

文章编号: 1001-4179(2005)11-0068-03

多元线性网络法在区域电力负荷预测中的应用

万星^{1,2} 丁晶¹ 张晓丽²

(1. 四川大学 水电工程学院, 四川 成都 610065; 2. 成都信息工程学院, 四川 成都 610041)

摘要: 区域电力负荷具有特殊的变化规律, 但表现出一定的周期波动性。对区域电力负荷变化影响因素分析后, 分别利用多元线性回归和人工神经网络模型对电力负荷变化进行了电力消耗预测。对二者预测结果还进行了误差分析, 并提出了高、中、低3个水平的电力负荷消耗。结果表明: 预测结果与当地电力负荷消耗增长规律相符, 且人工神经网络预测结果较多元线性回归模型稍好。预测的结果数据可以作为当地决策部门的资料参考。

关键词: 电力负荷; 多元线性回归; 人工神经网络; 预测
中图分类号: TP183 **文献标识码:** A

近年来, 随着电力市场化的进一步深入, 未来电力负荷以及供电需求预测成为众多学者的研究热潮, 电力负荷预测方法主要分为两类: 即传统预测方法和基于人工智能的预测方法。在基于人工智能的方法中, 人工神经网络^[1-2] (ANN) 由于无需预先给出数学模型和具有强非线性拟合能力而被广泛采用。然而, 对于 ANN 这类数据驱动型预测技术, 如何选取最能反映电力负荷动态变化信息的输入变量将直接关系到预测的精度和鲁棒性。由于负荷的变化有一定规律又有很强的随机性, 不同地区用户的组成、一年的不同时期、经济因素以及温度及其他气候条件, 甚至当地的风俗习惯均会导致负荷曲线呈现出变化的特征。事实上不同区域和不同季节可能呈现出完全不同的负荷特征, 因此也不存在普适的电力负荷预测输入变量。对于一个特定的区域而言, 只有经过负荷特征识别才能制定特定的模型自适应预测方案, 实现对日负荷特性的长期、可靠预测。

电力负荷具有特殊的变化规律, 但主要表现为两种不同的趋势^[3-4]: ① 逐渐增长的趋势; ② 以天、周、月、年为周期波动。在此基础上, 影响负荷变化的因素主要包括: 负荷构成、负荷随时间变化的规律、气象变化以及负荷的随机波动, 按照以年为周期来说, 还包括当地经济发展目标、产业发展以及人口增长规模等等因素。根据短期负荷预测的性质, 知其基本变化规律可由线性变化模型和周期变化模型来描述, 同时还存在随机负荷预测模型。另外, 气温、阴晴、降水和大风等都会引起负荷的变化。因此, 负荷变化将存在波动性, 所以各类电力需求信号一般都表现为连续波动、随机性。

本文采用多元线性回归与神经网络相结合的预测方法, 对四川电力的历年负荷曲线特征进行了识别, 实现了未来中长期电力需求预测的动态模拟, 为当地决策部门提供参考。

1 多元线性回归预测模型

因素变量之间的关系一般分为两类: ① 确定性关系, 也可

叫函数关系; ② 相关关系, 其特征主要表现为: 各个变量之间存在着一定关系, 但是这种关系很难用一种精确的方法表示出来。比如, 人的身高与腿长之间有一定关系, 但是身高与腿长之间没有确定性直接关系, 由身高不能确定计算出腿长, 由腿长也不能确定计算出身高。又如, 农作物化肥需求量与农作物产量之间的关系等等。由上述可见, 这些变量之间的关系是一种相关关系, 而变量之间的这种相关关系通常可用回归分析的方法来处

理。多元线性回归^[5-6]的目的是建立预报变量组 (X 的序列阵) 与响应变量 Y 之间的一种定量关系。其主要目的是用于分析出哪些预报因素是主要因素, 且还能够确切知道该影响变量的变化趋势以及用所建立的模型去预测未来的响应值。多元线性回归的一般结构为:

$$y = X\beta + \epsilon \quad (1)$$

式中 y 为 $n \times 1$ 阶观测序列值; X 为 $n \times p$ 阶回归矩阵; β 为 $p \times 1$ 阶参数向量; ϵ 为 $n \times 1$ 阶随机干扰向量值。

利用最小二乘法, 其最小二乘估计为:

$$Y = Xb = X(X'X)^{-1}X'Y \quad (2)$$

其中该模型问题的解就是一个向量 b , 即为估计未知参数 β , 为:

$$b = \beta = (X'X)^{-1}X'Y \quad (3)$$

模型中与随机干扰误差 ϵ 对应的即为模型预测残差, 指观察值与预测值之差, 它主要用于探测模型假设是否合理。

2 人工神经网络模型预测

BP 网络是目前预测领域中被广泛应用的一种方法^[7,8], 其多采用多个神经元函数来描述多因子之间复杂的输入输出关系。

一个多输入、单输出的 3 层神经网络模型的数学表达式为:

$$y = \sum_{i=1}^m w_i \varphi_i \left(\sum_{j=1}^p w_{ij} X_j - \theta_i \right) \quad (4)$$

