

DOI:10.3969/j.issn.1001-4551.2016.03.018

基于自组织特征映射的连续多日负荷 预测方法研究^{*}

马立新,周尚璠

(上海理工大学 光电信息与计算机工程学院,上海 200093)

摘要:针对影响连续多日每日最大负荷的因素较多且构成复杂,连续多日负荷预测方法少难度大、含节假日的连续多日负荷预测精度低等问题,分析了近几年工作日电力负荷数据特点,研究了自组织特征映射(Self-Organizing Feature Map, SOM)聚类算法并将其用于负荷数据的预处理,研究了节假日负荷的特性,总结了其负荷变化规律并加以区分预测,提出了一种基于自组织特征映射神经网络的连续多日负荷预测新方法。该方法区分普通工作日与节假日,普通工作日采用自组织特征映射神经网络聚类方法对日最大负荷进行特征提取,建立了以周期特征相似的历史数据作为训练样本的神经网络模型,节假日设定假日影响因子单独预测。运用某市近年的负荷数据进行预测,算例结果显示综合预测误差为3.21%,表明该方法预测精度完全满足实际需求,为连续多日最大负荷预测提供了一种可行的方法。

关键词:连续多日负荷预测;特征提取;自组织特征映射聚类;日最大负荷

中图分类号:TM715

文献标志码:A

文章编号:1001-4551(2016)03-0342-05

Multi-Days load forecasting base on self-organizing feature map

MA Li-xin, ZHOU Shang-jun-xi

(School of Optical-Electrical and Computer Engineering, University of Shanghai for
Science and Technology, Shanghai 200093, China)

Abstract: Aiming at the problem of the factors affecting maximum load in continuous multi-days are complex, hard to forecast the multi-days load, multi-days load including holidays forecasting has low accuracy, characteristics of power load in recent years were analyzed, the Self-Organizing Feature Map cluster method was used in data processing, loading data, characteristics of holidays were studied, a new method of multi-days load forecasting base on Self-Organizing Feature Map(SOM) neural network was proposed. The characteristics of the holiday and the non-holiday daily peak load were investigated respectively. For the non-holiday, the Elman neural network uses the training data which was selected by SOM clustering to forecast. For the holiday, the holiday factor is added. The experimental results show that the prediction accuracy of this method can meet the industry requirement, the average error is 3.21%. It provides a feasible way for multi-day load forecasting.

Key words: multi-days load forecasting; feature extraction; Self-Organizing Feature Map clustering; maximum load

0 引 言

负荷预测是电力系统调度运营的重要环节,其精度直接影响着电网的安全稳定。按预测时长,可将负荷预测分为:超短期、短期、中期、中长期负荷预测。目

前研究最广泛的是短期及中长期负荷预测,而对于时长介于两者之间的连续一周至一个月的负荷预测研究相对较少。连续多日负荷预测便是处于中期预测和短期预测间的新型负荷预测,对未来连续多日甚至一个月的负荷进行预测对电力调度和电力营销有着重要意

收稿日期:2015-10-28

基金项目:上海张江国家自主创新重点资助项目(201310-PI-B2-008)

作者简介:马立新(1960-),男,陕西安康人,工学博士,教授,主要从事电力系统分析与优化运行,智能电网与智能科学,电气设备状态监测与诊断方法,电能质量监控与能效测评技术,风光发电智能控制系统方面的研究。E-mail:malx_aili@sina.com

义。但影响连续多日负荷的因素较多且构成复杂,因此对连续多日的负荷进行预测的难度较大,通常通过分析历史负荷曲线,挑选与预测日负荷特征相似的数据构造样本进行预测,但预测时间往往较短^[1],相关研究也较少。

日最大负荷是指一昼夜时间内负荷最大峰值,是估算电力系统备用容量、日负荷率、制定峰谷电价等指标的重要依据,影响着整个电网、尤其是网架结构薄弱地区电网运行的安全稳定性。目前,研究者对日最大负荷的研究通常集中在短期负荷预测,并且大多致力于气象因素和日最大负荷的关系^[2-3]。由于未来连续多日的气象信息难以获取且准确度不高,因此对于连续多日日最大负荷的研究也较少。

