程。无后效性是指:当过程在 t_m 时刻所处的状态为已知时,过程在大于 t_m 的时刻 t所处的状态的概率特性只与过程在 t_m 时刻所处的状态有关,而与过程在 t_m 时刻以前的状态无关。时间离散、状态离散的马尔科夫过程,简称马尔科夫链。其数学表达为:

定义在概率空间 (Ω, F, P)上的随机序列 $\{X(t),$

基金项目: 湖南省电力科学基金资助项目(20043005)

$t\in T\}$,其中 $T=(0, 1, 2, \cdots)$,状态空间 $I=\{0, 1, 2, \cdots\}$,称为马尔可夫链,如果对任意正整数 l, m, k及任意非负整数 $j_1>\cdots>j_2>j_1\ (m>j_1)$, i_{m+k}, i_m, i_{j1}, $\cdots$, i_{j2}, i_{j1}有

$$P\{X(m+k)=i_{m+k}\,|\,X(m)=i_m, X(j_l)=i_{jl}, \cdots, X(j_2)=i_{j2}, X(j_1)=i_{j1}\}=P\{X(m+k)=i_{m+k}\,|\,X(m)=i_m\} \quad (1)$$

成立。

如果对任意的 $k, n\in N^+$,有

$$P_{ij}(n, k)=P_{ij}(k),\quad i, j=0, 1, 2, \cdots \quad (2)$$

其中:$P_{ij}(n, k)$表示"系统时刻 n的状态为 i,经 k步转移至状态 j的概率",$P_{ij}(k)$表示"从状态 i经 k步转移至状态 j的概率",称马尔可夫链是齐次的。齐次马尔可夫链 $\{X(t)\}$完全由其初始分布 $\{P_{(i)}, i=0, 1, 2, \cdots\}$及其状态转移概率矩阵所决定。

由于负荷分解后的随机项是一相依随机变量,各阶自相关系数描述了序列之间的相关关系及其强弱,因此,可以考虑利用不同时段的历史负荷值对未来时段的负荷分别预测,然后根据不同历史时段与待测时段负荷的相依关系的强弱对各预测结果进行加权求和,这就可以达到充分、合理利用历史负荷所包含的气象信息进行预测的目的。

2　模糊有序聚类

模糊有序聚类是对有序变量进行分类的一种方法,应用模糊有序聚类的方法来划分负荷的变化区间,可以更加充分地考虑负荷序列的数据结构,使划分的区间更加合理。

设变量 $x_1, \cdots, x_n$的某一归类是 $\{x_i, \cdots, x_j\}$, $j\geq i$,定义其均值向量为:

$$\bar{x}_{ij}=\frac{1}{j-i+1}\sum_{l=i}^{j}x_l \quad (3)$$

将公式

$$D(i, j)=\sum_{l=i}^{j}(x_l-\bar{x}_{ij})'(x_l-\bar{x}_{ij}) \quad (4)$$

定义为 $\{x_i, \cdots, x_j\}\ (j\geq i)$的直径,其含义表示该变量段内部各变量之间的差异情况。$D(i, j)$越小,表示该段内变量之间差异越小,或相互间越接近;反之,$D(i, j)$越大,表示该段内变量之间差异越大,或相互间越分散。

设将 n个有序变量分为 K类,某一分法为:

$$P(n, K):\{i_1=1, \cdots, i_2-1\};\{i_2, \cdots, i_3-1\};\cdots;\{i_k, \cdots, n\}$$

定义这一分类的误差函数为:

$$e[P(n, K)]=\sum_{j=1}^{K}D(i_j, i_{j+1}-1) \quad (5)$$

作 $\min\{e[P(n, K)]\}$与 K的关系曲线,拐弯处的 K值为最优分类数,根据 $\min\{e[P(n, K)]\}$的递推公式可以求出相应的最优分类标准。

3　加权马尔可夫链预测步骤

1) 利用小波分析的方法将负荷序列进行分解,确定趋势项、周期项和随机项。对于趋势项可利用多项式拟和或趋势外推的方法预测[9],得到预测值 $\hat{x}_1$,对周期项可利用周期图法[8]进行预测,得到预测值 $\hat{x}_2$,限于篇幅,前两项的预测不再详述。对随机项则利用加权马尔可夫链预测。