收稿日期: 2005-08-29

基金项目: 国家自然科学基金资助项目 (40271024; 50279023)

作者简介: 万星, 男, 四川大学水电工程学院, 博士研究生; 成都信息工程学院, 讲师。

式中 m 为隐层节点的个数; p 为输入层节点个数, 即系统输入变量的个数; w_i 为第 i 隐层节点到输出节点的权值; φ 为隐层中第 i 个神经元函数, 一般为 S 型函数: $\varphi(x) = 1/(1 + e^{-x})$; w_{ji} 为第 j 输入节点到第 i 个隐层节点的权值; θ_i 为第 i 个神经元的域值。

3 区域电力负荷预测研究

以四川电力负荷为例, 因西部大开发战略的实施, 为当地经济发展提供了难得机遇。区域经济发展, 能源需求尤其是电力需求就显得尤为重要。因此, 为实现区域经济的可持续发展, 就要对未来电力负荷需求进行预测规划, 以保证提供当地的能源需求, 就成为经济发展的关键环节。

经过各种努力, 通过多种渠道, 广泛收集四川历年国民经济主要指标及其有关资料⁹。具体见表 1 所示。经采用各种指标分析后, 得出具体的数据资料如表 1 所示。

表 1 四川省年用电量变化特性分析资料

年份	国民总产值/ 亿元	第一产业/ 亿元	第二产业/ 亿元	第三产业/ 亿元	总人口/ 万人	年用电量/ (亿 kW·h)
1978	185.76	82.20	65.55	38.01	7 071.9	80.13
1979	205.76	91.95	72.31	41.50	7 120.5	90.78
1980	229.31	101.68	81.05	46.58	7 154.8	93.95
1981	242.32	108.02	83.36	50.94	7 215.6	94.25
1982	275.23	125.36	92.84	57.03	7 300.4	100.29
1983	311.00	138.17	105.69	67.14	7 336.9	108.15
1984	358.05	156.11	121.68	80.27	7 364.0	116.37
1985	421.15	172.90	148.11	100.14	7 419.3	124.92
1986	458.23	181.20	160.62	116.41	7 511.9	129.83
1987	530.86	202.25	187.88	140.73	7 613.2	151.12
1988	659.69	241.95	238.32	179.42	7 716.4	159.92
1989	744.98	263.15	266.16	215.67	7 803.2	172.67
1990	890.95	321.41	313.64	255.90	7 892.5	177.11
1991	1 016.31	339.00	378.48	298.83	7 947.8	188.15
1992	1 177.27	372.04	441.57	363.66	7 992.2	194.37
1993	1 486.08	449.38	580.38	456.32	8 037.4	218.91
1994	2 001.41	597.37	802.77	601.27	8 098.7	247.55
1995	2 504.95	725.46	1 020.91	758.58	8 161.2	273.50
1996	2 985.15	860.02	1 229.01	896.12	8 215.4	292.10
1997	3 320.11	919.28	1 358.38	1 015.45	8 264.7	299.17
1998	3 580.26	941.24	1 527.07	1 111.95	8 315.7	304.08

注: 1997 年前包含重庆地区。

根据表 1 所示数据, 首先用简单因素分析法进行因素分析, 具体见图 1 所示。

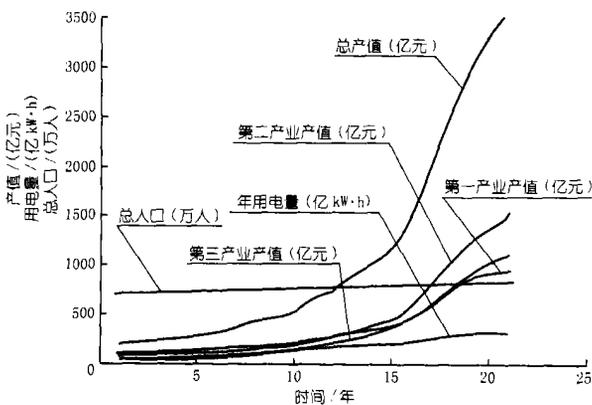


图 1 年用电量与其主要影响因素

从图 1 中可以看出, 年电力负荷和国民生产总值等指标序列的时序过程线基本上是上升趋势。通常说来, 影响电力消耗的因素错综复杂, 但可以进行科学分析, 寻其主次因素, 图 1 表明, 国民生产总值, 第一、二、三产业总值是影响电力消耗的主要因素。但这些因素的影响作用也较为复杂, 关系不明显, 因此首先考虑用(1)式的多元线性回归模型进行分析。

