

# 基于迟滞 ELM 模型的短期风速预测

余 敖, 陈 亮, 彭敬涛

( 东华大学 信息科学与技术学院, 上海 201620 )

**摘 要:** 为了应对能源危机, 许多国家开始大力发展最有发展前景之一的风能, 而风速预测是进行风电场出力预测的前提条件。目前常用的风速预测方法没有得到很高的预测精度以及预测时间。为了改善风速时间序列的预测精度和预测时间, 提出了一种基于迟滞极速学习机 (Extreme Learning Machine) 模型的风速预测方法。ELM 算法是一种新型神经网络, 计算效率高, 性能优越, 能避免局部最小化。通过改变神经元激励函数的方式将迟滞特性引入神经网络中, 以增强历史输入对当前响应的影响, 从而提高有用信息的利用率, 提高风速时间序列的预测性能。仿真结果表明, 与 ELM 模型等方法相比, 迟滞 ELM 模型能够有效减小风速时间序列的预测误差, 提高了预测精度以及减少了预测时间。

**关键词:** 极速学习机; 迟滞; 风速预测; 机器学习

**中图分类号:** TP391

**文献标识码:** A

**文章编号:** 1673-629X(2017)06-0130-06

**doi:** 10.3969/j.issn.1673-629X.2017.06.027

## Short-term Wind Speed Forecasting by Using Hysteretic ELM Model

YU Ao, CHEN Liang, PENG Jing-tao

( College of Information Science and Technology, Donghua University, Shanghai 201620, China )

**Abstract:** The wind, as one of the most promising energy, has been developed by many countries in response to the energy crisis, and the wind speed prediction is the premise of wind farm output prediction. Many wind speed prediction methods are adopted to get high prediction accuracy and prediction of time presently. In order to improve the prediction accuracy and time of wind speed time series, a wind speed prediction method based on hysteresis ELM algorithm has been proposed. ELM algorithm is a new type of neural network, and it has high computational efficiency and superior performance, which can avoid local minimum. Hysteresis characteristics have been introduced in neural network via changing neuronal excitation function to enhance influence of the response history to the current input and thus to improve the utilization rate of useful information and to promote the wind speed time series prediction performance. The simulation results show that compared with the ELM model, the hysteresis ELM model can reduce prediction error of wind speed time series effectively for improvement of prediction accuracy and decreasing of prediction time.

**Key words:** ELM; hysteresis; wind speed prediction; machine learning

### 1 概 述

随着世界经济的快速发展和人口的快速增长, 世界各国对能源的需求量也日益增大, 而传统的不可再生能源将逐渐枯竭, 这可能导致全球能源危机。为了应对能源危机, 世界各国对风能源产生了浓厚的兴趣。风能是一种清洁的可再生能源, 风能利用的主要形式是风力发电, 也是目前可再生能源中技术最成熟、最具规模化开发条件和商业化发展前景的发电方式之一<sup>[1]</sup>。

风速的随机性很强, 而这种随机性给电力系统带来很大的冲击, 对电能质量、系统稳定性、电力调度等

有较大的影响。对风力发电的出力进行预测可以在很大程度上降低风力发电的缺点对电网带来的不利影响, 而风速预测是进行风电场出力预测的前提条件<sup>[2]</sup>。

风速时间序列的预测研究在风电并网、气象监测等许多领域具有广泛的应用前景, 是当前研究的热点问题之一<sup>[3]</sup>。目前, 常用的风速时间序列预测分析方法有神经网络法、持续法、时间序列法、模糊逻辑法、卡尔曼滤波法、混沌时间序列预测分析法等<sup>[4-8]</sup>。这些预测方法运用不同的理论对风速序列展开预测研究, 从不同的角度对风速序列的预测性能起到改善作用, 同时也证实了风速序列预测的可行性。在这些方法

收稿日期: 2016-06-21

修回日期: 2016-09-22

网络出版时间: 2017-03-13

基金项目: 上海市自然科学基金资助项目 (14ZR1400700)

作者简介: 余 敖 (1990-), 男, 硕士研究生, 研究方向为机器学习、图像处理; 陈 亮, 副教授, 研究方向为机器学习、图像处理。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20170313.1547.104.html>

中,由于本身的自适应学习能力和灵活性,神经网络已经越来越多地应用于风速预测<sup>[9]</sup>。神经网络方法在黑箱系统建模方面拥有很大的优势,因此在时间序列预测分析中具有良好的预测性能。

极速学习机(Extreme Learning Machine, ELM)是一种新型的单隐层前馈神经网络,ELM 算法提供了一个结合支持向量机(SVM)、径向基函数(RBF)以及标准的前馈神经网络优势的统一学习平台。在计算方面,ELM 算法克服了昂贵的学习代价,比如反向传播<sup>[10]</sup>。因为这些优势,ELM 算法或者基于 ELM 的改进算法广泛应用于时间序列预测方面。Ronay A K 等<sup>[11]</sup>利用 ELM 算法和最近邻算法进行短期风速预测。王智慧<sup>[12]</sup>利用移动加权 ELM 算法进行股票预测以及带有遗忘机制的在线序列 ELM 算法进行股票预测。李彬等<sup>[13]</sup>利用 ELM 算法进行混沌时间序列预测等。和传统的学习算法相比,ELM 算法不需要同时训练隐含层的参数以及输出权值,只是随机给定隐含层参数,因此节省了大量的计算时间。由于 ELM 的快速学习和不严格遵守优化原则,使得 ELM 的随机分配更像生物神经网络的学习机制。同时,ELM 算法的训练过程是一个能够直接得到输出矩阵的广义逆的线性的最小二乘法计算过程。

