

# 基于新型试凑法 BP 模型的短期负荷预测

刘 明

(连云港供电公司, 江苏 连云港 222000)

**摘 要:** 为了以较简单的模型提高短期负荷预测的准确率和效率, 深入分析影响短期负荷预测的各种因素, 确定主要的并可以处理的影响因子。对负荷数据进行纵向垂直处理、横向水平处理、基于全体负荷数据的归一化处理, 对其他影响因子分别做出合适的处理后, 突破经验公式确定隐含层节点数及各层的传递函数、训练函数, 并随之提出了一种新型的方法确定误差反传算法的模型结构, 将所得到的模型应用于南京市短期负荷预测中, 得到了较同类方法更好的结果, 准确性高, 计算时间短, 很大程度上减小了预测误差。

**关键词:** 负荷预测; 新型方法; 误差反传算法; 影响因素; 预处理

## 0 引言

电力系统负荷预测是根据历史的负荷数值、天气状况、温度高低、湿度情况、是否节假日等, 建立预测模型, 预测未来时间内的负荷数值及其变化规律。准确的负荷预测对电力系统的经济调度、合理运行起着重要作用。短期负荷预测是指预测未来一天或几天的负荷, 可以确定机组工作状态、负荷分配调度等。

## 1 短期负荷预测的影响因素

电力系统负荷序列周期性与随机性都很强, 对负荷序列进行预测之前, 需要先整体把握负荷序列的变化规律和发展走向。针对负荷序列的变化规律, 同时考虑其他各种因素如天气、温度等的影响, 用这样的预测模型才能得到较好的负荷预测结果<sup>[1]</sup>。

### 1.1 基本正常负荷分量

又称为典型负荷分量, 它与天气状况等没有关系, 该负荷分量呈现出线性规律变化的特点, 同时具有一定的周期性, 负荷的组成内容和方式、负荷的组成种类以及各种负荷所占的比例因各个地区的不同而不同<sup>[2]</sup>。

### 1.2 天气敏感负荷分量

天气状况极大地影响着该部分分量, 如阴晴雨雪、有无阳光、温度高低、湿度大小等都会改变天气敏感负荷分量。一年之中, 不同的季节, 有不同的天气, 负荷也会跟着变化, 例如, 夏天处于高温期, 空调电扇等降温设备的大量使用会带动用电负

荷达到高峰期, 冬季的取暖负荷也会增加一定的负荷量, 春秋两季的用电负荷相对会少些, 一年四季, 周而复始。

### 1.3 特别事件负荷分量

也称为非正常负荷分量或者不确定负荷分量, 这部分负荷分量会使得负荷特性和正常的负荷特性有较大偏差, 由于这类事件的不确定性往往需要由相关工作人员来主观地加以判断, 因此, 这部分负荷分量在负荷预测模型中需要通过人工处理加以改进<sup>[3]</sup>。

### 1.4 随机负荷分量

随机负荷分量不可预知, 是负荷预测中难以加以说明的一部分, 例如在传统的时间序列预测方法中, 各个时刻的随机负荷分量可以考虑用同时刻的随机时间序列来代替。

## 2 对数据的预处理

对原始负荷数据进行预处理, 保证负荷序列的水平平滑性和垂直平滑性, 将天气状况、温度、日期类型等归一化以便于模型的辨识, 前期工作会在很大程度上决定网络的性能和最终的预测精度。

### 2.1 对负荷数据的纵向垂直预处理

借鉴作者的文献[1]2.2 部分的内容, 将南京市 2010 年 8 月 25 日至 12 月 7 日共 105 天的每天 10:00 点的负荷数据进行垂直预处理, 得到图 1 所示的结果。虚线表示的是对负荷数据序列进行垂直预处理之前的负荷序列曲线, 实线表示的是对异常数据点修正以后的负荷序列曲线, 可以看出, 修正后的负荷序列更加平滑, 趋于合理。

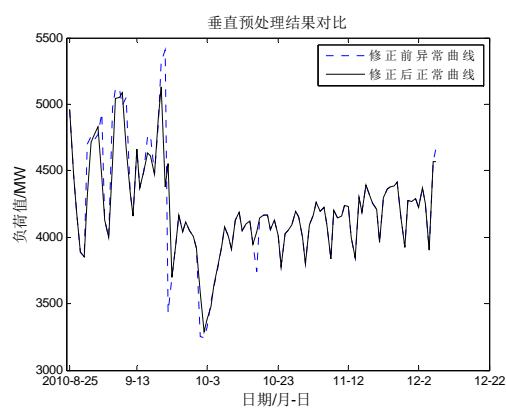


图 1 垂直预处理前后的负荷序列曲线对比

2.2 对负荷数据的横向水平预处理

对原始负荷数据进行了纵向垂直预处理后，还要保证每天的各个时刻负荷数据组成的横向负荷序列具有一定的平滑性，采用中位数法对其做水平预处理。对南京市 2010 年 8 月 25 日至 9 月 4 日的负荷序列水平平滑处理后，结果如图 2 所示。

