

# 小波网络和 BP 网络在负荷预测中的比较研究

冯再勇<sup>1,3</sup>, 钟康惠<sup>2</sup>, 马永旺<sup>1</sup>

(1. 成都理工大学 信息管理学院, 成都 610059; 2. 成都理工大学 地球科学学院, 成都 610059  
3. 苏州机电高等职业技术学校, 江苏 苏州 215031)

**摘要:** 从收敛速度、网络稳定性、泛化性能及预测结果等方面分析BP网络和小波神经网络的预测过程, 探讨了差异的原因. 发现: (1) BP网络预测稳定性强, 易取舍数值, 但预测速度较慢, 预测精度不够理想; (2) 小波网络预测速度快, 精度高, 泛化性能好, 但稳定性能差, 实际运用中还没有合适的数值取舍方法; (3) 小波网络总体超过BP网络, 但实际预测时技术问题较大.

**关键词:** BP网络; 小波网络; 收敛速度; 稳定性; 泛化性能; 预测精度

**中图分类号:** TM715 **文献标识码:** A **文章编号:** 1673-0143(2007)02-0051-05

精确的电力系统负荷预测既有利于电力部门合理安排机组起停, 减少旋转备用容量, 科学安排检修计划, 降低发电成本, 提高经济效益, 也有利于用电大户有效制订生产计划和准确分配生产费用. 然而统计技术、专家系统、模糊预测等传统的负荷预测方法均有其不足<sup>[1-4]</sup>. 神经网络是目前人们普遍看好的预测工具, 以BP网络及其改进形式<sup>[5-8]</sup>和小波神经网络<sup>[9-11]</sup>应用较为广泛. 本文结合实例, 对BP网络和小波网络的预测过程及效果进行了比较研究.

## 1 小波神经网络及其特点

### 1.1 小波理论

小波理论源于傅立叶分析. 经傅立叶变换处理后的信号只能在时域或者频域进行分析. 经小波平移和伸缩变换处理后的信号, 可在时频域对其局部细节进行多分辨分析. 小波因而被形象地称为“数学显微镜”.

小波变换的变换因子一般由连续小波函数

$$\Psi_{(a,b)}(t) = \frac{1}{\sqrt{|a|}} \Psi\left(\frac{t-b}{a}\right) \quad (1)$$

给出, 其中  $\Psi(t) \in L^2(R)$ ,  $a$  为伸缩尺度,  $b$  为平移尺度.

若  $\Psi(t)$  满足可允许性条件

$$C_\Psi = \int_{-\infty}^{+\infty} \frac{|\Psi(v)|^2}{|v|} dv < +\infty, \quad (2)$$

其中  $\Psi(v) = \int_{-\infty}^{+\infty} \Psi(t)e^{-ivt} dt$ , 即函数  $\Psi(t)$  的傅立叶变换, 那么  $\Psi(t)$  称为小波母函数, 它可以进行小波变换. 对于任意  $f(t) \in L^2(R)$ , 其小波变换如下:

$$W_{f(a,b)} = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{+\infty} f(t) \overline{\Psi\left(\frac{t-b}{a}\right)} dt. \quad (3)$$

(2) 式可允许性条件的重要意义在于, 如果  $\Psi(t)$  的傅立叶变换  $\Psi(v)$  在原点  $v=0$  处连续, 则必有  $\Psi(0)=0$ , 即  $\int_{-\infty}^{+\infty} \Psi(t) dt = 0$ . 这说明函数  $\Psi(t)$  有“波动”的特点. 又由于  $\Psi(t) \in L^2(R)$ , 于是  $\Psi(t)$  只在原点附近才有明显的起伏, 而在远离原点的地方必将迅速向零衰减, 所以它被称为“小波”<sup>[12]</sup>. 从物理意义来看, 参数  $b$  决定  $\Psi_{(a,b)}(t)$  的波动时间, 参数  $a$  决定波动的幅度, 这体现了小波变换“时域-频域”分析的思想.

已经证明, 连续小波函数的线性组合在  $L^2(R)$  中是稠密的, 即, 其线性组合可以逼近任意  $f(t) \in L^2(R)$ , 这从理论上严密的保证了小波网络逼近的可靠性.