2) 计算随机项的各阶自相关系数

$$r_k=\frac{\sum_{t=1}^{n-k}(x_t-\bar{x})(x_{t+k}-\bar{x})}{\sqrt{\sum_{t=1}^{n-k}(x_t-\bar{x})^2\sum_{t=1}^{n-k}(x_{t+k}-\bar{x})^2}} \quad (6)$$

其中:x_t表示 t月份的负荷值;$\bar{x}$表示负荷序列均值;n表示负荷序列长度。

3) 对各阶自相关系数归一化。归一化值为:

$$\bar{\omega}_k=\frac{|r_k|}{\sum_{k=1}^{m}|r_k|} \quad (7)$$

其中:m为预测所需自相关系数的最大阶数。该归一化值即为各步长马尔可夫链的权值。

4) 将随机项负荷序列值由小到大排列,应用模糊有序聚类方法确定负荷的分级标准,并据此确定各月份负荷值所处的状态,从而生成不同步长的马尔可夫链的转移概率矩阵。

5) 分别以预测期前面各历史负荷值为初始状态计算待测期负荷的状态概率 $P_i^{(k)}$, i为状态, $i\in I$, k为步长, $k=1, \cdots, m$。

6) 将不同步长的转移概率的预测值加权求和作为待测期的状态概率,则该点状态概率

$$P_i=\sum\bar{\omega}_k P_i^{(k)} \quad (8)$$

$\max\{P_i, i\in I\}$所对应的 i为预测期负荷所处的状态,状态 i的区间中值即为随机项在待测点的预测值 $\hat{x}_3$。

7) 计算总的售电量预测值为:

$$\hat{x}=\hat{x}_1+\hat{x}_2+\hat{x}_3 \quad (9)$$

4　预测实例

本文利用湖南省 2000～2003年 48个月份的售电量数据,对 2004年 1到 10月份的售电量进行了预测。

1)采用 db10对月售电量序列进行四尺度分解,求出各层的小波变换序列,观察图形,将第四层的低频信号作为趋势项(如图 1(a)所示),第四层和第三层的高频信号作为周期项(如图 1(b)所示),其它各层的高频信号作为随机项(如图 1(c)所示)。

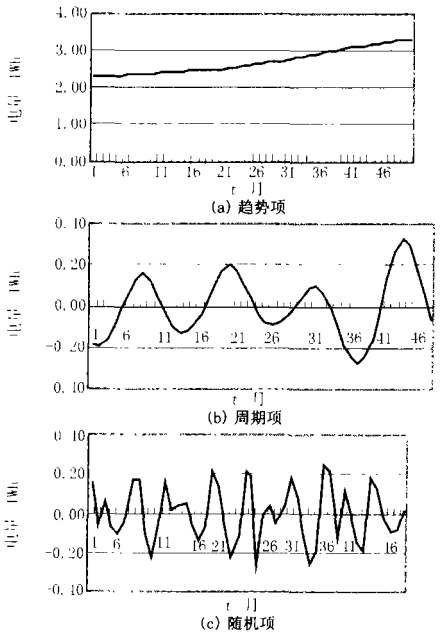

图 1 月售电量序列分解结果

Fig 1 Result of the monthly sales electric energy series decomposed

对于趋势项,采用趋势外推的方法预测,得到预测值为 $\hat{A}_1 = 3.321\,009$ TWh;对于周期项,采用周期图法预测,得到预测值为 $\hat{A}_2 = -0.219\,763$ TWh。对于受气象因素影响的随机项,采用马尔可夫预测方法。

2)利用式(6)计算随机项的各阶自相关系数为:

$r_1 = 0.132\,6,\ r_2 = -0.566\,4,\ r_3 = 0.053\,6$

3)利用式(7)将各阶自相关系数归一化,得到各步长的马尔可夫链预测权值

$\omega_1 = 0.176,\ \omega_2 = 0.753,\ \omega_3 = 0.071$

表 1 随机项序列值分类

Tab. 1 Classification of random series value

状态	类别	区间/GWh
1	负大	[-250.000, -191.144]
2	负小	(-191.144, -17.032]
3	正常	(-17.032, 82.710]
4	正小	(82.710, 189.721]
5	正大	(18.721, 250.000)