正是因为以上不同点,使得 ELM 算法的计算复杂度远低于传统算法。

但是,随机给定隐含层的参数,或多或少降低 ELM 算法的学习性能。除此之外,对隐含层节点的数量,并没有准确估计的方法,这是值得研究的一个点。在实际应用中,可以由交叉验证,但由于技术花费代价大,导致不能在线使用。

ELM 神经网络的记忆能力和泛化性能主要取决于网络的连接权值,ELM 神经网络根据当前输入得到未来的输出,而对于训练数据中所蕴含的历史变化趋势等信息利用不够充分,而在生物神经网络中普遍存在迟滞特性,这些迟滞特性可以提高神经元对历史变化信息的利用率<sup>[14]</sup>,从而改善 ELM 神经网络的性能。

使用迟滞激励函数的神经元的输出不仅取决于它的输入,还取决于输入的导数信息。最重要的是它为神经元提供记忆的信息。除此之外,当输入卡住在传统的 Sigmoid 激励函数的饱和区域,迟滞函数可以在两个迟滞 Sigmoid 激励函数之间进行自由切换,从而避免神经元陷入局部最小值以及使得整个神经网络获得更快的学习速度。

因此,提出了一种基于迟滞 ELM 模型的算法来预测短期风速序列。利用迟滞 ELM 神经网络进行短期风速预测,不仅可以使当前数据信息,而且还可以使用历史数据信息,因此大大地提高了风速预测的性能。

## 2 ELM 算法

ELM 算法是新加坡南洋理工大学的 Huang 教授等在 2004 年提出的一种神经网络模型<sup>[10,15]</sup>。前馈神经网络通常用于解决回归问题和分类问题<sup>[16]</sup>。接下来,从回归的角度来介绍其数学模型。

选取  $N$  个样本数据  $(x_j, y_j)_{j=1}^N$ , 其中  $x_j = [x_{j1}, x_{j2}, \dots, x_{jm}]^T$ ,  $y_j = [y_{j1}, y_{j2}, \dots, y_{jm}]^T$ , 一个具有  $N$  个隐含层节点的单隐层前馈神经网络模型(SLFN)的具体描述如下:

$$y_j = \sum_{i=1}^N \beta_i g(\omega_i \cdot x_j + \alpha_i), j = 1, 2, \dots, N \quad (1)$$

其中,  $\omega_i = [\omega_{i1}, \omega_{i2}, \dots, \omega_{im}]^T$  为第  $i$  个隐含层节点的输入权值;  $\alpha_i \in R$  为第  $i$  个隐含层节点的阈值;  $\beta_i \in R^m$  为第  $i$  个隐含层节点的输出权值;  $\omega_i \cdot x_j$  为  $\omega_i$  和  $x_j$  的内积; 激励函数  $g(x)$  通常选取 Sigmoid 函数。

式(1)的简洁表达式如下:

$$H\beta = Y \quad (2)$$

其中

$$H = \begin{bmatrix} g(\omega_1 \cdot x_1 + \alpha_1) & \cdots & g(\omega_N \cdot x_1 + \alpha_N) \\ \vdots & \cdots & \vdots \\ g(\omega_1 \cdot x_N + \alpha_1) & \cdots & g(\omega_N \cdot x_N + \alpha_N) \end{bmatrix}_{N \times N}$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_N^T \end{bmatrix}_{N \times m} \quad Y = \begin{bmatrix} y_1^T \\ \vdots \\ y_N^T \end{bmatrix}_{N \times m}$$

其中,  $H$  为隐含层输出矩阵。

选取均方值误差  $E(W)$  作为评价标准。在 SLFN 模型中,问题是找到最优参数  $\omega, \alpha, \beta$  使得目标函数  $E(W)$  最小。具体的数学表达式如下所示:

$$\arg \min_{W=(\omega, \alpha, \beta)} E(W) = \arg \min_{W=(\omega, \alpha, \beta)} \frac{1}{N} \sum_{j=0}^N \frac{1}{2} \|\varepsilon_j\|_2^2 \quad (3)$$

$$\sum_{i=1}^N \beta_i g(\omega_i \cdot x_j + \alpha_i) - y_j = \varepsilon_j, j = 1, 2, \dots, N \quad (4)$$

其中,  $\varepsilon_j = [\varepsilon_{j1}, \varepsilon_{j2}, \dots, \varepsilon_{jm}]$  为第  $j$  个误差。

Huang<sup>[17]</sup>已经证明:如果  $(\omega, \alpha)$  是随机选定,那么隐含层输出矩阵  $H$  也会确定。然后,式(3)的问题是一个标准的最小二乘法问题,其解决方案如下:

$$\beta^* = H^+ Y \quad (5)$$

其中,  $H^+$  是  $H$  的 Moore-Penrose 广义逆。

Huang<sup>[17]</sup>把这种算法称为 ELM。

ELM 算法表现出一些不同于传统学习模式(比如误差反向传播算法)的地方:

(1)传统的训练过程是一个复杂的非线性优化问题,然而 ELM 算法的训练过程是一个能够直接得到输出矩阵的广义逆的线性的最小二乘法计算过程。

(2)不像传统的学习算法需要同时训练隐含层的

参数  $(\omega, \alpha)$  以及输出权值  $\beta$ , ELM 算法只是随机给定隐含层参数, 因此节省了大量的计算时间。

这种随机分配也更像生物神经网络的学习机制。正是这些不同点使得 ELM 算法的计算复杂度远低于传统算法。

由于 ELM 的快速学习和不严格遵守优化原则, 使得 ELM 的随机分配更像生物神经网络的学习机制。但是, 在这样重要复杂的推导中, 特别是随机给定隐含层的参数, 或多或少降低了 ELM 算法的学习性能。除此之外, 对隐含层节点的数量, 也没有准确估计的方法, 只是手动尝试。在实际应用程序中, 可以由交叉验证, 但技术花费代价大, 不能在线使用。

### 3 迟滞 ELM 模型

#### 3.1 迟滞函数

迟滞在自然界中是随处可见的一种非线性现象。迟滞的定义是一个滞后效应由于力量作用于自身的变化。许多工程系统具有迟滞现象, 例如机械结构和铁磁材料。滞后型非线性的识别也已经研究了很多年<sup>[18]</sup>。迟滞特性分为离散迟滞和连续迟滞。对于离散迟滞神经元模型来说:

(1) 如果一个神经元的输入低于某个阈值(低于跳变点), 神经元的输出是零。

(2) 如果一个神经元的输入超过一个阈值(高于跳变点), 就会使得神经元的输出为一个定值。

(3) 如果一个神经元的输入在这两个阈值之间, 神经元的输出等于它的前一个值。

图 1(a) 展示了离散迟滞。因为输出只有两种状态, 也叫二进制离散迟滞。离散迟滞应用在 Hopfield 神经网络中比较多, 可以求解组合优化的问题(如  $N$  皇后问题<sup>[19]</sup>、旅行商问题<sup>[20]</sup>以及交叉开关分组交换问题<sup>[21]</sup>等等), 同时离散迟滞也会引起系统的同步<sup>[22]</sup>。目前已有研究, 并且已应用于图像着色的问题<sup>[23]</sup>。

对于连续迟滞神经元模型来说:

(1) 如果一个神经元的输入低于某个阈值(低于跳变点), 神经元的输出是零。

(2) 如果一个神经元的输入超过一个阈值(高于跳变点), 就会使得神经元的输出为一个定值。

(3) 如果一个神经元的输入在这两个阈值之间, 神经元的输出通过激励函数的计算公式得到。

图 1(b) 展示了连续迟滞。连续迟滞和离散迟滞相比, 就是第三种情况不太一样。连续迟滞用在 Hopfield 神经网络上也很广泛。Bharitkar 等<sup>[24]</sup>在 2000 年提出了连续迟滞 Hopfield 神经网络。该神经网络的激励函数由两个加了迟滞参数的 Sigmoid 函数组成

(见图 1(b))。与普通的 Sigmoid 激励函数相比, 神经元的输出不仅依赖于输入, 而且也依赖于其输入的导数。最重要的是它为神经元提供记忆的信息。除此之外, 在神经网络的训练中, 迟滞特性还可以抑制故障饱和现象。连续迟滞 Hopfield 神经网络也已经应用于解决  $N$  皇后问题<sup>[24-25]</sup>、时分多址广播调度问题<sup>[26]</sup>、风速预测问题<sup>[27]</sup>、控制与优化问题<sup>[28]</sup>等等。

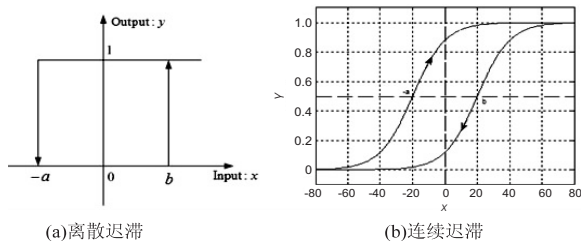


图 1 迟滞函数

从上述内容可以看到, 迟滞主要是从 Hopfield 神经网络以及典型的动力系统中引入的。

由于短期风速预测适合连续的迟滞函数, 所以只介绍连续的迟滞函数。迟滞神经元的激励函数数学描述如下所示:

$$g(x) = \begin{cases} \frac{1}{1 + e^{-c_a(x+a)}}, & \dot{x}(t - \delta t) \geq 0 \\ \frac{1}{1 + e^{-c_b(x+b)}}, & \dot{x}(t - \delta t) < 0 \end{cases} \quad (6)$$

$$\dot{x}(t - \delta t) = \frac{dx(t - \delta t)}{dt} = \lim_{\delta t \rightarrow 0} \frac{x(t) - x(t - \delta t)}{\delta t} \quad (7)$$

其中,  $x(t)$  为神经元的输入。

激励函数由两个加了迟滞参数的 Sigmoid 函数组成,  $a$  和  $b$  是最主要的参数,  $c_a$  和  $c_b$  都是正斜率参数。当  $a = b = 0$  和  $c_a = c_b = 1$  时, 该迟滞 Sigmoid 激励函数就变成了普通的 Sigmoid 激励函数。由图 1(b) 可以看到, 这种神经元的输出不仅取决于它的输入, 还取决于输入的导数信息。最重要的是它为神经元提供记忆的信息。除此之外, 当输入卡住在传统的 Sigmoid 激励函数的饱和区域时, 迟滞函数可以从两个迟滞 Sigmoid 激励函数之间切换, 从而避免神经元陷入局部最小值以及使得整个神经网络获得更快的学习速度。此外, 要注意到迟滞神经元的激励函数有四个迟滞参数, 即  $a, b, c_a, c_b$ 。所以与普通神经元相比, 迟滞神经元提供了更大的灵活性。

#### 3.2 迟滞 ELM 神经网络模型

提出了一种改进的迟滞 ELM 神经网络模型, 通过采用迟滞激励函数来提高泛化能力和联想记忆功能, 在某种程度上考虑到准确率和效率, 随机给定隐含层节点的参数。

将普通的 Sigmoid 激励函数换成迟滞 Sigmoid 激

励函数,保持其他的地方不变。选取  $N$  个样本数据  $(x_j, y_j)_{j=1}^N$ , 其中  $x_j = [x_{j1}, x_{j2}, \dots, x_{jN}]^T$  和  $y_j = [y_{j1}, y_{j2}, \dots, y_{jN}]^T$ 。而迟滞激励函数是分段函数,所以为了实际 MATLAB 程序编写方便,先要将 ELM 模型中  $(x_j, y_j)_{j=1}^N$  的  $x_j$  和  $y_j$  进行相关处理,得到

$$x_{aj} = [x_{aj1}, x_{aj2}, \dots, x_{ajn}]^T, x_{bj} = [x_{bj1}, x_{bj2}, \dots, x_{bjN}]^T$$

$$y_{aj} = [y_{aj1}, y_{aj2}, \dots, y_{ajn}]^T, y_{bj} = [y_{bj1}, y_{bj2}, \dots, y_{bjN}]^T$$

具体处理方法参考 3.4 节。

首先,随机给定迟滞参数  $a, b, c_a, c_b$ 。对于  $x_{aj} = [x_{aj1}, x_{aj2}, \dots, x_{ajn}]^T, y_{aj} = [y_{aj1}, y_{aj2}, \dots, y_{ajn}]^T$  来说,一个具有  $N$  个隐含层节点的单隐层前馈神经网络模型(SLFN)的具体描述如下:

$$y_{aj} = \sum_{i=1}^N \beta_{ai} g(\omega_i \cdot x_{aj} + \alpha_i), j = 1, 2, \dots, N \quad (8)$$

其中,  $\omega_i = [\omega_{i1}, \omega_{i2}, \dots, \omega_{iN}]^T$  为第  $i$  个隐含层节点的输入连接权值;  $\alpha_i \in R$  为第  $i$  个隐含层节点的阈值;  $\beta_i \in R^m$  为第  $i$  个隐含层节点的输出权值;  $\omega_i \cdot x_{aj}$  为  $\omega_i$  和  $x_{aj}$  的内积;激励函数  $g(x)$  选取带有迟滞特性的 Sigmoid 函数。

式(4)简化为:

$$H_\alpha \beta_\alpha = Y_\alpha \quad (9)$$

其中

$$H =$$

$$\begin{bmatrix} g(\omega_1 \cdot x_{a1} + \alpha_1) & \dots & g(\omega_{\tilde{N}} \cdot x_{a1} + \alpha_{\tilde{N}}) \\ \vdots & \dots & \vdots \\ g(\omega_1 \cdot x_{aN} + \alpha_1) & \dots & g(\omega_{\tilde{N}} \cdot x_{aN} + \alpha_{\tilde{N}}) \end{bmatrix}_{N \times \tilde{N}} =$$

$$\begin{bmatrix} \frac{1}{1 + e^{-c_a \langle \omega_1 x_{a1} + \alpha_1 \rangle + a}} & \dots & \frac{1}{1 + e^{-c_a \langle \omega_{\tilde{N}} x_{a1} + \alpha_{\tilde{N}} \rangle + a}} \\ \vdots & \dots & \vdots \\ \frac{1}{1 + e^{-c_a \langle \omega_1 x_{aN} + \alpha_1 \rangle + a}} & \dots & \frac{1}{1 + e^{-c_a \langle \omega_{\tilde{N}} x_{aN} + \alpha_{\tilde{N}} \rangle + a}} \end{bmatrix}_{N \times \tilde{N}}$$

$$\beta_\alpha = \begin{bmatrix} \beta_{\alpha 1}^T \\ \vdots \\ \beta_{\alpha \tilde{N}}^T \end{bmatrix}_{\tilde{N} \times m} \quad Y_\alpha = \begin{bmatrix} y_{a1}^T \\ \vdots \\ y_{aN}^T \end{bmatrix}_{N \times m}$$

其中,  $H_\alpha$  为隐含层输出矩阵。

选取均方值误差  $E_\alpha(W)$  作为评价标准。在 SLFN 模型中,问题是找到最优参数  $\omega, \alpha, \beta_\alpha$ , 使得目标函数  $E_\alpha(W)$  最小。具体的数学表达式为:

$$\arg \min_{W = \langle \omega, \alpha, \beta_\alpha \rangle} E_\alpha(W) =$$

$$\arg \min_{W = \langle \omega, \alpha, \beta_\alpha \rangle} \frac{1}{N} \sum_{j=0}^N \frac{1}{2} \|\varepsilon_j\|_2^2 \quad (10)$$

$$\sum_{i=1}^N \beta_{ai} g(\omega_i \cdot x_{aj} + \alpha_i) - y_j = \varepsilon_j, j = 1, 2, \dots, N \quad (11)$$