### 1.2 小波神经网络

小波神经网络同时继承了小波分析及神经网络的优点<sup>[13]</sup>, 是逼近单变量函数的最佳逼近器<sup>[14]</sup>. 其常见小波网络的结构特点和训练算法如下:

网络结构和表达式和BP网络基本一致, 即由输入层、隐含层、输出层构成. 不同之处主要

收稿日期: 2006-09-18

基金项目: 国家“973”计划资助项目(2002CB412607)

作者简介: 冯再勇(1982-), 男, 江苏南京人, 硕士生, 主要从事非线性理论研究.

在于, BP网络隐含层神经元的激励函数取 Sigomoid 函数  $f(x)=1/(1+e^{-x})$ , 小波网络则采取满足可允许条件的小波函数  $\Psi(t)$  为激励函数, 结构如图 1 所示,  $\Psi(t)$  的具体取法可以视实际需要进行选择. 输出层的激励函数仍然采用 Sigomoid 函数.

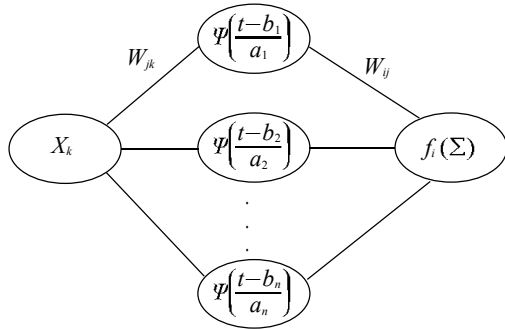


图 1 小波网络结构

网络训练过程基于误差逆传播思想, 按梯度下降方向调整权值  $w$ , 及小波参数  $a$ 、 $b$ . 由于隐含层采取了不同的激励函数, 因而在调整权值和小波参数时, 所采用的算法有所变化. 小波网络的具体实现过程如下<sup>[15]</sup>:

设  $x_k$ —输入模式向量;

$y_i$ —输出模式向量;

$w_{ij}$ —中间隐层到输出层的权值;

$w_{jk}$ —输入层到中间隐层的权值;

$d_i$ —输出向量模式  $y_i$  相应的期望输出.

输入层、隐含层及输出层的神经元个数分别为  $m$ 、 $n$ 、 $N$  个, 于是在  $t$  时刻

$$y_i(t) = f\left(\sum_{j=0}^n w_{ij} \Psi_{(a,b)}\left(\sum_{k=0}^m w_{jk} x_k(t)\right)\right), \quad i=1, 2, \dots, N,$$

表 1 山东省 1999~2003 年用电量

年份	第一产业GDP /亿元	第二产业GDP /亿元	第三产业GDP /亿元	总人口 /万	用电总量 /亿 kw·h
1999	1221.00	3705.44	2735.66	8883	914.96
2000	1268.57	4226.57	3047.30	8975	1000.85
2001	1359.49	4654.51	3424.31	9024	1104.47
2002	1390.00	5309.00	3851.00	9082	1241.84
2003	1505.00	6650.00	4275.00	9125	1395.72

同时建立 BP 网络和小波网络, 分别利用已有数据对网络进行训练. 为了更具有代表性, 小波网络中的小波函数采取常用的 Morlet 小波函数  $f(x) = \cos(1.75x) \times \exp(-x^2/2)$ , 这是已被证明满足可允许条件的小波函数. 网络运行的具体实现利用 Matlab 7.1 编制相应的程序并运行, 训练过程中,

$$E = \frac{1}{2} \sum_{i=1}^N (y_i(t) - d_i)^2, \quad E \text{ 为误差函数.}$$

若令  $net_j = \sum_{k=0}^m w_{jk} x_k(t)$ , 则有  $\Psi_{(a,b)}(net_j) = \Psi\left(\frac{net_j - b_j}{a_j}\right)$ ,

于是  $y_i(t) = f\left(\sum_{j=0}^n w_{ij} \Psi_{(a,b)}(net_j)\right)$ ,  $i=1, 2, \dots, N$ ,

$$f(x) = \frac{1}{1+e^{-x}}.$$

在梯度下降的思想下相应的参数调整过程如下:

$$w_{jk}(t+1) = -\eta \frac{\partial E}{\partial w_{jk}} + w_{jk}(t), \quad (4)$$

$$w_{ij}(t+1) = -\eta \frac{\partial E}{\partial w_{ij}} + w_{ij}(t), \quad (5)$$

$$a_j(t+1) = -\eta \frac{\partial E}{\partial a_j} + a_j(t), \quad (6)$$

$$b_j(t+1) = -\eta \frac{\partial E}{\partial b_j} + b_j(t). \quad (7)$$

其中  $\eta$  为学习系数. 公式的具体推导过程参见文献 [15].