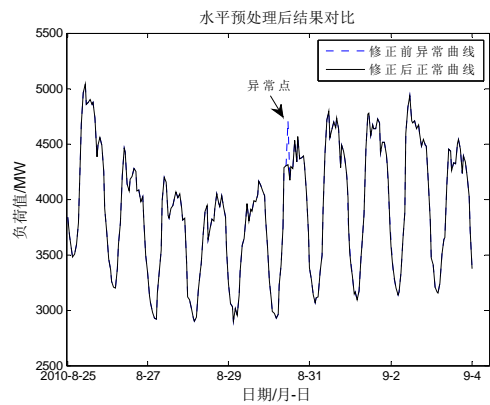


图 2 水平预处理前后的负荷序列曲线对比

如果需要加大修正力度，使得负荷序列更加趋于平滑，可采取减小偏离率阈值、即提高要求的方法，当偏离率为 0.02 时，修正结果如图 3 所示。

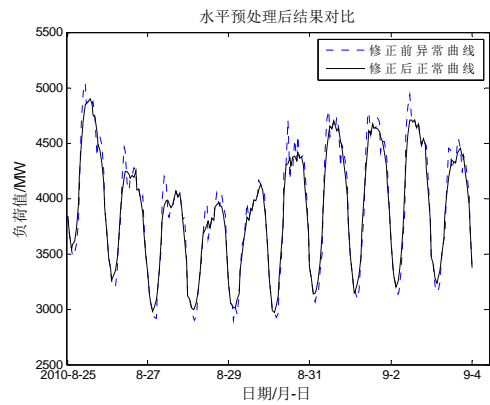


图 3 加大修正力度后水平预处理前后的曲线对比

2.3 对负荷数据的归一化处理

文献[4]提到：以恰当的方式对数据进行归一化处理，使得输入的负荷数据在[-1,1]范围之内，可以加速神经网络的收敛，本文所采用的归一化方法如式（1）。

$$y = \frac{2 \times (x - x_{\min})}{x_{\max} - x_{\min}} + (-1) \tag{1}$$

不同于MATLAB中的平均数方差法（ $x_{\max}$ 、 $x_{\min}$ 分别为待处理负荷样本中每行的最大值和最小值）的是，上式中， $x$ 为原始负荷数据， $y$ 为归一化之后的负荷数据，而 $x_{\max}$ 、 $x_{\min}$ 分别为待处理的整体负荷样本中的最大值和最小值。

3 建立新型 BP 预测模型

对每天的 24 小时负荷分别建立 24 个不同的神经网络预测模型，每个神经网络有一个输出量，对应于一天的某个时刻的整点负荷数值，这样的处理方法具有针对性，使得单个的网络较小，不容易产生过拟合，在对各天同一时刻的负荷数据进行训练后，网络的各项参数趋于合理化，加上网络本身的联想记忆和学习能力，将使网络训练和预测的效率和精确性大大提高。

假设负荷序列用  $l(h,n)$  表示，温度序列用  $T(s,n)$  表示，日期类型序列用  $D(n)$  表示，天气状况序列用  $W(n)$  表示，其中  $h=1,2,3 \cdots 24$ ，代表一天中的 24 个小时； $s=L, H$  分别为一天中的最低温度、最高温度； $n=1,2,3 \cdots N$ ，负荷数据是第  $n$  天的。