4)将随机项由小到大排列,利用 2所述有序聚类的方法将该序列进行分类,得到的最优分类标准见表 1。

5)由表 1确定随机项中各月份售电量的状态,生成各种步长的马尔可夫转移概率矩阵

$$P_1 = \begin{bmatrix} \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{2}{6} & \frac{1}{6} \\ \frac{3}{16} & \frac{6}{16} & \frac{2}{16} & \frac{3}{16} & \frac{2}{16} \\ 0 & \frac{5}{10} & \frac{3}{10} & \frac{2}{10} & 0 \\ \frac{2}{11} & \frac{3}{11} & \frac{3}{11} & \frac{3}{11} & 0 \\ 0 & \frac{1}{4} & 0 & \frac{2}{4} & \frac{1}{4} \end{bmatrix}$$

$$P_2 = \begin{bmatrix} 0 & 0 & \frac{1}{6} & \frac{2}{6} & \frac{3}{6} \\ \frac{2}{15} & \frac{4}{15} & \frac{3}{15} & \frac{5}{15} & \frac{1}{15} \\ \frac{1}{10} & \frac{4}{10} & \frac{3}{10} & \frac{2}{10} & 0 \\ \frac{1}{11} & \frac{6}{11} & \frac{4}{11} & 0 & 0 \\ \frac{2}{4} & \frac{1}{4} & 0 & \frac{1}{4} & 0 \end{bmatrix}$$

$$P_3 = \begin{bmatrix} 0 & \frac{1}{6} & \frac{3}{6} & \frac{1}{6} & \frac{1}{6} \\ \frac{3}{14} & \frac{2}{14} & \frac{2}{14} & \frac{5}{14} & \frac{2}{14} \\ 0 & \frac{5}{10} & \frac{2}{10} & \frac{3}{10} & 0 \\ \frac{3}{11} & \frac{5}{11} & \frac{2}{11} & 0 & \frac{1}{11} \\ 0 & \frac{1}{4} & \frac{2}{4} & \frac{1}{4} & 0 \end{bmatrix}$$

6)利用随机项在 2003年 10～12月份序列值及相应的状态转移概率矩阵对 2004年 1月份的售电量状态进行预测,12月份初始状态为 3,其一步状态转移概率为:

$$P_i^{(1)} = \left(0 \quad \frac{5}{10} \quad \frac{3}{10} \quad \frac{2}{10} \quad 0 \right)$$

11月份初始状态为 2,其两步转移概率为:

$$P_i^{(2)} = \left(\frac{2}{15} \quad \frac{4}{15} \quad \frac{3}{15} \quad \frac{5}{15} \quad \frac{1}{15} \right)$$

10月份初始状态为 2,其三步状态转移概率为:

$$P_i^{(2)} = \left(\frac{3}{14} \quad \frac{2}{14} \quad \frac{2}{14} \quad \frac{5}{14} \quad \frac{2}{14} \right)$$

由式(8)预测 2004年 1月份的状态概率为:

$$P_i = \left(0.116 \quad 0.299 \quad 0.214 \quad 0.312 \quad 0.060 \right)$$

显然,当 $i=4$时, P_i最大,故可以预测随机项在2004年 1月份的状态为 4,即随机项在 2004年 1月份的区间值为 (82. 710, 189. 721) GWh,如果将区间中值作为预测值,则随机项在 2004年 1月份的预测值为 $\Delta_3 = 136.216$ GWh。

7)利用式 (9)计算的 2004年 1月份售电量预测值为 3. 237 862 TWh。

同理,对 2004年 2～10月份售电量进行预测时,将前一个月份的实际售电量值加入到原始序列中,去除原始序列中第一个点的数据,得到等维新息序列,重复 1)～7)步,便得到各待测点的预测值。

此外,笔者采用时间序列法 (这里采用的是 AR(2)模型)以及组合预测法 (这里采用滑动平均、指数平滑、时间序列、神经网络 4种模型的组合预测模型)对待测点的售电量进行了预测,并与本文方法的预测结果进行了比较,结果如图 2所示,三种方法的预测误差见表 2,可见本文方法具有较好的预测精度。