在网络的权值和阈值随机确定后, 首先进行网络前向传播, 然后进行误差的计算以及逆传播的计算, 此过程中需要利用 (4)~(7) 式对网络的参数进行调整, 反复训练, 直到满足终止条件.

## 2 小波网络及 BP 网络在负荷预测中的应用实例

本文以山东省年用电量为例<sup>[16]</sup> (见表 1), 环境变量采取和年用电量相关程度比较高的 4 个因素: 当年的第一, 第二, 第三产业及全省的人口总量.

两种网络取相同的学习速率  $\eta=0.2$ , 隐含层神经元个数为 5 个, 控制误差限为  $10^{-18}$ . 每次实验时采取“批训练”方式, 进行轮回训练, 以减小全局误差. 两种网络的收敛速度及预测结果如表 2、表 3 所示.

表2 BP网络和小波网络的收敛速度比较

年份	1999	2000	2001	2002	2003
BP网收敛次数	3217	2462	1962	1684	1351
小波网络收敛次数	6818	1452	619	668	679

表3 BP网络模型方法和小波网络的结果对比

年份	实际用电/亿kw·h	BP网络		小波网络	
		预测值/亿kw·h	相对误差/%	预测值/亿kw·h	相对误差/%
2002	1241.84	1092.65	-12.01	1291.23	3.98
2003	1395.72	1219.29	-12.64	1332.61	-4.52

### 3 实验结果对比分析

由于两种网络算法本身要求初始权值、阈值等随机化,因此每次的训练结果会有不同,这里采取“概率意义”下较为稳定的结果对BP网络及小波网络的实验过程进行比较。

#### 3.1 收敛速度

由于两种网络在计算过程中分别采取了不同的函数作为隐含层神经元,而且为了使误差函数下降更快,小波网络中还要将伸缩,位移尺度向误差的负梯度方向进行调整及更新迭代,因此两种网络的固有特性决定了对收敛速度的比较不能以“时间代价”作为衡量标准.用首轮直到收敛为止的训练次数作为速度标志更为合理.由表2很容易看出,小波网络只是在初次训练时收敛较慢,一旦初次收敛,“记忆了”非线性函数后,以后的训练中表现出明显的优势,在这里的网络条件下,收敛速度几乎为BP网络的2~3倍.这正说明,小波网络在提取信息的非线性特征时需要较多的“时间代价”,一旦提取完毕后,对这种非线性函数的记忆能力极强.事实上,小波网络加入了伸缩参数和平移参数.一方面,这两个参数在提取、学习函数非线性特征时要“耗费”一定的时间,从而造成了初次收敛时间较BP网络“落后”;另一方面,这两个参数也增加了小波网络记忆非线性函数的能力,因此后面的训练中,收敛时间较BP网络明显“超前”。

#### 3.2 网络稳定性

在网络训练过程中,发现BP网络在收敛速度上处于劣势,然而其稳定性却明显胜出小波网络.BP网络的稳定性优势主要表现在以下两个方面:一方面,第一轮训练中小波网络除了在逼

近第一个样本目标时,收敛次数相对稳定外,后面的逼近过程中,收敛时的训练次数没有BP网络集中,“离散度”较为严重,这里给出的是概率意义下具有代表性的结果;另一方面,网络第一轮训练结束后,BP网络每轮的各个目标收敛时,训练次数几乎没有多大变化,而小波网络则变化显著,甚至有较大的“反弹”,过学习现象稍强.同时,由于网络的初始值是随机选取的,可以看出这反映了小波网络对初值的选取比较敏感.从逼近论的角度来看,BP网络用于函数逼近的“基” $f(x) = 1/(1 + e^{-x})$ 始终保持不变,而要逼近的函数虽然是隐含的,但却是确定的,因此在“基”和“逼近目标”都不变的情况下,逼近的过程也比较稳定.而小波网络用于函数逼近的“基” $\Psi((net_j - b_j)/a_j)$ 对于尺度因子、伸缩因子的初始值比较敏感.随机选取的 $a_j$ 、 $b_j$ 初始值不同,后面的更新迭代也会变化较大.“基”的变动使得对于目标函数的逼近过程波动较大,直观表现就是网络不稳定。