通过对负荷数据特性以及其他影响因素的深入分析，选取表 1 所示的参数作为网络的输入量和输出量。

表 1 网络的参数模型

输入节点	相应输入量
1	预测日前一天同一时刻的负荷值 $l(h,n-1)$
2	预测日前两天同一时刻的负荷值 $l(h,n-2)$
3	预测日一周同一时刻的负荷值 $l(h,n-7)$
4	预测日前一时刻负荷值 $l(h-1,n)$
5	预测日前两时刻负荷值 $l(h-2,n)$
6	预测日一天前一时刻的负荷值 $l(h-1,n-1)$
7	预测日一天前两时刻的负荷值 $l(h-2,n-1)$
8	预测日一天后一时刻的负荷值 $l(h+1,n-1)$
9	预测日的最高温度 $T(H,n)$
10	预测日的最低温度 $T(L,n)$
11	预测日一天前的最高温度 $T(H,n-1)$
12	预测日一天前的最低温度 $T(L,n-1)$
13	预测日的天气状况 $W(n)$
14	预测日的日期类型 $D(n)$
输出节点	相应输出量
1	预测日的负荷值 $l(h,n)$

预测模型的输入量为相关的影响因素共 14 个，输出量为预测时刻的负荷数据共 1 个，即输入层有 14 个节点，输出层有 1 个节点，下面给出隐含层节点数的确定过程。

BP 神经网络的预测精度和学习效率等都受到隐含层节点数的较大影响，最佳隐含层节点数的选择可参考公式 (2)。

$$\begin{cases} N < M - 1 \\ N < \sqrt{(L+M)} + a \\ N = \log_2 M \end{cases} \quad (2)$$

式中，M 为输入层节点数，N 为隐含层节点数，L 为输出层节点数，a 为 0-10 之间的常数，根据上面内容知道 M=14，L=1，可以确定隐含层节点数大概为 4。为了确定最佳的隐含层节点数，作者进行了大量的试验，以预测的南京市 2010 年 12 月 1 日至 12 月 7 日共七天的负荷预测平均相对误差的平均值（取 10 次结果的平均值）作为衡量隐含层节点数优劣的标准，具有较深的纵向适用性；同时隐含层节点数的范围从 1 至 50，以尽可能大的范围覆盖最优隐含层节点数。最终得到试验结果如表 2 所示。

表 2 隐含层节点数与误差关系

隐含层节点数	相对误差	隐含层节点数	相对误差
1	0.022695048	13	0.030549771
2	0.033406983	14	0.030322962
3	0.027387774	15	0.030713469
4	0.028185145	16	0.032009651
5	0.025379264	17	0.03169469
6	0.030078753	18	0.030214853
7	0.029000686	19	0.033754961
8	0.029589668	20	0.031187821
9	0.029069384	25	0.029779343
10	0.028022946	30	0.033043573
11	0.031073854	40	0.03471407
12	0.03120911	50	0.035229814

相对误差与隐含层节点数的关系如图 4 所示。

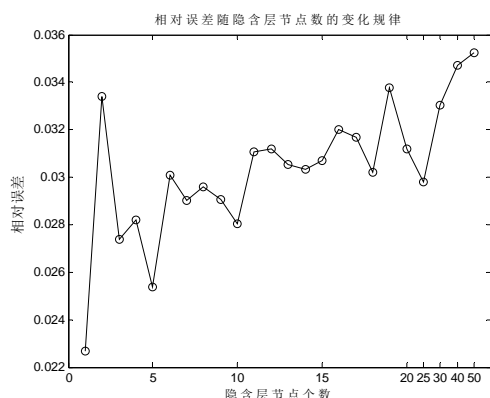


图 4 预测误差和隐含层节点数的关系

由试验结果可以得到：预测误差随着隐含层节点数的增加高低起伏变化，节点数小于 30 时，误差变化较平稳；节点数增加，误差有变大的趋势，当隐含层节点数取 1 时，预测误差最小。如果按照经验公式计算，隐含层节点数为 4，但经验公式毕竟只是一种参考，它没有确定的理论依据，实践是检验真理的唯一标准，因此，BP 神经网络的隐含层节点数取 1。

因此最终所采用的 BP 神经网络模型结构为 14-1-1，当隐含层节点的传递函数采用正切 S 型传递函数  $\tan sig$ ，输出层节点传递函数采用线性传递函数  $purelin$ ，训练函数采用动态自适应学习率的梯度下降 BP 算法训练函数  $traingda$ ，建模和预测效果较好。

#### 4 预测仿真及结果分析

BP 神经网络训练数据为南京市 2010 年 9 月至 11 月共 3 个月的每天整点负荷数据，预测 12 月 1 日至 7 日的负荷数据，由于训练需要一周前的负荷数据，因此训练和预测负荷数据样本集为 8 月 25 日至 12 月 7 日的所有整点负荷。

采用上面的新型 BP 模型进行负荷预测，得到最终的回归结果如图 5 所示。从宏观上可以得出：负荷预测过程中的训练、验证、测试的回归率均超过 90%，总体回归率为 94.678%，预测效果比较理想。

