

96 点短期负荷预测方法及其应用

鞠 平¹, 姜 巍¹, 赵夏阳¹, 王俊锴², 张世学², 刘 琰²

(1. 河海大学电力系, 南京 210098; 2. 陕西省电力公司, 西安 710004)

摘要: 结合相似日方法和人工神经网络方法的优点, 提出一种基于相似日和径向基函数网络的新方法。该方法具有较好的稳定性和精度, 与普通相似日方法相比, 减少了对人的经验的依赖, 但仍然保持了应用简便、实用有效的特点。开发了 96 点短期负荷预测软件, 并在陕西电网成功运行 1 年多。给出了几种方法的实际应用结果, 并进行了比较, 统计表明新方法的平均均方根误差较普通相似日方法从 3.153% 降低到 2.954%, 用综合方法则进一步降低到 2.854%。

关键词: 电力系统; 负荷预测; 相似日方法; 径向基函数网络

中图分类号: TM 715; TP 18

0 引言

随着电力市场的发展, 负荷预测越来越受到重视, 英国的研究结果表明, 短期负荷预测的误差每增加 1% 将导致每年运行成本增加约 1 770 万元。在挪威, 每增加 1% 的短期负荷预测误差将导致 455 万元~910 万元的附加运行成本^[1]。

负荷预测研究早在五六十年代就开始了, 自 80 年代以来, 随着计算机技术的迅速发展, 负荷预测的速度和精度逐步提高。短期负荷预测的方法有许多^[1~5], 常用的预测方法有时间序列方法、相似日方法和神经网络方法等。

负荷的历史记录总是构成一个时间序列。目前, 时间序列方法是处理随机序列的基本方法, 也是负荷预测中应用最广和较为成功的方法之一^[2,3]。其优点是: ①原理成熟、应用简便; ②主要根据过去的负荷值及干扰值来推算未来的负荷, 不需要相关因素的资料, 因此在一些相关因素的预测值和某些常数难以得到时, 不失为一种可行的方法。缺点是: ①对数据的要求较高; ②用线性模型表达负荷之间非线性关系有一定局限性; ③不能方便地考虑天气情况等对负荷有重要影响的相关因素。

人们在进行短期负荷预测时, 发现气象状况、日类型(指工作日或休息日)等影响因素比较相似的两天, 负荷也比较接近。在实际运行中, 有丰富经验的负荷预测工作人员在进行负荷预测时, 往往能找出与预测日较为相似的某天的负荷, 再进行修正。相似日方法^[4]正是沿着这一被实践证明有效的思路, 由以下 2 个步骤组成: ①采用某种差异评估函数, 寻找最有可能与预测日负荷相似的某天(相似日); ②根

据待预测负荷日的参数进行修正。然而该方法存在以下缺点: ①找到的相似日不一定具有与预测日最相近的负荷, 有时甚至相差较大, 这是因为存在以下 2 个不确定性因素: 一是影响因素相近的两天具有相近的负荷这条规律具有一定的不确定性, 二是用差异评价函数描述两天影响因素的相近程度存在一定的不确定性; ②由于各个因素共同影响负荷的复杂非线性特性, 用曲线拟合或经验方法来修正由于各个因素的差异而引起的负荷差异, 难以达到好的效果和稳定性。

由于电力负荷变化受天气情况和人们的社会活动等因素的影响较大, 且一般为非线性关系, 而人工神经网络(ANN)能够实现复杂的非线性映射。90 年代以来, 国内外利用 ANN 进行负荷预测已有许多成功的实例, 但大多采用某种网络对负荷与影响因素之间的关系进行大范围映射^[5~7]。误差较大, 而且需要大量的训练样本。

本文研究 96 点的日负荷预测, 即预测次日各时段(15 min/点)的负荷值。兼取相似日方法和 ANN 方法的优点, 提出一种基于相似日和径向基函数网络的新方法。该方法减少了相似日方法对人的经验的依赖, 又保持了其实用有效的特点。开发了实用化软件, 在陕西省网的实际应用表明, 新方法有效改进了普通相似日方法的预测精度和稳定性, 其平均均方根误差较普通相似日方法从 3.153% 降低到 2.954%。