3.1 电力负荷的多元线性回归模型预测

根据电力负荷需求的历年统计关系, 发现区域电力需求主要与当地的国民生产总值, 第一、第二、第三产业总值以及人口因素等有关。电力负荷需求预测用多元线性回归模型分析方法, 就是根据过去的历史资料确定电力需求与其主要因素之间的线性依赖关系, 从而实现对其电力负荷需求的预测估计。

设第 t 年的国民生产总值和第一、第二、第三产业产值以及总人口分别记为 $y(1, t)$, $y(2, t)$, $y(3, t)$, $y(4, t)$ 和 $y(5, t)$, 它们都为影响因素, 年电力负荷消耗记为 $x(t)$, 则用多元线性回归模型为:

$$x(t) = c(0) + c(1)y(1, t) + c(2)y(2, t) + c(3)y(3, t) + c(4)y(4, t) + c(5)y(5, t) \quad (5)$$

式中 $c(j)$ ($j = 0 \sim 5$) 为模型参数。该模型的建立过程就是确定这些模型参数, 参数估计的原则是残差绝对值的和最小, 本文采用公式(2)与(3)来确定模型参数。

根据表 1 所示历年电力负荷数据, 用公式(2)与(3)的方法来对模型进行优化求解, 可以得到:

$$\begin{cases} c(0) = -210.5 \\ c(1) = 0.015 \\ c(2) = 0.068 \\ c(3) = 0.034 \\ c(4) = 0.030 \\ c(5) = 0.044 \end{cases} \quad (6)$$

利用式(6)的这些参数, 又对表 1 中的数据进行了重新拟合, 具体结果见表 3。

3.2 电力负荷的人工神经网络模型预测

考虑用线性回归模型要求变量之间的线性关系要较好, 但从图 1 中可以发现, 电力负荷影响的主要因素的非线性关系不明显。因此, 鉴于神经网络模型的非线性映射能力, 所以考虑用人工神经网络(ANN)来进行预测估计。

首先以国民生产总值、前一年用电量作为输入变量, 再轮流预测当年的电力负荷。因神经网络模型是用于长期预测的, 因此在考虑建立模型时不仅要考虑拟合效果, 而且还要特别考虑预测效果。由于原有资料各变量最大值对原始资料初始变化不具有大幅度外推预测能力, 因此在建立模型时必须预先估计各变量可能最大值, 再对原始数据进行初始化。最大值的选择直接影响模型拟合和预测效果, 选择原则是要求评定系数后验差比值小和小误差概率大, 另外要保证结果呈现较合理的增长趋势。模型结构用 5 个影响因素为输入变量, 以年电力负荷消耗为目标值, 建立 3 层 BP 模型网络为 $5 \times 4 \times 1$ 结构, 隐含层含 4 个神经元, 选取输入层神经元数的 75% 为作为隐含层数¹⁰。全部数据采用批处理(batch processing)方式输入网络。采用 L-M 规则训练, 精度达到 0.001 停止训练, 训练次数为 5 000 次, 训练 1 210 次就达到训练精度。

模型的拟合结果如图 2 所示, 评判模型效果的指标分别为后验差比值 $c = 0.023$, 小误差概率 $p = 98\%$ 。

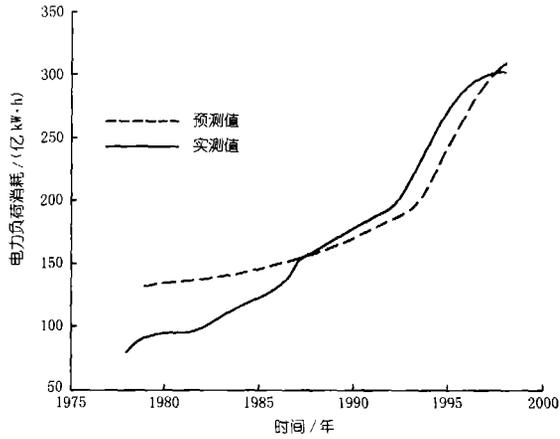


图2 BP模型拟合结果

上述两种模型的预测计算值见表2示。

此外,根据上述两种模型预测拟合的结果分析,还进行了长期的电力负荷消耗预测,结果见表3所示。

表2 电力负荷消耗预测结果 亿 kW·h

年份	实际值	多元回归模型	人工神经网络模型	年份	实际值	多元回归模型	人工神经网络模型
1978	80.13	98.19	80.13	1989	172.67	160.41	168.34
1979	90.78	101.49	132.11	1990	177.11	172.82	176.31
1980	93.95	104.34	135.80	1991	188.15	181.33	179.55
1981	94.25	107.71	136.96	1992	194.37	191.50	187.38
1982	100.29	113.39	137.15	1993	218.91	209.95	193.09
1983	108.15	116.99	139.47	1994	247.55	240.91	214.39
1984	116.37	120.89	142.64	1995	273.50	270.58	243.05
1985	124.92	126.63	146.17	1996	292.10	299.17	273.56
1986	129.83	132.46	149.96	1997	299.17	318.21	298.12
1987	151.12	140.71	152.40	1998	304.08	332.60	309.26
1988	159.92	152.23	163.28				

表3 各模型电力负荷预测的计算值 亿 kW·h

时间/年	多元线性回归模型估计水平			人工神经网络模型估计水平		
	高	中	低	高	中	低
2005	543.58	506.63	467.24	357.43	347.94	339.64
2010	825.50	719.49	614.08	522.98	468.73	426.46
2015	1 318.66	1 063.12	827.19	1 053.00	830.37	657.64
2020	2 195.86	1 626.61	1 139.94	1 726.51	1 548.28	1 241.06

为了预测四川未来20多年的电力负荷消耗,考虑到影响经济和人口增长的因素极其复杂,不确定性成分较大。又为全面考虑多种可能性,并使最终决策能多方案选择,在估计经济、人口增长率时,分为3种情况:即乐观估计、平均估计和悲观估计。具体预测结果见表3所示。

根据表3中所示的预测结果,用曲线表示后见图3、图4所示。

从图3、图4中可以看出,两种模型预测结果都有较为显著的增长趋势,这与国民经济的发展对电力总需求一致,表明当地经济发展分为3个不同发展速度来体现。但是,人工神经网络预测结果较多元线性回归模型预测结果更为平缓,说明人工神经网络的非线性影射能力更好。

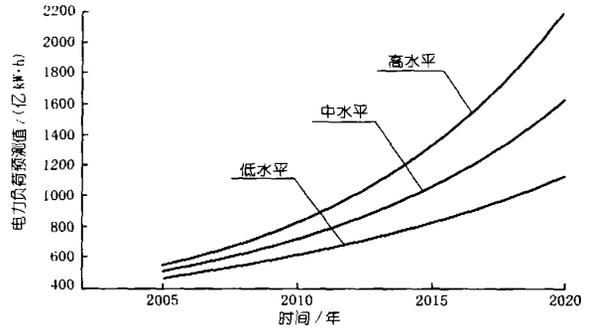


图3 多元回归模型预测结果

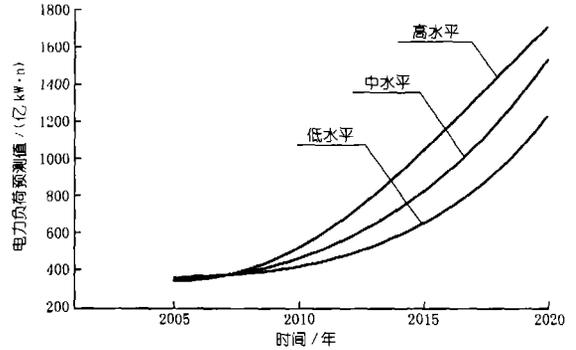


图4 BP模型预测结果

4 结论

文中分别运用多元线性回归模型和人工神经网络模型进行当地电力负荷消耗预测,预测结果较好。但是,多元线性回归主要体现因素之间的线性关系,而人工神经网络则具有较强的非线性影射能力,预测结果表明人工神经网络能较好的拟合非线性特性。还表明二者结合能够获得较为真实的预测数据,为当地决策部门指定经济发展政策提供决策参考。

参考文献:

- [1] 彭清娥,刘兴年,曹叔尤.流域年均含沙量的人工神经网络模型.水利学报,2001,(11):79~83.
- [2] R Hecht-Nielsen. Application of counter propagation networks. Neural Networks Vol. 1; 131~139.
- [3] 邵能灵,侯志俭等.基于小波分析的电力系统短期负荷预测方法.中国机电工程学报,2003(1).
- [4] 谢敬东,唐国庆.组合预测方法在电力负荷预测中的应用.中国电力,1998,(6):3~5.
- [5] 何仁斌编著. MATLAB 6 工程计算及应用.重庆:重庆大学出版社,2001:226~238.
- [6] 唐嗣政编著.概率论与数理统计.成都:成都科技大学出版社,1995:144~202.
- [7] 聶,刘权授.随机水文学.北京:中国水利水电出版社,1997:17~28.
- [8] 聶.人工神经前馈网络模型用作过渡期径流预测的探索.水电站设计,1997,13(2):69~74.
- [9] 四川省统计局.统计年鉴.北京:统计出版社,1999:11~24.
- [10] Paya BA, Esat II, Badi MNM. Artificial neural network based fault diagnostics of rotating machinery using wavelet transforms as a pre-processor. Mech Sys and Signal Processing, 1997 11(5): 751~765.

(编辑:徐诗银)