#### 3.3 网络泛化性能

对比研究两种网络泛化能力,发现小波网络的泛化能力具有一定的比较优势.每次实验中,首先用两种网络分别对同样的样本进行训练至收敛,再加入一个新的样本分别训练至收敛,然后对前面的样本进行仿真,检查新输出与原始值的“偏离”程度,以便说明新样本的训练对前面已经记忆的样本的影响.为了更好地显示两种网络的性能,采取多次实验.实验显示BP网络具有很好的稳定性,BP网络各个样本在不同次实验中的新输出基本一致,但与原始值偏差较大,且大多偏向原始值的同一侧.而小波网络虽然有一定的波动性,但在统计意义下,只在原始值小范

围内波动. 用数学语言可以描述为: 若设前  $k$  个样本对应的原始输出分别为  $y_1, y_2, \dots, y_k$ , 分别用两种网络进行  $p$  轮训练, 由 BP 网络进行第  $p$  次泛化训练后得到的新输出分别为  $b_1^{(p)}, b_2^{(p)}, \dots, b_k^{(p)}$ , 由小波网络得到的相应值为  $w_1^{(p)}, w_2^{(p)}, \dots, w_k^{(p)}$ , 则有

$$\sum_{p=1}^P |b_i^{(p)} - y_i| > \sum_{p=1}^P |w_i^{(p)} - y_i|, \quad i=1, 2, \dots, k, \quad (8)$$

$$\text{以及} \left| \sum_{p=1}^P (b_i^{(p)} - y_i) \right| > \left| \sum_{p=1}^P (w_i^{(p)} - y_i) \right|, \quad i=1, 2, \dots, k, \quad (9)$$

小波网络中, 实验观察显示近似的有  $y_i \approx \frac{1}{P} \sum_{p=1}^P w_i^{(p)}$ , 显然这个结果和式 (8)、式 (9) 能够很好的吻合. 于是有理由猜想是否有事件  $y_i = \lim_{p \rightarrow \infty} \frac{1}{P} \sum_{p=1}^P w_i^{(p)}$  ( $i=1, 2, \dots, k$ ) 依概率收敛, 即对于任意的  $\varepsilon > 0$ , 有  $\lim_{p \rightarrow \infty} P \left\{ \left| \frac{1}{P} \sum_{p=1}^P w_i^{(p)} - y_i \right| < \varepsilon \right\} = 1, \quad i=1, 2, \dots, k$ .

这一猜想的具体论证, 有待作一步的研究.

### 3.4 预测精度

将 1999~2001 年的已知数据作为网络的训练样本, 预测 2002 年的用电量, 同样将 2000~2002 年的数据作为预测 2003 年用电量的样本, 分别用两种网络进行训练、预测. 两种网络用于预测的理想结果如表 3 所示. 显然小波网络预测的精度比 BP 网络有了大幅提高, 印证了小波函数的非线性程度更高, 对非线性特征的提取能力, 记忆能力也更强, 泛化性能较好. 然而实际用于预测的过程中, 一方面, 为了改善小波网络的不稳定性以及网络的泛化性能, 每次实验要求尽量进行多轮训练以减小全局误差, 而多轮训练又会产生“过学习”现象, 这是一对矛盾. 另一方面, 不同的实验中由于网络随机选取的各参数  $w_{ij}, w_{ik}$  及  $a_j, b_j$  初始值不同, 训练结束后进行的预测结果也会有差异. 增加训练次数可以得到统计意义上比较可信的结果, 然而进行多少次实验比较合适, 研究还不够充分, 这就带来了决定数值取舍的技术问题.

## 4 结论与展望

综上所述, 可以得到以下结论:

1) 用 BP 网络预测负荷, 稳定性强, 易取舍数值, 但预测速度较慢, 预测精度不够理想;

2) 用小波网络预测速度快, 泛化性能好, 精度高, 但稳定性能差, 实际运用中还没有合适的数值取舍方法;

3) 预测时, 理论上小波网络总体优于 BP 网络. 但实践上, 尚存在技术问题, 优势难显.