目前,自组织特征映射网络(self-organizing feature map,SOM)与神经网络相结合的预测模型多用于对疾病、天气的预测^[4],本研究基于此提出一种基于 SOM-Elman 神经网络的连续多日负荷的预测方法,同时充分考虑连续多日中存在节假日的可能,针对节假日负荷规律与工作日不同的特点,分别分析二者日最大负荷变化规律并加以利用。非节假日,本研究通过自组织特征映射网络对历史负荷数据进行特性分析并结合 Elman 神经网络进行预测,该方法通过特征提取消除冗余信息,精简输入样本数量,利用 Elman 神经网络具有适应时变能力、预测精度准确、预测时间短的优点,以提高对未来一个月每日最大负荷预测的精度。节假日,本研究通过分析假日期间负荷特点,设定假日影响因子,算例分析结果表明预测精度显著提高。

1 SOM-Elman 组合预测方法

1.1 SOM 神经网络聚类

自组织特征映射网络(Self-Organizing Feature Map, SOM)是一种特征聚类和高维可视化的无教师学习算法,它是通过模拟大脑神经元的侧向抑制、自组织等特性而产生的一种人工神经网络^[5],网络结构如图1所示。

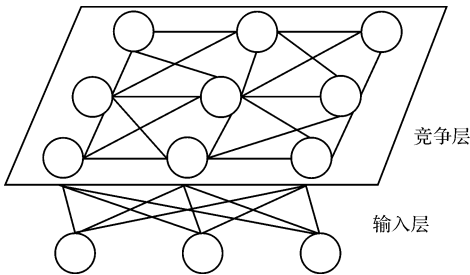


图1 自组织映射网络结构

算法流程如下:

(1)初始化

取 $[0,1]$ 区间的随机值作为神经网络的连接权值 W_{ij} ,初始领域 $N_c(0)$ 设置为较大范围,确定网络的学习速率 $\eta(0)$, $0 < \eta(0) < 1$;迭代次数为 T ,归一化:

$$\bar{X}_k = \frac{X_k}{\|X_k\|} = \frac{(X_{1k}, X_{2k}, \dots, X_{ik}, \dots, X_{nk})}{\sqrt{(X_{1k})^2 + \dots + (X_{ik})^2 + \dots + (X_{nk})^2}} \quad (1)$$

$$\bar{W}_j = \frac{W_j}{\|W_j\|} = \frac{(W_{1j}, W_{2j}, \dots, W_{ij}, \dots, W_{nj})}{\sqrt{(W_{1j})^2 + \dots + (W_{ij})^2 + \dots + (W_{nj})^2}} \quad (2)$$

X_k — n 维, $k=1,2,\dots,N$; W_j — n 维, $j=1,2,\dots,M$ 。

(2)接受输入

取任意一个 $X_k: X_k = (X_{1k}, X_{2k}, \dots, X_{nk})$ 从网络的输入层输入,作为输入向量。

(3)寻找获胜节点

选择输入向量 $\bar{X}_k$ 和连接权向量 $\bar{W}_j$ 间欧式距离最小的神经元 c 为获胜神经元,即: $\|X_k - W_c\| = \min\{d_{jk}\}$

(4)参数调整

对式(3)中的获胜节点 c 及其领域内所有节点的权值按公式(3)进行调整。

$$W_{ij}(t+1) = W_{ij}(t) + \eta(t)h_{c,j}(t)(X_i - W_{ij}(t)) \quad (3)$$

式中: $0 < \eta(t) < 1$ —学习速率,是增益函数; $h_{c,j}(t)$ —领域函数。

(5)更新学习速率 $\eta(t)$ 和领域函数 $h_{c,j}(t)$

$h_{c,j}(t)$ 通常采用高斯函数,如下:

$$h_{c,j}(t) = \exp\left(-\frac{d_{cj}^2}{2r^2(t)}\right) \quad (4)$$

式中: d_{cj} —获胜神经元 c 和任一领域内被激活的输出神经元 j 之间的距离,调整规律如下式所示:

$$r(t+1) = \text{INT}\left(\left(r(t) - 1\right) \times \left(1 - \frac{t}{T}\right)\right) + 1 \quad (5)$$

$$\eta(t+1) = \eta(t) - \frac{\eta(0)}{T} \quad (6)$$

(6)循环学习

将其他输入模式依次重复步骤(3),直到所有输入模式学习完毕;

(7)令 $t=t+1$,返回步骤(2),直至 $t=T$ 或网络收敛为止^[6-8]。

SOM 神经网络聚类的显著特征是它赋予质心一种空间组织,这是它与其他基于原型的聚类的根本差别,它将相邻关系强加于簇质心上,因此互为邻居的簇之间比非邻居之间更相关,聚类准确度更高^[9-10]。

1.2 Elman 神经网络

Elman 是一种典型的动态神经网络,它在 BP 人工神经网络输入层、隐含层、输出层的基本结构基础上增加一个承接层,它用来存储内部状态,从而使系统具有适应时变特性的能力^[11],由于电力系统负荷具有动态特性,动态神经网络对电力负荷的预测准确度更高,时间更短^[12],结构如图 2 所示。

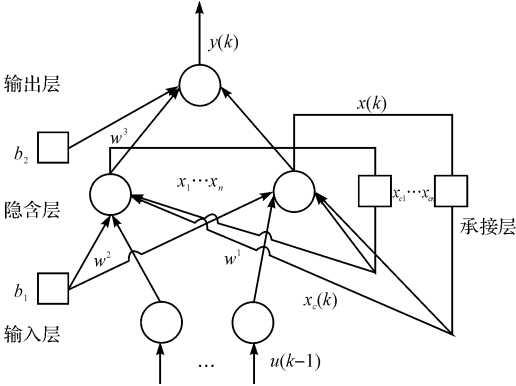


图 2 Elman 神经网络结构

Elman 网络空间表达式为:

$$y(k) = g(w^3x(k)) \tag{7}$$

$$x(k) = f(w^1x_c(k) + w^2(u(k-1))) \tag{8}$$

$$x_c(k) = x(k-1) \tag{9}$$

式中: y —输出节点向量; x —中间层节点向量; u —输入向量; x_c —反馈状态向量; w^3 —中间层至输出层连接权值; w^2 —输入层至中间层连接权值; w^1 —承接层到中间层的连接权值; $g(*)$ —输出神经元的传递函数,是中间层输出的线性组合; $f(*)$ —中间层神经元的传递函数,常用 S 函数^[13]。

1.3 SOM-Elman 神经网络组合预测方法

连续多日负荷预测是介于短期负荷预测和中期负荷预测间的新型负荷预测,它主要针对未来一周至一月的负荷进行分析预测。但电力负荷本身具有非线性及随机性的特点,导致负荷波动的因素很多。未来连续多日的最大负荷变化特点与过去近一周的负荷数据相似度并不高,若模仿短期负荷预测方法,直接利用前一周的负荷数据对日最大负荷进行预测准确度不高。因此需找出与待测日类似,相关度较高的历史负荷数据作为输入。对日大负荷的特征进行提取,寻找相似序列是连续多日负荷预测的关键。

1.3.1 日最大负荷特征提取

本研究利用 SOM 网络对输入的历史日最大负荷数据进行聚类分析,经过迭代计算,得到各个神经元之间的权值,各个类别的神经元会将数据之间存在的相

似性体现出来,使得同一类别的神经元的权值比较相近,最后将训练样本数据分类,作为训练样本集。

本研究取某市 2011 年 1 月 1 日至 2013 年 8 月 30 日的(除春节、国庆节所在月份)共 37 组电力负荷数据,编号后通过 SOM 算法对输入样本进行聚类,学习次数 $T=500$ 。考虑到特征向量的维数以及实用性,分为 4 类,设置网络的输出节点矩阵为 2×2 结构,其中分类结果由各组数据的编号表示,聚类结果如表 1 所示。