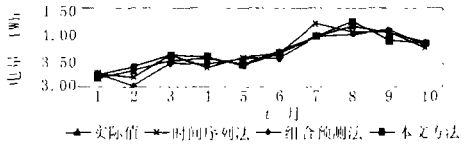

图 2　预测结果比较

Fig. 2　Comparison of forecasted results

表 2　湖南省网 2004年月售电量预测

Tab. 1　Electric energy forecasting of Hunan Provincial grid in 2004

月份	实际值	时间序列法		组合预测法		本文方法	
	/TWh	预测值/TWh	误差	预测值/TWh	误差	预测值/TWh	误差
1	3.178 342	3.268 700	2.84%	3.257 200	2.48%	3.237 862	1.87%
2	3.317 567	3.195 900	-3.67%	3.038 800	-8.40%	3.391 579	2.23%
3	3.507 998	3.609 800	2.90%	3.460 300	-1.36%	3.613 541	3.01%
4	3.539 624	3.377 000	-4.59%	3.437 550	-2.88%	3.586 694	1.33%
5	3.439 313	3.561 500	3.55%	3.549 500	3.20%	3.419 987	-0.56%
6	3.688 267	3.633 300	-1.49%	3.538 750	-4.05%	3.649 657	-1.05%
7	3.974 849	4.201 000	5.69%	3.963 550	-0.28%	3.976 997	0.05%
8	4.169 901	4.065 600	-2.50%	4.012 750	-3.77%	4.261 455	2.20%
9	4.016 975	4.033 300	0.41%	4.096 350	1.98%	3.889 212	-3.18%
10	3.864 049	3.757 100	-2.77%	3.808 900	-1.43%	3.847 559	-0.43%

5　结论

由于电力负荷受许多外界因素的影响,现在对负荷的预测逐渐进入细化阶段,小波分析是解决这个问题的有效工具,通过对负荷的分解,明晰了负荷的各种变化规律。然而,由于月气象资料的不准确性,进行月售电量预测时不能象超短期负荷预测那样建立随机项与气象因素的关联模型。小波分解重构所得到的月售电量序列的随机项波动较大,利用模糊聚类的方法可以将包含相似气象因素的负荷归为一类,然后采用加权马尔可夫链对随机项进行预测。这种方法避免了精度较低 (甚至难以得到)的月气象资料对预测精度的影响,不仅可以预测出未来的月售电量,而且可以预测出未来月售电量所处的区间范围,这更加符合实际情况。

参考文献:

[1] 李金颖,牛东晓. 非线性季节型电力负荷灰色组合预测研究[J]. 电网技术, 2003, 27(5): 26-29.
LI Jin-ying, NIU Dong-xiao. Study on Grey Combined Forecasting of Nonlinear Seasonal Power Load[J]. Power System Technology, 2003, 27(5): 26-29.

[2] 张国江,邱家驹,李继红. 基于模糊推理系统的多因素电力负荷预测[J]. 电力系统自动化, 2002, 26(5): 49-53.
ZHANG Guo-jiang, QIU Jia-ju, LI Ji-hong. Multi-factor Short-term Load Forecasting Based on Fuzzy Inference System[J]. Automation of Electric Power Systems, 2002, 26(5): 49-53.

[3] 王文圣,丁 晶,赵玉龙,等. 基于偏最小二乘回归的年用电量预测研究[J]. 中国电机工程学报, 2003, 23(10): 17-21.
WANG Wen-sheng, DING Jing, ZHAO Yu-long, et al. Study on the Long Term Prediction of Annual Electricity Consumption Using Partial Least Square Regressive Model[J]. Proceedings of the CSEE, 2003, 23(10): 17-21.

[4] Hippert H S, Pefreira C E, Souza R C. Neural Network for Short-term Load Forecasting: a Review and Evolution[J]. IEEE Trans on Power Systems, 2001, 16(2): 44-54.

[5] 邰能灵, 侯志俭. 小波模糊神经网络在电力系统短期负荷预测中的应用[J]. 中国电机工程学报, 2004, 24(1): 24-29.
TAI Neng-ling, HOU Zhi-jian. New Short-term Load Forecasting Principle with the Wavelet Transform Fuzzy Neural Network for the Power System[J]. Proceedings of the CSEE, 2004, 24(1): 24-29.

[6] Nason G P, Sachs R V. Wavelets in Time Series Analysis[J]. Philosophical Transactions of the Royal Society Series A, 1999: 2511-2516.

[7] 谢宏,陈致业,牛东晓,等. 基于小波分解与气象因素影响的电力系统日负荷预测模型研究[J]. 中国电机工程学报, 2001, 21(5): 5-10.
XIE Hong, CHEN Zhi-ye, NIU Dong-xiao, et al. The Research of Daily Load Forecasting Model Based on Wavelet Decomposing and Climatic Influence[J]. Proceedings of the CSEE, 2001, 21(5): 5-10.

[8] 马社祥,刘贵忠,曾召华. 基于小波分析的非平稳时间序列分析与预测[J]. 系统工程学报, 2000, 15(4): 305-311.
MA She-xiang, LIU Gui-zhong, ZENG Zhao-hua. Wavelet-based Analysis and Forecasting for Non-Stationary Time Series[J]. Journal of System Technology, 2000, 15(4): 305-311.

[9] 牛东晓,曹树华,赵磊,等. 电力负荷预测技术及其应用[M]. 北京:中国电力出版社, 1998.
NIU Dong-xiao, CAO Shu-hua, ZHAO Lei, et al. Power Load Forecasting Technology and Its Application[M]. Beijing: China Electric Power Press, 1998.

[10] 任雪梅,陈逊,徐腊元. 基于马尔可夫链和模糊聚类的电力系统短期负荷预测[J]. 北京理工大学学报, 2004, 24(5): 416-422.
REN Xue-mei, CHEN Xun, XU La-yuan. Power System Short Term Load Forecasting Based upon a Combination of Markov Chain and Fuzzy Clustering[J]. Transactions of Beijing Institute of Technology, 2004, 24(5): 416-422.

收稿日期: 2005-06-28; **修回日期**: 2005-10-05

作者简介:

董继征 (1980 -),男,硕士研究生,研究方向为电力系统稳定性及电力市场; E-mail: djzh1105@163.com

王 桓 (1980 -),男,博士研究生,研究方向为电力系统稳定性及电力市场;

何怡刚 (1966 -),男,教授,博士生导师,从事电力市场、小波、神经网络研究。

Application of Markov chain with weights to load forecasting

DONG Ji-zheng[1], WANG Huan[1], HE Yi-gang[1], CHEN Hong-yun[1], WANG Wei[2]
(1. School of Electrical Engineering and Information, Hunan University, Changsha 410082, China;
2. School of Machine and Automobile Engineering, Hunan University, Changsha 410082, China)

Abstract: The series of the monthly sales electric energy is resolved by utilizing the wavelet. To the trend term and cycle term free of meteorological factor, this paper applies trend extrapolate and cyclic graph methods to predict them separately. To the last term at random that is influenced by meteorological factor, this paper employs fuzzy sequential cluster method to set up the classification standard, then it regards the correlation coefficients of record values as weights and predicts the future loads by using Markov chain model with weights. And all results are synthesized at last. This method avoids the influence to the load forecasting prediction of the moon meteorological data, which has a bad prediction. Not only the concrete value of the monthly sales electric energy but its range in the future is gained. So the prediction is more practical. A forecasting instance is provided and the validity of the method proposed is tested finally.

This project is supported by Electrical Science Foundation of Hunan Province (No. 20043005).

Key words: load forecasting; Markov chain; sales electric energy; meteorological data; part of the normal route

欧美国家核电态度由冷变热

面对国际油价高涨并且难有较大回落的局面,欧美国家近年来纷纷制定本国电源结构调整政策,修建核电站,扩大核电比重,以此解决日益增长的电力需求问题。

美国总统布什 2005年上半年公开表示,美国的能源战略调整任务是要减少对石油的依赖程度,因此要重新启动核电站建设计划。美国政府拟新建 70万 kW级核电站 2座。

另外,上世纪 70年代率先冻结核电计划的欧盟国家,近来也纷纷从国家新能源政策调整角度来重新审视对待核电站问题。在欧洲,芬兰首先带头开工建设自上世纪 70年代之后的第一座核电站;法国政府拟投资 36亿美元兴建核电站;英国第四大电力公司也准备重新启动核电站建设;曾受过切尔诺贝利核电站泄漏严重影响的乌克兰政府为减少对俄罗斯石油进口的过分依赖,也准备建设核电站;对核电站管理最严格、原计划要分阶段关闭国内所有核电站的德国,态度也在发生变化,政府表示要重新审议对核电站的有关政策,放宽原来的过分限制。这是德国重新启动核电站建设,让该国重新认识核电问题的舆论和思想准备。