1 基于相似日和径向基函数网络的方法

1.1 基本原理

针对相似日方法存在的第 1 个问题, 由于根据差异评估函数值最小的原则找出的相似日不一定具有与预测日最相近的负荷, 而差异评估函数值不是

最小但是比较小(如次小等)的一批比较相似的负荷日,也有较大可能具有与预测日负荷很相近的负荷,因此,本文提出取一批比较相似的负荷日,而非 1 个相似日。

针对相似日方法存在的第 2 个问题,本文采用 ANN 来描述负荷与各个影响因素之间的复杂非线性关系。但是,ANN 大范围映射往往带来较大的误差,而且需要大量的训练样本。而相似日(包括比较相似的负荷日)毕竟只有几个,因此,本文在文献[8]的基础上,提出采用一种样本量少、具有局部逼近能力的网络——径向基函数网络,将一批比较相似的负荷日的信息输入径向基函数网络进行处理,然后进行预测日的负荷预测。

这种基于相似日(similar day)和径向基函数(radical basis function)网络的短期负荷预测方法简称为 SD & RBF 方法。

1.2 径向基函数网络^[8,9]

径向基函数网络通常是一种 2 层前传网络,其输出层节点的输出为隐节点给出的基函数的线性组合。隐层中的基函数对输入激励产生一个局部化的响应,即仅当输入落在输入空间中一个很小的指定区域中时,隐单元才做出有意义的非零响应。因此,这种网络也被称为局部化接收场网络(localized receptive field network)。

图 1 是径向基函数网络的神经元,其中径向基函数 $\text{radbas}(n)$ 取为:

$$\begin{aligned} \text{radbas}(n) &= \exp(-n^2) \\ n &= \|W - P\|b \end{aligned} \quad (1)$$

径向基函数的输出随输入向量 P 和权向量 W 之间距离的减小而增大,当 P 和 W 完全相同时,输出为极大值 1。因此可以直观地想象径向基函数扮演了信号检测器的角色,阈值 b 可以调节敏感程度。图 2 是径向基函数网络的结构,由一个包含 S^1 个神经元的隐含层和一个包含 S^2 个单元的线性输出层组成。

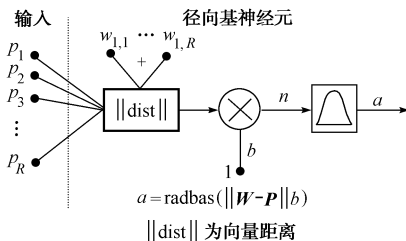


图 1 径向基函数神经元

Fig. 1 A neuron of radial basis function network

径向基函数网络的训练方法有多种,本文限于篇幅只介绍其中的一种——精确训练(exact training)法,其余方法可参见文献[8,9]。

精确训练法的步骤为:

a. 径向基函数层生成 S^1 个单元,并将每个单

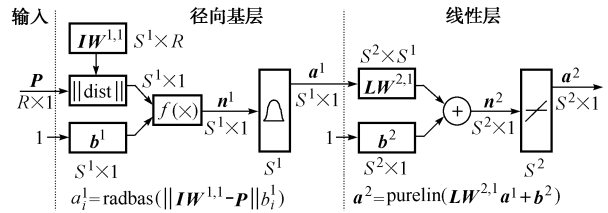


图 2 径向基函数网络的结构

Fig. 2 Structure of a radial basis function network

元的权值向量分别设为和一个输入样本相同。这样,径向基函数层就成为输入样本的检测器,对每一个输入样本,总有一个对应单元的输出为 1。

b. 由步骤 a 可以确定径向基函数层的输出即线性层的输入 a^1 ,目标样本向量已知,网络是线性的。因此,解一个线性方程组就可以确定线性层的权值矩阵。

这种训练方法保证了对于每一个输入样本,网络的输出误差为 0 的目标样本,故称之为精确训练法。

径向基函数网络有以下特点:

a. 很强的泛函逼近能力,原理上能逼近任意的非线性函数。

b. 典型的局部逼近网络,对于每一个输入输出样本,网络中只需调整少量的权值,从而具有训练速度快的优点。

c. 不需要大量样本,在每一个输入、输出样本附近有较好的泛化能力。

1.3 具体步骤

a. 将每日的非数值因素数值化,例如对光照量晴天取 4,多云取 3,阴天取 2,下雨取 1;最高温度和最低温度可取实际值。

b. 为降低问题求解规模,对一天 96 点的每一点分别建立其预测模型。

c. 异常样本处理。考虑前后时段负荷之间的比例,当它出现一个超出常规的数值时,可以肯定发生了异常数据现象,修正值可采用正常比例范围内的估计值,对每天的负荷数据逐一加以检查,便可以异常数据选出并校正。

d. 取预测日前 2 星期的已知日作为待选的相似日,以差异评估函数小于给定的小常数 r 的已知日作为比较相似的负荷日, r 为一经验值。

e. 将这批相似日的影响因素值作为 RBF 网络的输入样本,每一相似日第 i 点负荷值作为输出样本,用精确训练法对 96 个 RBF 网络分别进行训练。由于差异评估函数小于 r 的约束,每一个相似日的影响因素向量与由预测日影响因素向量在向量空间

上都比较接近。利用 RBF 网络的局部泛函逼近的能力,在向量空间中预测日影响因素向量的附近,拟合负荷和影响因素之间的非线性关系。其中 b 的选取一般遵循以下原则^[9]:大于输入样本两两之间向量距离的最小值,小于输入样本两两之间向量距离的最大值。在满足这一原则的前提下,可取几个值试探,选其中预测效果较好者。

f. 将预测日影响因素向量输入 96 个 RBF 网络,即可得到预测日 96 点负荷。最后还应修正预测值。由于负荷受多种不确定因素的影响,纯粹根据历史负荷的预测有时难以令人满意,需要运行人员根据经验进行修正,如阴雨天,可适当将晚峰提前,并将晚峰值提高 1%~3%。有时预测出的值在某个点可能同负荷变化的规律相违背,如急剧增加或减少,在负荷曲线上表现为突然的尖峰或低谷,这种预测值一般不可靠,可以用其前后的预测值的平均值来代替。

本文提出的 SD & RBF 方法有以下几个优点:

a. 充分合理地利用了与预测日比较相似的一批已知日的信息,而不是 1 个已知日的负荷信息,保证了算法较好的稳定性和精度。由于 RBF 网络在样本点附近具有较好的泛化能力,能自动根据已知日的影响因素向量和预测日影响因素向量的距离大小,在输出结果中加以不同的影响。距离近的影响较大;距离远的影响较小。

b. 采用 RBF 网络直接拟合负荷和影响因素之间的非线性关系,不存在修正算法的问题。

c. RBF 网络训练速度快,和相似日方法结合后,整个方法仍保持了应用简便、实用有效的特点。

2 在陕西电网中的应用

海大学与陕西省电力公司调度中心合作开发完成了 96 点短期负荷预测软件。由于不同预测方法各有其特点,所以本软件分别采用 3 种方法对陕西电网的负荷进行预测,即时间序列方法、普通相似日方法和 SD & RBF 方法。然后对 3 种方法的结果进行加权求和,最简单的情况是平均加权。

陕西电网的负荷预测具有如下技术难点:①该省负荷峰谷差较大,因而负荷预测的难度也就较大;②原始负荷数据中存在一定数量的不良数据或错误数据;③温度对负荷的影响大,但对温度的影响难以准确把握;④该省属于干旱地区,旱情直接影响农业灌溉用电从而对负荷有较大影响,而旱情的准确资料难以得到。所开发的软件中对这些问题进行了适当考虑。

日负荷预测的误差指标较多,本文取 4 个相对误差指标作为各种方法预测效果判断的根据。

a. 平均相对误差 E_{MAPE} :

$$E_{MAPE} = \frac{1}{N} \sum \left| \frac{Y_i - \hat{Y}_i}{Y_i} \right|$$
 (2)

其中 Y_i 表示实际负荷; $\hat{Y}_i$ 表示负荷预测值; N 表示一天内负荷数据由 N 个数据点组成,这里为 96 点。

这是一种直观的误差表示方法。

b. 均方根相对误差 E_{MSE} :

$$E_{MSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N \left(\frac{Y_i - \hat{Y}_i}{Y_i} \right)^2}$$
 (3)

E_{MSE} 由于对误差进行了平方,加强了数值大的误差在指标中的作用,从而提高了该指标的灵敏性。

c. 最大负荷相对误差 E_{max} 和最小负荷相对误差 E_{min} :

$$E_{max} = \left| \frac{Y_{max} - \hat{Y}_{max}}{Y_{max}} \right|$$
 (4)

$$E_{min} = \left| \frac{Y_{min} - \hat{Y}_{min}}{Y_{min}} \right|$$
 (5)

采用上述 4 种方法对相同时期每天的负荷进行预测。各方法每天的误差指标见表 1~表 4。

表 1 均方根相对误差
Table 1 Mean square error

				(%)
日期	时间序列方法	相似日方法	SD & RBF 方法	综合方法
2000-04-05	2.37	2.39	2.63	2.39
2000-04-06	2.99	3.02	3.79	3.24
2000-04-07	4.27	4.10	5.26	4.50
2000-04-08	2.96	3.49	2.71	2.74
2000-04-09	3.16	3.12	3.11	3.02
2000-04-10	7.17	3.43	2.97	3.25
2000-04-11	1.59	1.64	2.33	1.56
2000-04-12	1.97	2.05	2.31	2.06
2000-04-13	3.76	3.80	5.72	4.41
2000-04-14	1.96	2.23	1.20	1.61
2000-04-15	3.68	4.57	2.92	3.53
2000-04-16	1.22	1.32	1.79	1.38
2000-04-17	7.40	3.97	2.27	2.53
2000-04-18	2.37	2.16	2.08	2.11
2000-04-19	4.02	4.05	3.41	3.79
2000-04-20	3.16	3.11	2.77	2.96
2000-04-21	2.78	2.81	2.58	2.60
2000-04-22	4.47	5.67	3.67	4.45
2000-04-23	2.88	2.87	3.20	2.94
2000-04-24	2.05	3.54	2.04	1.97
2000-04-25	2.17	2.28	2.22	2.18
2000-04-26	3.01	3.11	3.08	3.05
2000-04-27	2.35	2.27	2.83	2.43
2000-04-28	3.45	3.39	2.86	3.01
2000-04-29	3.41	4.44	4.10	3.65
平均值	3.224 8	3.153 2	2.954 0	2.854 4

表 2 平均相对误差
Table 2 Mean relative error

(%)				
日期	时间序列方法	相似日方法	SD & RBF方法	综合方法
2000-04-05	1.95	1.98	2.14	1.95
2000-04-06	2.66	2.69	3.42	2.92
2000-04-07	3.96	3.77	4.85	4.19
2000-04-08	2.50	2.83	2.17	2.31
2000-04-09	2.76	2.75	2.62	2.60
2000-04-10	6.24	2.79	2.30	2.60
2000-04-11	1.30	1.33	1.81	1.22
2000-04-12	1.64	1.73	1.86	1.71
2000-04-13	3.42	3.46	5.21	4.01
2000-04-14	1.66	1.80	1.01	1.28
2000-04-15	3.49	4.13	2.37	3.28
2000-04-16	1.01	1.07	1.44	1.12
2000-04-17	6.97	3.30	1.82	2.31
2000-04-18	1.87	1.70	1.71	1.69
2000-04-19	3.43	3.45	2.77	3.17
2000-04-20	2.68	2.64	2.28	2.50
2000-04-21	2.14	2.13	2.07	1.94
2000-04-22	3.99	4.96	3.01	3.81
2000-04-23	2.48	2.46	2.71	2.52
2000-04-24	1.40	3.07	1.62	1.62
2000-04-25	1.83	1.90	1.80	1.81
2000-04-26	2.45	2.51	2.53	2.48
2000-04-27	1.88	1.81	2.38	1.99
2000-04-28	2.92	2.87	2.20	2.42
2000-04-29	2.62	3.77	3.45	3.00
平均值	2.771 2	2.676	2.462	2.418