因此, 如果能够找到一种逼近非线性目标能力足够强, 且稳定性能更好的小波函数, 将会给小波网络的实际运用带来新的生机.

### 参考文献:

- [1] 飞思产品研发中心. 神经网络理论与 Matalab7 实现[M]. 北京: 电子工业出版社, 2005.
- [2] 孙振, 路洋. 电力系统预测方式综述[J]. 黑龙江电力, 2005, 27(4): 260-262, 299.
- [3] 张步涵, 刘小华, 万建平, 等. 基于混沌时间序列的负荷预测及其关键问题分析[J]. 电网技术, 2004, 28(13): 32-35, 49.
- [4] 沈百新. 利用专家系统预测地区用电负荷[J]. 电力需求管理, 2005, 7(2): 49-50.
- [5] 杨奎河, 王家宝, 赵玲玲, 等. 神经网络矫正的非线性短时负荷预测模型[J]. 系统工程与电子技术, 2004, 26(11): 1710-1713.
- [6] 何飞, 张国忠, 刘亚, 等. 基于 BP 网络的改进负荷预测方法[J]. 华东电力, 32(3): 31-33.
- [7] 阮萍, 骆力明, 王华, 等. 基于灰色系统和人工神经网络的中长期电力负荷预测[J]. 首都师范大学学报(自然科学版), 2004, 25(2): 22-25.
- [8] 姚李孝, 宋玲芳, 李庆宇, 等. 基于模糊聚类分析与 BP 网络的电力系统短期负荷预测[J]. 电网技术, 2004, 29(1): 20-23.
- [9] 董景荣. 基于小波网络的电力系统短期负荷预报研究[J]. 预测, 2000, (4): 66-69.
- [10] 王家红, 黄阿强, 熊信良. 基于小波网络的短期负荷预测方法[J]. 电力自动化设备, 2003, 3(3): 11-13.
- [11] 张步涵, 赵剑剑, 刘小华, 等. 一种基于小波神经元网络的短期负荷预测方法[J]. 电网技术, 2004, 28(7): 15-18.
- [12] 冉启文. 小波变换与分数傅里叶变换理论及应用[M]. 哈尔滨: 哈尔滨工业大学出版社, 2001.
- [13] 陈哲, 冯天瑾, 陈刚, 等. 一种基于 BP 算法学习的小波神经网络[J]. 青岛海洋大学学报, 2001, 31(1): 122-128.
- [14] Kreinovich V, Sirisaengtaksin O, Cabren S. Wavelet neural networks are asymptotically optimal approximators for functions of one variable[A]. Florida, USA: Proceeding of IEEE 1994 Int'l Conf Neural Networks (1) [C]. Florida, USA, 1994, 1: 299-304.
- [15] Chen Z, Feng T J, Meng Q C. The Application of Wavelet Neural Network in Time Series Prediction and System Modeling Based on Multiresolution Learning [J]. IEEE

Conference on Systems, Man, and Cybernetics, 1999,(1):  
425-430.

及其在电力需求预测中的应用[J].上海理工大学学报,  
2005, 27(4): 319-322.

[16] 靳中伟,叶舟,陈康民,等.  $\alpha$ -加权模糊线性回归模型

## Comparison of BP and WNN in Power Load Forecasting

FENG Zai-yong<sup>1,3</sup>, ZHONG Kang-hui<sup>2</sup>, MA Yong-wang<sup>1</sup>

(1. College of Information Management, Chengdu University of Technology, Chengdu 610059, China;

2. College of Earth Science, Chengdu University of Technology, Chengdu 610059, China;

3. Suzhou Higher Vocational School of Mechatronic Technology, Suzhou 215031, Jiangsu, China)

**Abstract:** Combining with the practical power load forecasting example, analysing the prediction courses of BP and WNN in convergence speed, stability, generalization and prediction accuracy. It is found that: ① BP model forecasting is better in stability, selects data easier while operates slower, gets inaccurate results; ② WNN model predicts faster, brings out more accurate results, generalizes better while stability is worse, has difficult in selecting data; ③ On the whole, WNN model performs better than BP model while it has some technical problems.

**Key words:** BP model; WNN model; convergence speed; stability; generation; prediction accuracy