其中,  $\varepsilon_j = [\varepsilon_{j1}, \varepsilon_{j2}, \dots, \varepsilon_{jm}]$  为第  $j$  个误差。

然后,式(9)的问题是一个标准的最小二乘法问

题,其解决方案如下所示:

$$\beta^* = H^+ Y \quad (12)$$

其中,  $H^+$  是  $H$  的 Moore-Penrose 广义逆。

根据该公式可以得到  $\beta_\alpha$ , 同理得到  $\beta_b$ 。

最后,将  $\beta_\alpha$  和  $\beta_b$  相加得到  $\beta$ 。

$$\beta = \beta_\alpha + \beta_b \quad (13)$$

### 3.3 迟滞参数调整

迟滞神经元的激活函数有四个相关联的迟滞参数,即  $a, b, c_a, c_b$ , 这些参数为神经网络提供了一个高度的灵活性。使用梯度下降法来优化迟滞参数。对于第  $i$  个样本来说,当第  $i$  个隐含层神经元处在上升分支的激励响应时,迟滞参数的人工智能训练如下:

$$\alpha_i(t+1) = \alpha_i(t) - \eta \frac{\partial E(W)}{\partial \alpha_i}, i = 1, 2, \dots, N, j = 1, 2, \dots, N \quad (14)$$

$$\frac{\partial E(W)}{\partial \alpha_i} = \varepsilon(j) \frac{-c_a e^{-c_a \langle \text{net}_i + \alpha_i \rangle}}{(1 + e^{-c_a \langle \text{net}_i + \alpha_i \rangle})^2}, \text{ for } \text{net}_i(t) - \text{net}_i(t-1) \geq 0 \quad (15)$$

其中,  $t$  为第  $i$  次迭代;  $\eta$  为学习速率;  $E(W)$  和  $\varepsilon(j)$  同式(10)和式(11)。

对于第  $i$  个隐含层节点神经元有:

$$\text{net}_i = \omega_i \cdot x_j + \alpha_i \quad (16)$$

同样地,当第  $i$  个隐含层神经元处在下降分支的激励响应时,迟滞参数的人工智能训练如下:

$$b_i(t+1) = b_i(t) - \eta \frac{\partial E(W)}{\partial b_i}, i = 1, 2, \dots, \tilde{N}, j = 1, 2, \dots, N \quad (17)$$

$$\frac{\partial E(W)}{\partial b_i} = \varepsilon(j) \frac{-c_b e^{-c_b \langle \text{net}_i + b_i \rangle}}{(1 + e^{-c_b \langle \text{net}_i + b_i \rangle})^2}, \text{ for } \text{net}_i(t) - \text{net}_i(t-1) < 0 \quad (18)$$

对于两个正斜率参数  $c_a$  和  $c_b$  来说,它们也同样可以用梯度下降法进行调整。引进一个中间变量  $d$  以便于防止它们训练成为负数。令  $d = \sqrt{c}$ , 斜率参数可以用梯度下降法写成如下形式:

$$d_i(t+1) = d_i(t) - \eta \frac{\partial E(W)}{\partial d_i}, i = 1, 2, \dots, N, j = 1, 2, \dots, N \quad (19)$$

$$\begin{cases} \frac{\partial E(W)}{\partial d_i} = \varepsilon(j) \frac{-2d_\alpha (\text{net}_i + \alpha_i) e^{-d_\alpha^2 \langle \text{net}_i + \alpha_i \rangle}}{(1 + e^{-c_a \langle \text{net}_i + \alpha_i \rangle})^2}, \\ \quad \text{for } \text{net}_i(t) - \text{net}_i(t-1) \geq 0 \\ \frac{\partial E(W)}{\partial d_i} = \varepsilon(j) \frac{-2d_b (\text{net}_i + b_i) e^{-d_b^2 \langle \text{net}_i + b_i \rangle}}{(1 + e^{-d_b^2 \langle \text{net}_i + b_i \rangle})^2}, \\ \quad \text{for } \text{net}_i(t) - \text{net}_i(t-1) < 0 \end{cases} \quad (20)$$

迟滞网络参数的优化训练提高了神经网络适应历史输入信息的能力,使网络可以充分利用训练样本中

包含的有用信息,并能有效地改善泛化能力和联想记忆能力。此外,有别于传统的 Sigmoid 激励函数,在输入饱和和区域卡住,迟滞函数可以从两个迟滞 Sigmoid 激励函数之间切换,从而避免神经元陷入局部最小值以及使得整个神经网络获得更快的学习速度。

### 3.4 迟滞 ELM 算法的实现

与传统 ELM 神经网络相比,迟滞 ELM 网络的每个神经元的激励函数都具有迟滞参数。为了使网络具有良好的泛化性能,迟滞神经网络在训练过程中除了要进行权值的训练,还要对迟滞参数进行训练,这增加了神经网络训练的复杂性。

为此,提出一种网络参数的混合训练方法,即先采用最小二乘法训练 ELM 神经网络的连接权值,后采用梯度下降法获取最优的迟滞参数。整体思路:先确定输出权值,再确定迟滞参数,再训练好的参数进行短期风速预测。