表 3 最大负荷相对误差
Table 3 Relative error of maximum load

(%)				
日期	时间序列方法	相似日方法	SD & RBF方法	综合方法
2000-04-05	5.25	5.25	4.60	4.92
2000-04-06	1.20	1.50	1.20	1.20
2000-04-07	4.01	1.81	5.64	3.69
2000-04-08	2.17	1.23	0.01	0.92
2000-04-09	0.91	1.21	0.91	0.91
2000-04-10	7.84	0.90	2.80	3.12
2000-04-11	1.52	1.21	3.40	2.13
2000-04-12	1.52	1.22	0.60	1.22
2000-04-13	3.03	3.03	4.29	3.66
2000-04-14	2.92	4.58	1.77	3.20
2000-04-15	3.10	3.10	0.60	2.46
2000-04-16	0.30	0.01	0.30	0.01
2000-04-17	6.35	3.46	0.59	0.60
2000-04-18	0.29	0.58	1.78	0.88
2000-04-19	2.98	2.69	3.26	2.98
2000-04-20	1.15	1.15	1.15	1.15
2000-04-21	1.78	0.29	0.59	0.88
2000-04-22	3.63	3.31	1.78	3.00
2000-04-23	2.03	2.03	2.03	2.03
2000-04-24	0.59	2.28	0.58	0.58
2000-04-25	2.91	2.91	2.61	2.91
2000-04-26	1.15	1.15	1.44	1.15
2000-04-27	0.57	0.85	0.57	0.57
2000-04-28	0.29	1.15	1.43	0.86
2000-04-29	1.15	1.18	1.15	0.29
平均值	2.345 6	1.922 8	1.802 8	1.812 4

表 4 最小负荷相对误差
Table 4 Relative error of minimum load

(%)				
日期	时间序列方法	相似日方法	SD & RBF方法	综合方法
2000-04-05	0.01	0.50	0.99	0.01
2000-04-06	3.90	3.90	3.90	3.90
2000-04-07	0.01	0.01	4.74	1.53
2000-04-08	3.05	4.10	2.53	2.53
2000-04-09	1.49	1.49	1.53	1.00
2000-04-10	1.06	3.55	1.04	1.04
2000-04-11	1.97	2.45	0.01	1.49
2000-04-12	1.98	1.98	0.50	1.98
2000-04-13	2.48	2.48	5.61	3.50
2000-04-14	1.91	1.91	0.99	0.97
2000-04-15	3.05	3.57	2.53	3.05
2000-04-16	0.49	0.49	0.50	0.49
2000-04-17	5.18	0.98	0.49	1.00
2000-04-18	0.95	0.01	0.97	0.01
2000-04-19	0.97	1.46	0.95	0.48
2000-04-20	5.16	5.16	4.27	4.72
2000-04-21	0.97	0.97	1.44	0.97
2000-04-22	1.49	2.50	0.49	1.49
2000-04-23	1.90	1.44	1.44	1.90
2000-04-24	1.44	0.49	0.49	0.01
2000-04-25	1.43	1.43	0.48	0.96
2000-04-26	2.83	3.32	3.32	3.32
2000-04-27	2.35	1.89	1.89	2.35
2000-04-28	5.71	5.71	4.72	5.21
2000-04-29	5.66	2.91	5.66	4.76
平均值	2.296 8	2.187 2	2.058 8	1.945 6

分析以上结果可见：

- a. 从这段时间误差平均值来看,综合方法最好,SD & RBF 方法次之,两者明显优于相似日方法,更优于时间序列方法。
- b. 从各天的误差来看,没有一种方法均是最好的,也没有一种方法均是最差的。综合方法的误差均介于其余 3 种方法最好与最差之间。这是由于综合方法采用了等权平均值,对每个预测值同等加权而导致的必然结果。因此可以说,综合方法总体是最好的。
- c. 从各天的误差来看,SD & RBF 方法并非均优于相似日方法和时间序列方法,比如 4 月 5 日~4 月 7 日、4 月 13 日。这主要可能是因为天气信息不准、不全造成的,受条件所限,采用的历史数据是预报值而非实际值。
- d. 从 4 种误差指标来看,均方根相对误差与平均相对误差的大小基本一致,即均方根相对误差大时平均相对误差也大,均方根相对误差小时平均相对误差也小。但与最大负荷相对误差、最小负荷相对误差并不一致,即时间序列方法、相似日方法的误差