表 1 聚类结果

输出神经元	聚类结果
1	1、2、3、4、5、8、9、10、11、12、13、14、19、23
2	6、15、21、24、25、29、32、34
3	7、18、22、28、33
4	16、17、20、26、27、30、31、35、36、37

分类拓扑结构如图 3 所示。

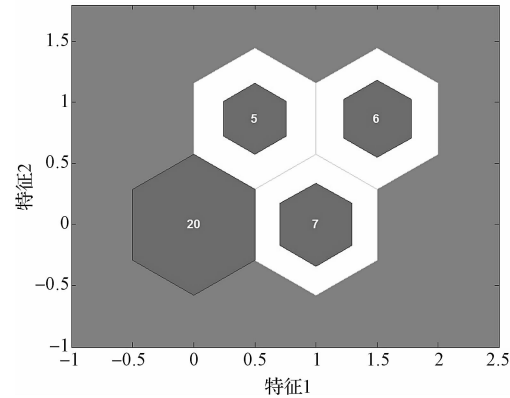


图 3 分类拓扑结构

1.3.2 SOM-Elman 神经网络组合模型

本研究所用预测方法流程图如图 4 所示。

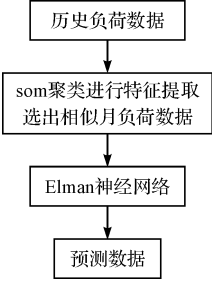


图 4 SOM-Elman 神经网络组合预测流程图

本研究对历史负荷数据利用自组织特征映射算法进行特征提取,对多组历史数据进行分类,选出与预测月相似的同类型月,以其历史负荷作为神经网络的输入样本集,通过建立 Elman 神经网络进行负荷预测。

2 法定假日预测方法

清明节、五一劳动节、中秋节等重大节假日(本研

究仅针对法定假期为3天的节日进行分析)。由于法定节假日大多数工业暂停,假日期间的最大负荷与其他时期存在较大差异^[14],将节假日与日常负荷用相同方法进行预测效果不佳。为了充分合理考虑实际情况,对节假日单独进行分析建模预测十分必要。

2.1 假日最大负荷特征分析

通过对比2010年-2013年某市节假日(本研究中特指放假为期3天的法定节假日)的最大负荷数据,分析后发现:

- (1)节假日期间,日最大负荷明显降低,且节日当天负荷达到最低值;
- (2)法定节日出现在假期第一天,则第一天负荷最低,随后逐日升高;法定假日出现在第二日,则第一天负荷最大,第三天次之,节日当天最低;法定假日出现在第三天,则与日递减。

2011年某市元旦、清明、五一节假日3天的最大负荷曲线如图5所示,图中节日分别出现在第一天、第二天、第三天。

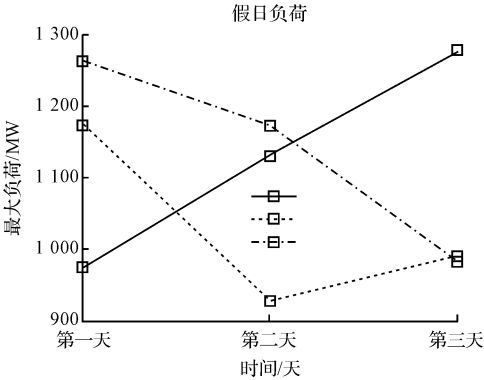


图5 2011年某市节假日日最大负荷曲线

2.2 法定假日预测方法

根据本研究中分析的节假日最大负荷变化特点,分别针对节日所在位置的不同情况设置节假日影响因子,对其进行编码:

- (1)节日为第一天:3天假日的影响因子编码分别为0.15 0.30 0.55;
- (2)节日为第二天:3天假日的影响因子编码分别为0.55 0.15 0.30;
- (3)节日为第三天:3天假日的影响因子编码分别为0.55 0.30 0.15;

取某市2010年1月至2013年4月期间3天节假日的负荷数据作为神经网络的训练样本,将历史负荷数据与假日影响因子作为Elman神经网络的输入,进行建模预测,可得到精度较高的假日最大负荷。

3 算例分析

本研究用2011年1月到该年4月每月每日最大负荷对该年5月每日最大负荷进行预测,将预测结果与BP神经网络预测的结果进行对比,其中两种预测方法的隐含层节点数相同,验证本研究方法的可行性。

预测误差采用相对误差、平均绝对百分误差(MAPE)衡量^[15],即:

$$APE = \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100\% \tag{10}$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100\% \tag{11}$$

通过SOM神经网络对历史负荷数据进行聚类分析,得到的聚类结果如表2所示。

表2 SOM聚类结果

输出神经元	聚类结果
1	16、17、26、27、30、31、34
2	7、18、22、28、33
3	6、15、20、21、25、29、32
4	1-5、8-14、19、23、24

经聚类后的样本数据分别建立Elman神经网络模型进行预测,选取预测误差最小的一组作为结果。某市某年5月日最大负荷预测误差分析如表3所示。

表3 某市某年5月日最大负荷预测误差分析

日期	实际值	BP 预测方法		本研究方法	
		预测值	APE	预测值	APE
1日	950.9	1 185.0	24.64	956.8	0.63
2日	1 244.5	1 214.0	2.45	1 345.4	8.11
3日	1 353.0	1 388.7	2.64	1 313.8	2.90
4日	1 376.3	1 511.4	9.82	1 291.5	6.16
5日	1 295.9	1 654.6	27.68	1 292.0	0.30
6日	1 420.3	1 452.4	2.26	1 373.7	3.28
7日	1 429.2	1 467.2	2.66	1 368.9	4.22
8日	1 448.2	1 424.5	1.64	1 394.8	3.69
9日	1 436.9	1 411.1	1.80	13 87.0	3.47
10日	1 427.7	1 143.0	19.94	1 380.7	3.29
11日	1 393.0	1 238.9	11.06	1 344.7	3.47
12日	1 271.8	1 180.8	7.16	1 327.6	4.39
13日	1 458.2	1 484.6	1.81	1 434.5	1.63
14日	1 489.2	1 536.4	3.17	1 436.2	3.56
15日	1 489.3	1 332.1	10.56	1 421.3	4.57
16日	1 442.7	1 288.5	10.69	1 403.5	2.72
17日	1 451.2	1 213.9	16.35	1 436.3	1.03
18日	1 392.0	1 411.9	1.43	1 345.6	3.33
19日	1 249.8	1 537.5	23.02	1 322.9	5.85
20日	1 441.0	1 200.8	16.67	1 416.4	1.71
21日	1 471.5	1 502.1	2.08	1 430.6	2.78

续表 3

日期	实际值	BP 预测方法		本研究方法	
		预测值	APE	预测值	APE
22 日	1 514.2	1 374.8	9.21	1 440.6	4.86
23 日	1 477.5	1 512.6	2.38	1 463.3	0.96
24 日	1 509.3	1 441.3	4.51	1 439.3	4.64
25 日	1 442.3	1 447.7	0.37	1 406.1	2.51
26 日	1 277.0	1 449.4	13.50	1 374.0	7.60
27 日	1 471.7	1 569.4	6.64	1 462.2	0.65
28 日	1 446.3	1 611.2	11.40	1 464.7	1.27
29 日	1 522.3	1 427.1	6.25	1 464.4	3.80
30 日	1 449.7	1 444.0	0.39	1 427.1	1.56
31 日	1 438.2	970.0	32.55	1 427.9	0.72
MAPE		8.74		3.21	