迟滞激励函数是分段函数,所以为了便于实际 MATLAB 程序编写,先要将 ELM 模型中  $(x_j, y_j)_{j=1}^N$  的  $x_j$  和  $y_j$  进行处理,具体方法如下:

令  $x_{a1} = x_{j1}, x_{b1} = 0, y_{a1} = y_{j1}, y_{b1} = 0$ , 然后再将  $x_{j2}$  与  $x_{j1}$  进行对比,如果  $x_{j2}$  大于或者等于  $x_{j1}$ , 那么令  $x_{a2} = x_{j2}, x_{b2} = 0, y_{a2} = y_{j2}, y_{b2} = 0$ 。否则令  $x_{a2} = 0, x_{b2} = x_{j2}, y_{a2} = 0, y_{b2} = y_{j2}$ , 一直循环到最后一个  $x_{jn}$ , 如果  $x_{jn}$  大于或者等于  $x_{j(n-1)}$ , 那么令  $x_{ajn} = x_{jn}, x_{bjn} = 0, y_{ajn} = y_{jn}, y_{bjn} = 0$ , 否则令  $x_{ajn} = 0, x_{bjn} = x_{jn}, y_{ajn} = 0, y_{bjn} = y_{jn}$ 。

整个迟滞 ELM 算法的具体实施步骤如图 2 所示。

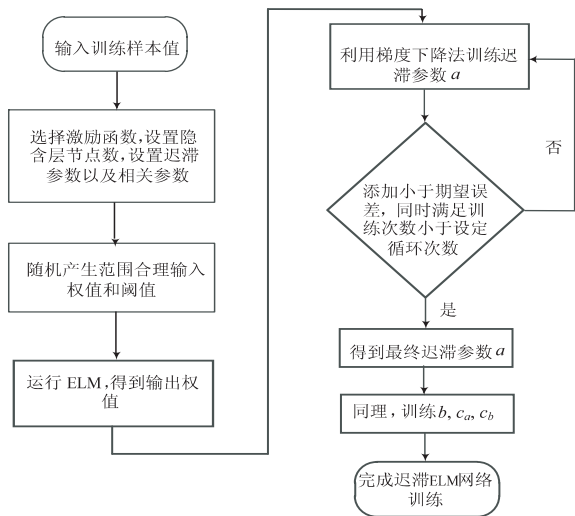


图 2 实施步骤

## 4 短期风速预测应用

风速时间序列具有很强的波动性和随机性,传统的神经网络由于神经元响应机制单一,很难适应风速的剧烈波动,因此预测结果并不理想。采用上述迟滞 ELM 神经网络开展风速时间序列预测分析,与 ELM 模

型的预测结果进行对比。

采取加拿大的 ACADIA VALLEY 地区的每小时的平均风速数据,时间跨度是从 2015 年 1 到 7 月的风速数据,利用 1 至 6 月的数据来预测 7 月风速数据,总共 650 组样本数据。所有实验都是在同一台标准 PC 上,该 PC 使用的是 2.60-GHz Intel Pentium(R) Dual-Core CPU 和 2 G 的运行内存, MATLAB 版本是 R2013a。迟滞 ELM 网络采用 3 层结构,输入层神经元数量  $m = 6$ , 隐含层神经元数量分别为 5, 10, 20, 50, 100 (在经过预测结果对比之后采用 10 个神经元), 输出层神经元数量为 1。在实验过程中,分别从隐含层节点数和训练样本数进行实验,评价标准是均方误差、最大误差、平均误差、训练时间、测试时间,其中均方误差最重要。

实验结果如下:

(1) 用第 1 组至第 600 组数据作为训练样本, 第 601 组至第 650 组作为测试样本, 求取 50 次的平均值作为计算结果, 根据手动实验对比, 发现隐含层节点数为 10 时, 实验结果较好。隐含层节点数为 10 时的实验结果如图 3 和图 4 所示。

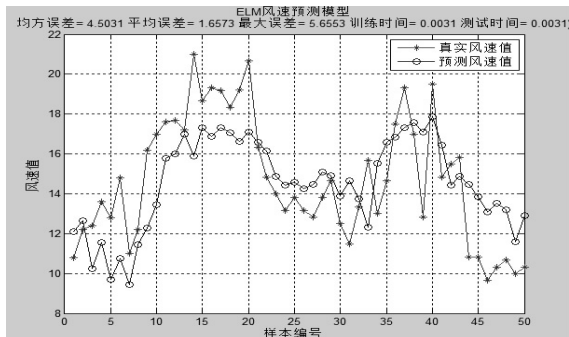


图 3 ELM

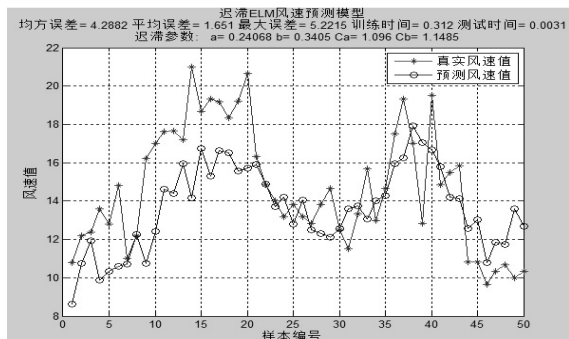


图 4 迟滞 ELM

通过对表 1 中的数据进行对比分析,可以得到:对于迟滞 ELM 算法来说,均方误差、平均误差和最大误差均要比 ELM 算法好。它的训练时间和测试时间要比 ELM 算法稍微长点,是因为在训练阶段引入梯度下降法来优化迟滞函数的参数和在测试阶段随机给定隐含层神经元连接参数。无论如何,迟滞 ELM 所耗费的时间是可以承受的,均方误差和泛化能力却得到显著提高。

表1 不同算法的性能比较

算法	均方 误差	平均 误差	最大 误差	训练 时间	测试 时间
ELM	4.503 1	1.657 3	5.655 3	0.003 1	0.003 1
迟滞 ELM	4.288 2	1.651	5.221 5	0.312	0.003 1

(2)用第1组至第300组数据作为训练样本,第601组至第650组作为测试样本,求取50次的平均值作为计算结果作为比较的第一组;用第300组至第600组数据作为训练样本,第601组至第650组作为测试样本,求取5次的平均值作为计算结果作为比较的第二组;两组对比实验结果如表2所示。

表2 不同算法以及不同样本数据的比较

序号	算法	均方 误差	平均 误差	最大 误差	训练 时间	测试 时间
1	ELM	4.498 8	1.563 6	5.989 5	0.037 4	0.003 1
	迟滞 ELM	4.197 1	1.580 1	5.121 8	0.229	0.003 1
2	ELM	5.467 8	1.847 9	6.119 1	0.018 7	0.003 1
	迟滞 ELM	4.932 6	1.814 5	5.142 3	0.135	0.003 1

通过对表2中的数据进行对比分析,可以得到:不同的训练样本对预测结果是有影响的,但是,对于同一样本来说,迟滞 ELM 的预测均方误差还是要小于 ELM。

5 结束语

为了应对由于世界经济的高速发展带来的能源危机,世界各国大力发展清洁的可再生的风能。风能的发展与风电场的出力息息相关,风速预测又是进行风电场出力预测的前提条件。目前常用的风速预测方法预测精度不高、预测时间较短。为了改善风速时间序列的预测精度和预测时间,提出了一种基于迟滞 ELM 模型的风速预测方法,并且进行了对比实验。实验结果表明,与利用 ELM 模型进行短期风速预测的结果相比,利用迟滞 ELM 模型进行短期风速预测的预测精度更高,预测时间更短。因此,迟滞 ELM 模型能够有效减小风速时间序列的预测误差,提高预测精度以及减少预测时间,从效率上来说学习系统具有更好的误差和泛化能力。

参考文献:

[1] 邵 璠. 风电场短期风速预测研究[D]. 乌鲁木齐:新疆大学,2008.

[2] 韩 爽,杨勇平,刘永前. 三种方法在风速预测中的应用研究[J]. 华北电力大学学报,2008,35(3):57-61.

[3] 陈丹丹,李永光,张 莹,等. 风电场风速预测的研究方法[J]. 上海电力学院学报,2011,27(3):247-251.

[4] 蒋金良,林广明. 基于 ARIMA 模型的自动站风速预测[J]. 控制理论与应用,2008,25(2):374-376.

[5] 黄小华,李德源,吕文阁,等. 基于人工神经网络模型的风速预测[J]. 太阳能学报,2011,32(2):193-197.

[6] 吕 涛,唐 巍,所 丽. 基于混沌相空间重构理论的风电场短期风速预测[J]. 电力系统保护与控制,2010,38(21):113-117.

[7] 谷兴凯,范高峰,王晓蓉,等. 风电功率预测技术综述[J]. 电网技术,2007,31(S2):335-338.

[8] 黄文杰,傅 砾,肖 盛. 采用改进模糊层次分析法的风速预测模型[J]. 电网技术,2010,34(7):164-168.

[9] Wang X,Guo P,Huang X. A review of wind power forecasting models[J]. Energy Procedia,2011,12(39):770-778.

[10] Huang G B,Zhou H,Ding X,et al. Extreme learning machine for regression and multiclass classification[J]. IEEE Transactions on Systems Man & Cybernetics Part B Cybernetics A Publication of the IEEE Systems Man & Cybernetics Society, 2012,42(2):513-529.

[11] Ronay A K,Fink O. Two machine learning approaches for short-term wind speed time-series prediction[J]. IEEE Transactions on Neural Networks & Learning Systems,2015, 27(8):1734-1747.

[12] 王智慧. BP 神经网络和 ELM 算法研究[D]. 杭州:中国计量学院,2012.

[13] 李 彬,李贻斌. 基于 ELM 学习算法的混沌时间序列预测[J]. 天津大学学报,2011,44(8):701-704.

[14] Wang X,Qiao Q,Wang Z. Solving optimization problems based on neural network with hysteretic activation function[C]//International symposium on advances in neural networks. [s. l.]:[s. n.],2005:745-749.

[15] Huang Guangbin. Extreme learning machine;a new learning scheme of feedforward neural networks[C]//IEEE international joint conference on neural networks. [s. l.]:IEEE, 2004:985-990.

[16] Luzanin O,Plancak M. Hand gesture recognition using low-budget data glove and cluster-trained probabilistic neural network[J]. Assembly Automation,2014,34(1):94-105.

[17] Huang G B,Zhu Q Y,Siew C K. Extreme learning machine: theory and applications[J]. Neurocomputing,2006,70(1):489-501.

[18] Xie S L,Zhang Y H,Chen C H,et al. Identification of nonlinear hysteretic systems by artificial neural network[J]. Mechanical Systems and Signal Processing,2013,34(1-2):76-87.

[19] Nakaguchi T,Jin No K,Tanaka M. Hysteresis neural networks for N-queens problems[J]. IEICE Transactions on Fundamentals of Electronics Communications & Computerences, 1999,82(9):1851-1859.

[20] Nakaguchi T,Jin'No K,Tanaka M. Hysteresis neural networks for solving traveling salesperson problems[C]//IEEE interna-

从图 5 可见,采用提出算法的误差分布较传统方法较小,控制性能得到有效改善。

4 结束语

当前的时滞系统的控制采用 Smith 预估模型,当被控对象发生非线性变化和受到干扰时,控制误差的反馈调节性能变差。为此,提出一种加入前馈补偿的时滞系统鲁棒性跟踪控制器设计方法。实验结果表明,采用该方法设计的时滞系统鲁棒性跟踪控制器,构造简单,准确性好,控制性能优越,在工业过程控制等领域具有较好的应用价值。

参考文献:

[1] 董文瀚,马振强,解武杰,等. 非匹配不确定离散系统的无抖振积分滑模控制[J]. 控制与决策,2015,30(12):2181-2186.

[2] 章俊,张冬梅,王辛刚. 一类参数不确定连续正系统受限控制[J]. 信息与控制,2015,44(5):592-597.

[3] Zhang J F, Han Z Z, Wu H. Robust finite-time stability and stabilization of switched positive systems[J]. IET Control Theory and Applications,2014,8(1):67-75.

[4] Ngoc P H A. Stability of positive differential systems with delay[J]. IEEE Transactions on Automatic Control, 2013, 58(1):203-209.

[5] Smith S M, White K, Xu M. Fuzzy logic flight and bottom fol-

lowing controllers for the ocean voyager II AUV[C]//Proceedings of the joint conference on information sciences. Pinehurst:Association for Intelligent Machinery,1994:56-59.

[6] 程帅,孙俊喜,曹永刚,等. 多示例深度学习目标跟踪[J]. 电子与信息学报,2015,37(12):2906-2912.

[7] Zhong W, Lu H C, Yang M H. Robust object tracking via sparse collaborative appearance model[J]. IEEE Transactions on Image Processing,2014,23(5):2356-2368.

[8] Vincent P, Larochelle H, Lajoie I, et al. Stacked denoising autoencoders: learning useful representations in a deep network with a local denoising criterion[J]. Journal of Machine Learning Research,2010,11(12):3371-3408.

[9] 李天成,孙树栋. 采用双重采样的移动机器人 Monte Carlo 定位方法[J]. 自动化学报,2010,36(9):1279-1286.

[10] 穆施瑶,朱琦. 双层网络中一种新的接入控制机制[J]. 电子学报,2015,43(6):1231-1235.

[11] 万磊,张英浩,孙玉山,等. 基于自抗扰的自主水下航行器地形跟踪控制[J]. 兵工学报,2015,36(10):1943-1948.

[12] 韩京清. 自抗扰控制技术-估计补偿不确定因素的控制技术[M]. 北京:国防工业出版社,2009:243-261.

[13] Xiang M, Xiang Z R, Karimi H R. Stabilization of positive switched systems with time-varying delays under asynchronous switching[J]. International Journal of Control, Automation and Systems,2014,12(5):939-947.

[14] 张丽香,陈凤兰. 自适应 Smith 预估补偿的过热汽温控制系统[J]. 信息与控制,2015,44(5):513-518.

(上接第 135 页)

tional symposium on circuits and systems. [s. l.]: IEEE, 2000:153-156.

[21] Xia G P, Tang Z, Li Y, et al. A binary Hopfield neural network with hysteresis for large crossbar packet-switches[J]. Neurocomputing,2005,67(1):417-425.

[22] Chen L, Lian D H, Wang T. Synchronization stability of a distributed hysteresis controlled system[C]//Industrial electronics and applications. [s. l.]:[s. n.],2012:1616-1620.

[23] Jin N K. Dynamical hysteresis neural networks for graph coloring problem[C]//Applications of nonlinear dynamics. [s. l.]:[s. n.],2009:331-340.

[24] Bharitkar S, Mendel J M. The hysteretic Hopfield neural network[J]. IEEE Transactions on Neural Networks, 2000, 11(4):879-887.

[25] Thangavel P, Gladis D. Hysteretic Hopfield network with dynamic tunneling for crossbar switch and N-queens problem[J]. Neurocomputing,2007,70:2544-2551.

[26] Sun M, Xu Y Q, Dai X F, et al. Noise-tuning-based hysteretic noisy chaotic neural network for broadcast scheduling problem in wireless multihop networks[J]. IEEE Transactions on Neural Networks and Learning Systems, 2012, 23(12):1905-1918.

[27] Wang Hongfei, Xiu Chunbo. Hysteretic neural network and its application in the prediction of the wind speed series[C]//26th Chinese control and decision conference. [s. l.]:[s. n.],2014:985-990.

[28] 修春波,刘畅,郭富慧,等. 迟滞混沌神经元/网络的控制策略及应用研究[J]. 物理学报,2014,64(6):060504.