大,SD & RBF 方法和综合方法的误差不一定大。这是由于时间序列方法和相似日方法的误差是一天 96 点误差的平均值,而 SD & RBF 方法和综合方法的误差则是一个特殊时间点的误差。

e. 从不同天误差比较来看,某天误差大时各方法一般均较大,误差小时各方法一般均较小,但各方法的最大误差或最小误差并非在同一天。

3 结语

因条件所限,采用的历史气象数据是预报值而非实际值,因此有可能与实际值存在较大误差。而且气象数据不够全面,如没有雨量、光照等定量数据,而只有小雨、多云等描述性气象预报数据。相信如果能进一步增加负荷影响因素的信息量和信息的准确度,预测精度会进一步提高。当然,准确的信息需要一定的经济代价,我们认为这样的投入是必要的,也是有经济效益和社会效益的。

参考文献

- 1 Liu K. Comparison of Very Short-term Load Forecasting Technique. IEEE Trans on Power Systems, 1996, 11(2): 877~882
- 2 刘晨辉(Liu Chenhui). 电力系统负荷预报理论和方法(Theory and Method of Power Systems Load Forecasting). 哈尔滨: 哈尔滨工业大学出版社(Harbin: Harbin Institute of Technology Press), 1987
- 3 侯志俭(Hou Zhijian). 电力系统短期负荷预报的几种改进手段(Several Improvements for Short-term Load Forecasting of Power Systems). 电力系统自动化(Automation of Electric Power Systems), 1996, 20(7):

27~31

- 4 程松(Cheng Song). 一种新的基于相似负荷日的短期负荷预计方法(A New Approach to Load Forecasting Based on Similar Day). 江苏电机工程(Proceeding of Jiangsu Electrical Engineering Association), 1999, 18(4): 28~32
- 5 韩民晓(Han Minxiao). 短期负荷预测方法的研究及在线应用(Studies and Applications of Short-term Load Forecasting Methodology). 电力系统自动化(Automation of Electric Power Systems), 1998, 22(10): 34~37
- 6 Park D C, El-Aharkawi M A, Marks R J. Load Forecasting Using an Artificial Neural Network. IEEE Trans on Power Systems, 1991, 6(6): 442~449
- 7 Lee K Y, Cha Y T, Park J H. Short-term Load Forecasting Using an Artificial Neural Network. IEEE Trans on Power Systems, 1992, 7(1): 534~537
- 8 高山, 单渊达(Gao Shan, Shan Yuanda). 基于径向基函数网络的短期负荷预测(Short-term Load Forecasting Method Based on Radial Basis Function). 电力系统自动化(Automation of Electric Power Systems), 1999, 23(5): 31~34
- 9 徐秉铮, 张百灵, 韦岗(Xu Bingzheng, Zhang Bailing, Wei Gang). 神经网络理论与应用(Neural Network Theory and Applications). 广州: 华南理工大学出版社(Guangzhou: South China Technology University Press), 1994

鞠平, 男, 博士, 教授, 副校长, 主要研究电力系统模拟与控制。

姜巍, 男, 硕士, 主要研究电力系统负荷预测。

赵夏阳, 女, 硕士, 讲师, 主要研究电力系统优化调度。

NINETY-SIX POINTS SHORT-TERM LOAD FORECASTING—THEORY & APPLICATIONS

Ju Ping¹, Jiang Wei¹, Zhao Xiayang¹, Wang Junkai², Zhang Shixue², Liu Yan²

(1. Hohai University, Nanjing 210098, China)

(2. Shanxi Provincial Power Company, Xi'an 710004, China)

Abstract: Based on similar day method and radial basis function network, this paper proposes a new method for short-term load forecasting. The new method uses the parameters of several more similar days, instead of only one similar day in similar day method, as the inputs to radial basis function network for forecasting the loads of 96 points (15 minutes per point) of the next day. The method behaves the advantages of both similar day method and radial basis function network method, i. e. simple, practical, accurate and experience unreliable. The software developed has run well for over one year in Shanxi Provincial Power Company. The practical application results show that the mean error reduces from 3.153% to 2.954% compared with conventional similar day method.

This project is supported by Trapoyt.

Keywords: power systems; load forecasting; similar day method; radial basis function network