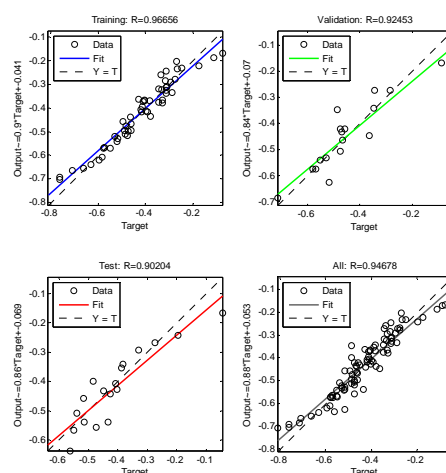


图 5 负荷预测回归曲线

最终预测的 12 月 1 日至 12 月 7 日七天的负荷和实际负荷对比情况如图 6 所示，预测相对误差曲

线如图 7 所示。

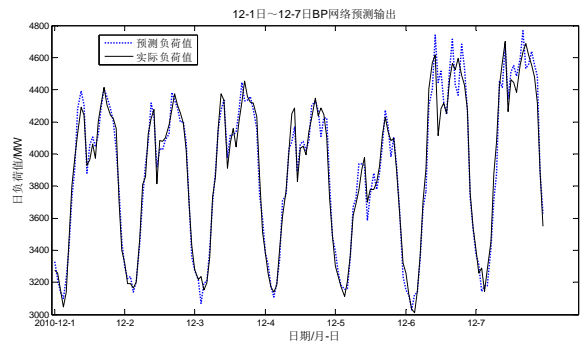


图 6 预测负荷和实际负荷对比情况

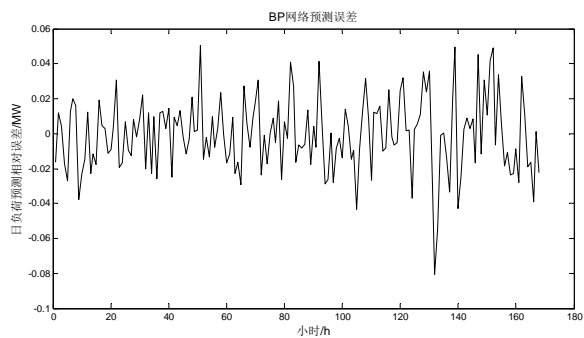


图 7 预测误差曲线图

表 3 给出七天的负荷预测的平均相对误差。

表 3 预测误差结果分析			%
日期	平均相对误差	平均均方根误差	
2010-12-1	1.600806267	1.805440165	
2010-12-2	1.210243329	1.407238076	
2010-12-3	1.479510529	1.882261005	
2010-12-4	1.403444733	1.842558603	
2010-12-5	1.445937408	1.76227647	
2010-12-6	2.282211663	3.062069492	
2010-12-7	2.137641117	2.512670542	
七天平均值	1.6514	2.0392	

七天的负荷预测相对误差的平均值为 1.6514%，最大相对误差为 2010-12-6 的

2.282211663%，最小相对误差为 12-2 的 1.210243329%，对应的均方根误差的平均值、最大均方根误差和最小均方根误差分别为 2.0392%、3.062069492%、1.407238076%，对比于参考文献[5]等采用 BP 模型进行负荷预测的结果有较高的提升。

5 结论

BP 神经网络的结构并没有根据经验公式确定，而是在进行了可能的所有试验后，最终确定最佳隐含层节点数、传递函数和训练函数，因此得到了类比同类文章更高的预测准确率。

参考文献：

[1] 刘明.基于小波和神经网络理论的电力系统负荷预测研究[D].南京:南京理工大学,2012.

[2] 张师玲.基于 RBF 神经网络与模糊控制的电力系统短期负荷预测[D].镇江:江苏大学,2010.

[3] 石万清.电力系统短期负荷预测技术的研究与实现[D].合肥:合肥工业大学,2003.

[4] J.Sola , J.Seville . Importance of Input Data Normalization for the Application of Neural Network to Complex Industrial Problems[J]. IEEE Trans on Nuclear Science, 1997, 44(3): 1152~1156.

[5] 吕婵. 基于 BP 神经网络的短期负荷预测[D]. 武汉: 华中科技大学, 2007.

作者简介：

刘 明（1986-），男，江苏徐州人，硕士研究生，从事调度运行和继电保护方面的研究， E-mail：Liudming@126.com。