从表 3 算例结果可知,采用 SOM-Elman 组合模型进行预测,平均绝对误差为 3.21%。由于预测前对历史负荷数据进行了聚类分析,预测时使用的样本数据与待预测月负荷变化规律具有相似性,相关程度较高因此预测准确度较高。Elman 神经网络具有适应时变特性的能力,能更加准确地反应负荷的动态变化特性,因而预测精确度比起静态神经网络 BP 神经网络来说较高,算例中精确度提高 5.53%。对于节假日,本研究用 BP 神经网络对五一节当天负荷进行预测,由于不能对节假日这一突变做出判断因此预测结果十分不理想,而将 5 月 1 日劳动节单独建立模型进行预测后,充分利用节假日的负荷特性将预测误差缩小到 0.63%,结果准确度大幅度提高。

4 结束语

本研究首次将 SOM 聚类用于连续多日的日最大负荷特征提取,并区分节假日对其负荷特性进行单独研究。提出的 SOM-Elman 组合预测模型实现了对未来一个月的每日最大负荷的预测,并且预测精度较高。对未来一个月的日最大负荷的预测有助于电力部门提前对重要设备进行巡查抢修,也对电力调度有所帮助,具有一定的实际应用价值。

经对算例结果的分析,预测结果中个别误差较大的数据是受当天气温及降雨量影响。因此,如何将气象因素对连续多日日最大负荷的影响考虑进预测模型及如何得到准确的未来一个月的气象预报都是有待深入研究的问题。

参考文献(References):

[1] 赵 傲,康重庆,葛 睿,等. 电力市场中多日负荷曲线的预测[J]. 电力自动化设备,2002,22(9):31-33.

[2] 李湘华,宋 尖,刘阳升,等. 典型气象因素对省网夏季日最大负荷的影响[J]. 电力科学与技术学报,2012,27(2):70-75.

[3] 张 斌,刘阳升,宋 尖,等. 考虑高温累积效应的夏季日最大负荷预测[J]. 湖南电力,2012,32(4):5-9.

[4] 王 杰,闫东伟. 提高预测精度的 ELMAN 和 SOM 神经网络组合[J]. 系统工程与电子技术,2004,26(12):1943-1945.

[5] 梁建武,陈祖权,谭海龙. 短期负荷预测的聚类组合和支持向量机方法[J]. 电力系统及其自动化学报,2011,23(1):34-38.

[6] 王艳华. SOM 研究的若干新进展[J]. 福建电脑,2013,11:1-4.

[7] 严金华. 基于 SOM 神经网络的负荷特性分类[J]. 大众科技,2013,15(172):31-33.

[8] 简献忠,曹树建,郭 强. SOM 聚类与 Voronoi 图在验证码字符分割中的应用[J]. 计算机应用研究,2007,32:1-7.

[9] Arroyo J R, Espinola. Different approaches to forecast interval time series: a comparison in finance[J]. **Computational Economics**,2011,37(2):169-191.

[10] Valero S, Aparicio J, Senabre C, et al. Analysis of d-different testing parameters in Self-Organizing Feature Maps for short-term load demand forecasting in Spain[C]. 2010 Proceedings of the International Symposium Modern Electric Power Systems (MEPS), Wroclaw, 20-22September, 2010: 1-6.

[11] 朱 晟,蒋传文,侯志俭. 基于气象负荷因子的 Elman 神经网络短期负荷预测[J]. 电力系统及其自动化学报,2005,17(1):23-26.

[12] 牛晓华,曹树华. 电力负荷预测技及其应用[M]. 北京:中国电力出版社,2009.

[13] 史 峰,王小川,郁磊,等. MATLAB 神经网络 30 个案例分析[M]. 北京:北京航空航天大学出版社,2010.

[14] 马立新,李 渊. 区间电力负荷特征提取及预测方法[J]. 控制工程,2015,22(4):645-648.

[15] 马立新,尹晶晶,郑晓栋等. 智能用电环境下电力负荷预测方法的研究[J]. 机电工程,2015,32(9):1233-1237.

[编辑:张 豪]

本文引用格式: