

文章编号: 0254-0096(2018)02-0315-09

基于投影寻踪主成分分析和耦合模型的风电功率预测

王 聪¹, 张宏立¹, 范文慧²

(1. 新疆大学电气工程学院, 乌鲁木齐 830047; 2. 清华大学自动化系, 北京 100084)

摘 要: 为提高搜索精度和解决样本数据高维性问题,以数值天气预报为基础,提出一种基于投影寻踪和改进状态转移算法优化的风电功率预测模型。该方法首先选取风电场周围多个位置多个高度的气象数值信息,采用投影寻踪主成分分析方法将高维的样本数据投影到低维空间,提取主成分,再建立投影寻踪耦合模型;同时通过加入正交变换的状态转移算法优化最佳投影方向、多项式系数和阈值项,确定网络结构以确保得到最佳模型。以某风电场为实例研究,表明基于投影寻踪和改进状态转移算法的方法可靠性高,能有效解决风电功率预测中存在的预测精度低、数据非线性和高维性等实际难题。

关键词: 风电功率; 预测; 投影寻踪; 主成分分析; 耦合模型; 状态转移算法

中图分类号: TK81; TP3 **文献标识码:** A

0 引 言

由于风速的随机性和间歇性,风电的大规模并网给电力行业带来严峻挑战,其中对电力系统运行调度的影响尤为明显,因此对风电场输出功率预测问题的研究变得尤为重要。风电输出功率预测的研究,是优化电网调度、解决风电波动和实现风电与电力系统传统电源联合优化运行的关键技术之一^[1]。

风电功率预测模型主要有持续性预测模型、物理预测模型和统计预测模型等。持续性预测模型是将最新的实测值作为下一时间步长的预测值,原理来源于大气系统的惯性原理^[2],该方法虽简单易行,但预测误差较大,结果不稳定。物理预测模型和统计预测模型多是基于数值天气预报数据的方法。物理预测模型是通过考虑地形、地表、湍流等多因素和发电功率曲线,达到预测结果,该方法存在成本高,且复杂的地形因素信息不易收集等缺点。

统计预测模型主要针对短期预测,是利用风

速、风向和输出功率的实测数据构成历史统计数据,采用线性预测或非线性预测来进行预测。统计方法的误差主要来源于气象数据的随机性,计算时间短,对短期的功率预测精度高,是应用最多的一种预测方法。常用的有人工神经网络方法^[3,4]、组合预测方法^[5]、灰色预测方法^[6]及高斯预测方法^[7]等。这些方法均能得到较好的短期预测结果,但人工神经网络方法网络结构较复杂,权值不易确定;组合优化方法利用风速的季节性具有一定的误差;灰色预测方法是通过间接的预测风速和风能进而预测风力功率等。同时统计方法在利用历史数据建模时,大部分是对历史数据进行去噪^[8]和降维^[9]后直接作为输入进行建模预测,但未考虑不同的数值天气预报的差异,进行集中建模,导致较大噪声,预测精度降低^[10],此外对于高阶模型的参数估计难度也偏大。

针对以上预测方法的不足,提出基于投影寻踪主成分分析和投影寻踪耦合模型的风电功率预测模型。投影寻踪方法是将高维的数据投影到低维空间,并对低维空间的投影数据进行分析 and 处理。

收稿日期: 2015-12-25

基金项目: 国家自然科学基金(51575469; 51767022); 新疆大学博士生创新项目(XJUBSCX-2015014); 新疆维吾尔自治区研究生创新项目(XJGR12016026)

通信作者: 张宏立(1972—), 男, 博士、副教授, 主要从事群智能计算、电力系统的建模及控制方面的研究。641087385@qq.com

在遥感和光谱图像处理、农业、水利、化工等领域均得到很好应用并取得较好的效果,但在风电领域尚少有使用。本文将改进的投影寻踪方法应用到短期风电功率预测中。首先收集风电场周围多个位置不同高度的气象数值信息,通过投影寻踪将高维的数据样本投影到低维空间,利用投影寻踪主成分分析找出主要的综合指标;接近风力机叶片的风速、风向正弦、风向余弦、气温、气压和湿度;再利用基于神经网络的投影寻踪耦合模型进行预测模型的建立,将 Laguerre 多项式替代 Hermite 多项式,减少了数据区间的转换;同时利用一种新型的智能优化算法——正交状态转移算法进行最佳投影方向、多项式系数和阈值项等的优化,确定网络结构,得到最佳的预测模型。通过实例研究,该风电功率预测方法在短期风电功率预测中有很好的精度和效率。

1 建模思路

由于风电场周围不同位置不同高度的气象信息都会对风力机发电功率造成影响,在选取样本数据时综合考虑各位置信息,选取不同位置不同高度的气象数据作为原始样本。由于选取的是高维的样本数据,在进行模型预测中会降低模型的泛化能力,因此引入投影寻踪主成分分析,选取影响风力发电的主要因素,最后根据影响因素建立投影寻踪回归与 Laguerre 神经网络的耦合预测模型。具体建模流程如图 1 所示。

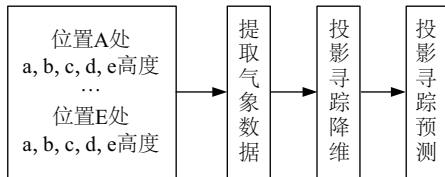


图1 建模流程图

Fig. 1 Modeling flowchart

2 选取样本数据

以每日数值天气预报时序数据(数据时间分辨率为 10 min)为基础,选取影响因素:不同位置不同高度的风速、风向余弦、风向正弦、气温、气压、湿度、辐射、地表粗糙度等,根据选取的数据建立样本记为 X 。影响风电场输出功率的 p 个影响因素作

为 p 个自变量 $\{x_1, x_2, \dots, x_p\}$, 观测 n 个样本点, 构成自变量和因变量的数据表:

$$X = [x_1, x_2, \dots, x_p]_{n \times p} \quad (1)$$

为了消除量纲的影响,需对数据进行归一化处理,对于越大越优的指标:

$$x'(i, j) = \frac{x^*(i, j) - x_{\min}(j)}{x_{\max}(j) - x_{\min}(j)}, \quad i = 1, 2, \dots, n; j = 1, 2, \dots, p \quad (2)$$

对于越小越优的指标:

$$x'(i, j) = \frac{x_{\max}(j) - x^*(i, j)}{x_{\max}(j) - x_{\min}(j)}, \quad i = 1, 2, \dots, n; j = 1, 2, \dots, p \quad (3)$$

标准化处理:

$$x(i, j) = \frac{x'(i, j) - \bar{x}'(j)}{S_{x'(j)}} \quad (4)$$

式中, $x_{\max}(j)$ 、 $x_{\min}(j)$ ——第 j 个指标的最大值和最小值; $x^*(i, j)$ ——指标特征值; $x'(i, j)$ ——指标归一化后的序列; $\bar{x}'(j)$ 、 $S_{x'(j)}$ ——第 j 个指标的均值和标准差。

3 投影寻踪主成分分析

投影寻踪是通过将高维的数据投影到低维空间,分析和处理非线性、非正态高维数据的统计方法,它是处理多因素复杂问题的一种统计方法^[11]。投影寻踪主成分分析的基本思想是,如果投影指标函数值的大于零的部分归因于前 d ($d \leq p$) 个影响因素,那么这些成分就可“取代”原来的 p 个因素,而且信息完全利用。

设 $a = \{a(1), a(2), \dots, a(p)\}$ 是 p 维单位向量,数据在 a 方向上的一维投影值 $z(i)$ 为:

$$z(i) = \sum_{j=1}^p a(j)x(i, j), \quad i = 1, 2, \dots, n \quad (5)$$

投影指标函数记为 $Q(a)$, 投影指标越大越好,这里 $Q(a)$ 就是主成分分析中协方差矩阵的最大特征值。记:

$$Q(a_1) = \max \left(\sum_{i=1}^n (z(i) - Ez)^2 / (n-1) \right) \quad (6)$$

$$Q(a_2) = \max \left(\sum_{i=1}^n (z(i) - Ez)^2 / (n-1) \right), \quad a_2 \perp a_1 \quad (7)$$

$$Q(a_3) = \max \left(\sum_{i=1}^n (z(i) - Ez)^2 / (n-1), a_3 \perp a_1, a_2 \right) \quad (8)$$

...

$$Q(a_d) = \max \left(\sum_{i=1}^n (z(i) - Ez)^2 / (n-1), a_d \perp a_1, a_2, \dots, a_{d-1} \right) \quad (9)$$

计算各个主成分:

$$F_{i'i} = \sum_{j=1}^p a_{i'i} \times x'(i,j), i = 1, 2, \dots, n; i' = 1, 2, \dots, d \quad (10)$$

构造各个评价样本的综合评价函数:

$$F_i = \sum_{i'=1}^d \alpha_d F_{i'i}, i = 1, 2, \dots, n; i' = 1, 2, \dots, d \quad (11)$$

式中, Ez ——投影值的平均值, $\alpha_1, \alpha_2, \dots, \alpha_d$ ——第1、第2...第 d 主成分的贡献率。

对于投影指标函数的优化,本文选用一种新型的智能优化算法——状态转移算法进行优化。

4 投影寻踪耦合模型

投影寻踪回归方法的主要目的是解决高维空间的回归问题,利用若干个岭函数加权的形式来逼近回归函数,具有预测的功能^[12]。将岭函数看作BP网络的神经元函数且不考虑阈值,则该模型可看作是一个三层的网络模型。但投影寻踪回归模型是借助高维数据的投影寻踪思想,BP网络则是模拟神经元的功能。从建模思路来看,都未假定输入到输出的映射关系,只对组成映射关系的单元进行处理,都存在结构不易确定、收敛速度慢等缺点。为了解决这两类模型的缺陷,取长补短,建立基于神经网络的投影寻踪耦合模型。

基于神经网络的投影寻踪耦合(coupling error back propagation and projection pursuit, BPPP)模型的结构等同于含一个BP网络,输出层的网络权值隐含在神经元函数的拟合多项式系数中。

神经网络耦合模型主要有以下优势:

1) 模型结构

耦合模型与BP网络相比,采用不同的神经元函数,使结构多了一个信息的载体,能取得更高的精度;与回归模型相比,多出阈值项,可加快模型的收敛速度。

模型的网络结构为一种多个输入,单个输出,

含一个隐层的简单形式,神经元函数多采用Hermite多项式。为了减少使用庞大的函数表,又能保证逼近的精度,将Laguerre多项式^[13]代替Hermite多项式。两种多项式相同的性质有:1)正交性:对于给定的权函数,两种多项式的序列均为正交序列;2)完备性:对于所有满足特定函数所构成的完备空间中,两种多项式的序列都可构成一组基;3)两种多项式都具有拟合概率密度函数的特点。此外,与Hermite多项式相比,Laguerre多项式的定义区间为 $[0, +\infty]$,在实际预测中避免了实际数据区间的换算。因此,Laguerre多项式代替Hermite多项式不仅满足耦合模型的要求,而且能大大减少数据处理的复杂性。

将Laguerre多项式代替Hermite多项式,采用可变阶的正交多项式拟合其中的一维岭函数,数学表达式为:

$$h_r(z) = (r!)^{\frac{1}{2}} \pi^{\frac{1}{4}} 2^{\frac{1-r}{2}} H_r(z) \varphi(z), -\infty < z < \infty \quad (12)$$

式中, $r!$ —— r 的阶乘; $z = a^T X$; φ ——标准高斯方程, $\varphi = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$, $H_r(z)$ ——Laguerre多项式采用递推的形式给出,如 $H_0(z) = 1$, $H_1(z) = 1 - Z$, $H_r(z) = (2k - 1 - x)H_{r-1}(z) - (r - 1)^2 H_{r-2}(z)$ 。

2) 学习策略

与BP网络相比,学习策略首先是对投影方向 a 的优化,而BP则是由一个初始点出发,迭代得到最优解;同时耦合模型还包括了阈值与多项式系数的优化,完成一个神经元函数的优化后再开始下一神经元的优化。

基于神经网络的投影寻踪耦合模型在数据降维和模型预测过程存在一个求解最佳投影方向、阈值项和多项式系数的高维有约束的优化问题。为了在保证精度的前提下获得高效率,本文选用新型的智能算法——正交状态转移算法进行寻优。

5 改进的状态转移算法

5.1 状态转移算法

状态转移算法(state transition algorithm, STA)^[14],借鉴状态及状态转移的概念,将待优化问题的解理解为状态,优化算法的思想描述为状态转移,则解待优化问题的过程便是一个状态转移过程。如式(13):

$$\begin{cases} x_{k+1} = A_k x_k + B_k u_k \\ y_k = f(x_{k+1}) \end{cases} \quad (13)$$

式中, x_k ——表示一个状态, 对应待优化问题的一个解, $x_k \in \mathbf{R}^n$; A_k 、 B_k ——状态转移矩阵, 可理解为优化算法思想的算子, A_k 、 $B_k \in \mathbf{R}^{n \times n}$; u_k ——状态 x_k 及历史状态的函数, $u_k \in \mathbf{R}^n$; f ——适应度函数或评价函数。

状态转移算法为了提高全局搜索能力和局部搜索能力, 并平衡两者之间的关系, 定义旋转变换、平移变换、伸缩变换、坐标变换 4 种操作算子。

旋转变换 (rotation transformation, RT):

$$x_{k+1} = x_k + \alpha \frac{1}{n \|x_k\|_2} \mathbf{R}_r x_k \quad (14)$$

平移变换 (translation transformation, TT):

$$x_{k+1} = x_k + \beta \mathbf{R}_t \frac{x_k - x_{k-1}}{\|x_k - x_{k-1}\|_2} \quad (15)$$

伸缩变换 (expansion transformation, ET):

$$x_{k+1} = x_k + \gamma \mathbf{R}_e x_k \quad (16)$$

坐标变换 (axesion transformation, AT):

$$x_{k+1} = x_k + \delta \mathbf{R}_a x_k \quad (17)$$

式中, α 、 β 、 γ 、 δ ——正常数的因子; \mathbf{R}_r ——在 $[-1, 1]$ 之间均匀分布的随机矩阵, $\mathbf{R}_r \in \mathbf{R}^{n \times n}$; \mathbf{R}_t ——元素在 $[0, 1]$ 之间的随机变量, $\mathbf{R}_t \in \mathbf{R}$; \mathbf{R}_e ——元素服从高斯分布的随机对角阵, $\mathbf{R}_e \in \mathbf{R}^{n \times n}$; \mathbf{R}_a ——服从高斯分布的随机对角阵且仅有一个随机位置为非零值, $\mathbf{R}_a \in \mathbf{R}^{n \times n}$ 。

5.2 改进策略

为了进一步保持搜索范围的广度和搜索的多样性, 本文在原有 4 种操作的基础上加入正交变换 (orthogonal transformation, OT)。在每次完成状态变换后, 从整体规模 $P(t)$ 中选取适应度值较差的 15% 的个体 $D(t)$, 并对它们进行正交变换, 得到标准正交基下的正交矩阵 x' , 对于任意 $x \in D(t)$, 如果它的正交 x' 适应值更优, 则 x 被 x' 取代; 否则 x 保留下来, 即只有有效的正交操作才能使好的解传递下去。

对任意 η , $\mu \in x_k$, 有 $(x' \eta, x' \mu) = (\eta, \mu)$, 则 x' 为 x_k 的正交变换, 且 $x' \varepsilon = \varepsilon x'$, 其中 ε 是 x_k 的标准正交基。

6 风电功率预测模型的建立

风速是影响风电场输出功率的重要参量之一,

除此之外, 风向、温度、湿度、地表粗糙度等都对风电场输出功率有一定的影响。通过文献[15]对影响风电场输出功率的影响因素分析, 选取影响因素: 不同位置不同高度的风速、风向余弦、风向正弦、气温、气压、湿度、辐射、地表粗糙度等, 分别以 10 min 为一个采样点的气象数据和输出功率作为样本数据。建模流程如下:

1) 样本数据的无量纲处理: 首先按照式(2)和式(3)进行归一化处理, 其次按照式(4)进行标准化处理;

2) 主成分的提取:

步骤 1 对样本数据进行求取投影值, 如式(5);

步骤 2 构造投影指标函数, 如式(6)~式(9);

步骤 3 计算各个主成分, 如式(10);

步骤 4 构造各评价样本的综合评价函数, 如式(11);

3) 确定影响风力发电功率的主要影响因素;

4) 构造投影寻踪耦合模型:

①步骤 1 投影数据的构造。假设实际风电场输出功率为因变量 $y_i (i=1, 2, \dots, n)$, 影响风电场输出功率的气象因子 (即 2) 中提取的主成分因素) 作为 d 个自变量 $\{x_1, x_2, \dots, x_d\}$, 观测 n 个样本点, 构成自变量和因变量的数据表 $X' = [x_1, x_2, \dots, x_d]_{n \times d}$ 和 $Y = [y]_{n \times 1}$ 。

②步骤 2 投影值的计算。设 $a = (a_1, a_2, \dots, a_d)$ 为最佳投影方向, 投影寻踪就是将 X' 投影到 a 上, 得到投影值 z_i :

$$z_i = \sum_{j=1}^d a_j x_{ij} - \theta, \quad i=1, 2, \dots, n; j=1, 2, \dots, d \quad (18)$$

式中, θ ——阈值项。

③步骤 3 对散点 (z, y) 用基于正交 Laguerre 多项式拟合, Laguerre 多项式的投影寻踪耦合模型为:

$$\hat{y} = \sum_{i=1}^m \sum_{j=1}^R c_{ij} h_{ij}(z), \quad i=1, 2, \dots, m; j=1, 2, \dots, R \quad (19)$$

④步骤 4 投影指标函数的建立。在投影寻踪耦合模型预测过程中, 存在 3 个优化参数 a, c, θ 。在优化投影方向 a 时, 要求 z_i 应尽可能大地提取 x_{ij} 中的变异信息, 即要求 z_i 的标准差 S_a 尽可能大, 局部密度 D_a 尽可能达到最大; 同时还要考虑

多项式系数 c 和阈值项 θ 的优化问题,因此可通过求解投影指标函数最小化问题来估计最佳 a 、 c 、 θ 即:

$$\min Q(a, \theta, c) = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (20)$$

约束条件为:

$$\text{s.t.} \sum_{j=1}^d a_{mj}^2 = 1, 1 \geq a(j) \geq -1 \quad (21)$$

⑤步骤5 参数优化。利用改进的状态转移算法进行优化求解,将待优化的参数 a 、 θ 、 c ,作为状态转移算法的状态,将求解 a 、 θ 、 c 的过程作为状态的变化;

⑥步骤6 计算第一次拟合误差 $r_1 = y - \hat{y}$,如果满足要求则输出模型参数,否则进行步骤7;

⑦步骤7 用 r_1 代替 y ,回到2),开始下一个岭函数优化,直至满足要求,停止增加岭函数个数,输出最后结果。

7 算例结果与分析

采用新疆某风电场6个月的数值气象预报数据和1.5 MW的风力机输出功率作为样本,选择时间分辨率为10 min。数值天气预报选取每日12:00发布的预报。

本文主要完成以下分析。

1) 主成分的提取

2)为了验证改进算法的有效性,分别采用粒子群(PSO)、状态转移算法(STA)、改进状态转移算法(OTSTA)进行优化投影寻踪耦合模型(BPPP),并比较这3类优化模型的预测效果。

3)为了验证该耦合模型的有效性,将改进算法优化的耦合模型(BPPP)、BP神经网络模型(BP)和投影寻踪回归模型(PPR)的预测效果进行比较分析。由于采样点太多,预测曲线图和预测误差图只选取了其中一部分。

为了对预测结果精度和模型的有效性判断,本文采用平均相对误差 MAE 、平均百分比绝对误差 $NMAE$ 和百分比均方根误差 $NRMSE$,定义如下:

$$MAE = \frac{1}{N} \sum_{i=1}^N \left| \frac{\hat{y}_i - y_i}{y_i} \right| \times 100\% \quad (22)$$

$$NMAE = \frac{1}{N} \sum_{i=1}^N \left| \frac{\hat{y}_i - y_i}{y_r} \right| \times 100\% \quad (23)$$

$$NRMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N \left(\frac{\hat{y}_i - y_i}{y_i} \right)^2} \times 100\% \quad (24)$$

式中, N ——数据长度; y_i ——实际功率; \hat{y}_i ——预测功率; y_r ——装机容量。

7.1 投影寻踪主成分分析

以数值天气预报数据为基础,采用式(5)~式(11)的方法进行主成分分析。对选取的样本数据建立投影寻踪主成分分析模型,由图2可知,前6个主成分几乎涵盖所有的特征,所以本文共提取6个主要的影响因素。

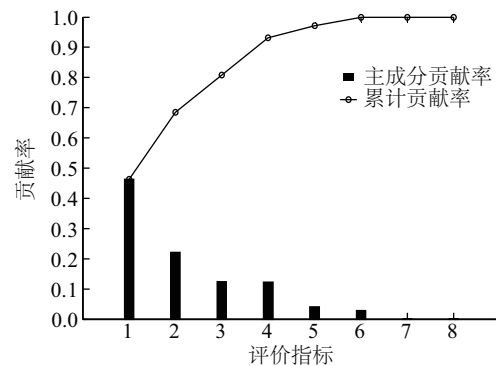


图2 主成分贡献率

Fig. 2 Contribution rates of principal components

分别为接近风力机叶片的风速、风向正弦、风向余弦、气温、气压和湿度。6个投影指标函数值和贡献值(α)如表1所示。

表1 主成分指标函数值

Table 1 Principal components index function value			
$Q(a)$	$Q(a_1)$	$Q(a_2)$	$Q(a_3)$
值	7.0469	3.3744	1.8819
α	46.32%	22.18%	12.37%
$Q(a)$	$Q(a_4)$	$Q(a_5)$	$Q(a_6)$
值	1.8576	0.6283	0.4245
α	12.21%	4.13%	2.79%

7.2 算法有效性分析

算法参数设置为:执行搜索群体规模(SE)设为80,迭代次数为500,间歇性交流频率(CF)为50, α 为 $1 \rightarrow 10^{-5}$, β 、 γ 、 δ 均为1, f_c 为2。粒子群算法参数设置为:种群规模80,迭代次数500,惯性权重0.9,加速因子均为2。

由于智能算法具有一定的随机性,对 3 种算法优化的 BPPP 预测模型独立运行 10 次,投影指标函数结果见表 2。

表 2 3 种算法的优化结果

Table 2 The results of three algorithms

算法	最优值	最差值	平均值	平均 用时/s	偏差/ %
OTSTA	2.68	2.60	2.604	87.8	3.08
STA	2.62	2.51	2.611	79.6	4.20
PSO	2.62	2.43	2.521	156.4	4.61

由表 2 可知,OTSTA 求得投影指标函数最优为 2.68,10 次运算的偏差最小为 3.08%。PSO 算法和 STA 算法的最优值和偏差接近,但 PSO 耗时 156.4 s,是 STA 算法耗时的 1.9 倍。OTSTA 算法的最优值比 STA 高了 0.06,偏差率下降到 3.08%,耗时却仅比 STA 多了 8.2 s。综上,OTSTA 优化的效果最好,好的优化结果可以保证一个好的预测精度。

根据主成分个数及算法训练的结果,BPPP 模型经训练后均采用 6-6-1 的网络结构。

根据 3 种算法优化得到的参数,分别用 OTSTA-BPPP、STA-BPPP 和 PSO-BPPP 进行预测风电功率。得到平均绝对误差最小时的功率误差曲线和功率预测曲线如图 3 和图 4。

由误差曲线和功率预测曲线可看出,在风电功率预测中,总体上 OTSTA-BPPP 模型具有较好的预测结果。OTSTA 优化的模型误差绝对值都小于 150 kW,且预测结果较稳定,相对于 STA 和 PSO,OTSTA 鲁棒性更好。

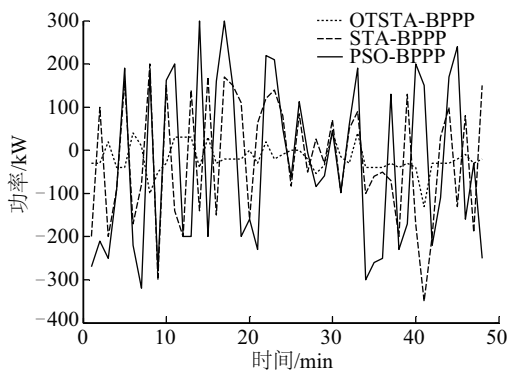


图 3 预测功率绝对误差曲线

Fig. 3 Power prediction error curves

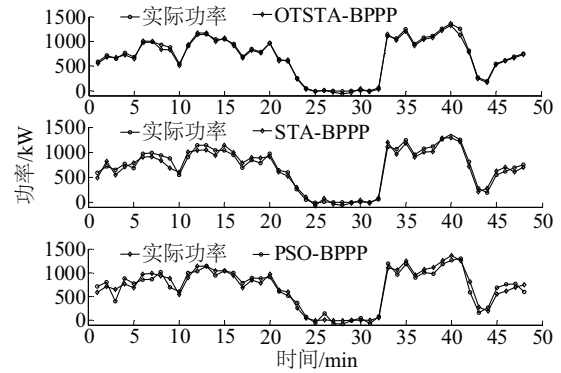


图 4 功率预测比较曲线

Fig. 4 Power prediction and comparison curves

表 3 给出了 3 种优化方法得到的预测误差。由表 3 可知,3 种优化方法得到的预测 MAE 均在 15% 以下, NMAE 和 NRMSE 均在 20% 以下。而 OTSTA 优化的耦合模型的 MAE、NMAE 和 NRMSE 都明显下降,能得到更高的预测精度,证明了该改进算法的有效性。

OTSTA 算法融合了 STA 算法的计算耗时短,结构简单的优点,又提升了算法的全局搜索能力,计算的速度和精度都要优于 PSO 算法,在保证效率的同时精度又优于 STA。

表 3 预测误差比较

Table 3 Forecasting error comparison

网络模型	预测误差/%		
	MAE	NMAE	NRMSE
OTSTA-BPPP	12.35	15.50	17.12
STA-BPPP	13.51	16.86	18.61
PSO-BPPP	13.74	16.92	18.49

7.3 模型有效性分析

为了验证该耦合模型的有效性,选取改进算法优化的投影寻踪耦合模型(BPPP)、投影寻踪回归模型(PPR)和 BP 网络(BP)进行预测。

通过仿真得到如图 5 的功率预测曲线和图 6 的相对误差散点图。

由图 5 可知,BPPP 耦合模型的功率预测曲线与实际值基本吻合,预测精度高,稳定性好,个别时刻点的预测误差较大,或是由于该时刻存在其他影响因素,未考虑在内,总体上该耦合模型达到很好的预测结果。而 PPR 模型和 BP 模型的预测结果

则存在较大误差。由图 6 的相对误差散点图可看出,BPPP 模型的预测误差在 10%以内的占 82.3%,在 20%以内的点占 88.2%,最大误差点小于 60%。PPR 模型的预测误差集中在 30%以内,最大误差为 48%,BP 神经网络模型的误差多散布在 40%以内,最大误差点达到 90%。

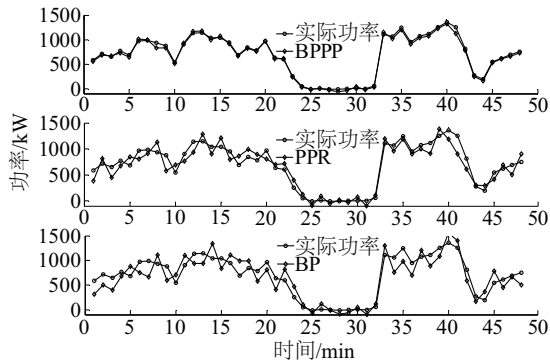


图5 功率预测比较曲线

Fig. 5 Power prediction and comparison curves

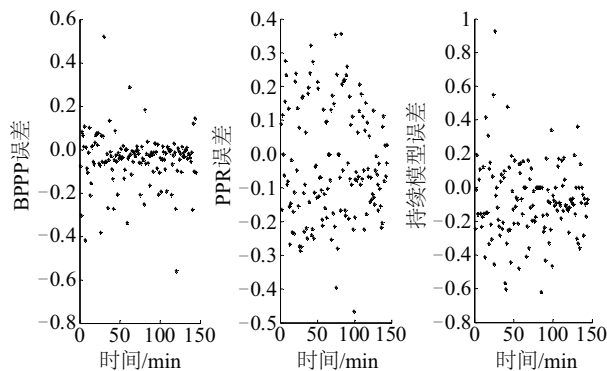


图6 预测相对误差点分布

Fig. 6 Power prediction error distribution

BPPP 与 PPR、BP 神经网络模型相比具有更好的预测结果和更强的泛化能力,证明了结合 BP 网络和回归模型优点的耦合模型的有效性。表 4 给出了 3 种模型的预测误差。

表4 预测误差比较

Table 4 Forecasting error comparison

网络模型	预测误差/%		
	MAE	NMAE	NRMSE
BPPP	12.35	15.50	17.12
BP	19.44	23.33	24.20
PPR	14.41	21.65	22.63

由表 4 可知,BPPP 模型与 PPR 和 BP 神经网络模型相比,相同的测试样本,MAE、NMAE 和 NRMSE 均较小,保持在 20%以内,证明了 BPPP 预测的结果更接近于实际的输出功率,精度高、稳定性好。

8 结 论

1)本文首先选用风电场周围不同位置不同高度的气象数据,采用投影寻踪主成分分析的方法确定影响风力发电功率的主要因素,再采用基于神经网络的投影寻踪耦合模型,融合了神经网络与投影寻踪回归的共同优势,取长补短,并选用 Laguerre 多项式作为神经元函数,是一种新型的网络形式。实验结果表明,对风电功率预测问题,BPPP 模型比 PPR 模型和 BP 网络模型具有更高的预测精度,更强的泛化能力。

2)状态转移算法是一种结构简单的智能算法,加入正交变换的状态转移算法,增加了状态的多样性,能进一步避免优化过程中陷入局部最优问题的发生,具有高于 STA 和 PSO 的优化性能,提高了 OTSTA-BPPP 模型的短期风电功率预测的准确性。

3)本文提出基于投影寻踪和状态转移算法的耦合模型,在风力发电功率预测过程中,耗时短,精度高,可有效减小预测误差,提高预测精度。

[参考文献]

- [1] 王彩霞,鲁宗相,乔 颖,等.基于非参数回归模型的短期风电功率预测[J].电力系统自动化,2010,34(16):78—82,91.
- [1] Wang Caixia, Lu Zongxiang, Qiao Ying, et al. Short-term wind power forecast based on non-parametric regression model[J]. Automation of Electric Power Systems, 2010, 34(16): 78—82, 91.
- [2] Wegley H, Formica W. Test applications of a semi-objective approach to wind forecasting for wind energy applications[R]. PNL-4403, Pacific Northwest Laboratory, 1983.
- [3] Bhaskar K, Singh S N. AWNN-assisted wind power forecasting using feed-forward neural network[J]. IEEE

- Transactions on Sustainable Energy, 2012, 3(2): 306—315.
- [4] 刘瑞叶, 黄磊. 基于动态神经网络的风电场输出功率预测[J]. 电力系统自动化, 2012, 36(11): 19—22, 37.
- [4] Liu Ruiye, Huang Lei. Wind power forecasting based on dynamic neural networks[J]. Automation of Electric Power Systems, 2012, 36(11): 19—22, 37.
- [5] 张国强, 张伯明. 基于组合预测的风电场风速及风电功率预测[J]. 电力系统自动化, 2009, 33(18): 92—95, 109.
- [5] Zhang Guoqiang, Zhang Boming. Wind speed and wind turbine output forecast based on combination method [J]. Automation of Electric Power Systems, 2009, 33(18): 92—95, 109.
- [6] El-Fouly T H M, El-Saadany E F, Salama M M A. Grey predictor for wind energy conversion systems output power prediction[J]. IEEE Transactions on Power Systems, 2006, 21(3): 1450—1452.
- [7] Chen N, Qian Z, Nabney I T, et al. Wind power forecasts using Gaussian processes and numerical weather prediction[J]. IEEE Transactions on Power Systems, 2014, 29(2): 656—665.
- [8] Liu Da, Niu Dongxiao, Wang Hui, et al. Short-term wind speed forecasting using wavelet transform and support vector machines optimized by genetic algorithm [J]. Renewable Energy, 2014, 62: 592—597.
- [9] Skittides C, Früh W G. Wind forecasting using principal component analysis[J]. Renewable Energy, 2014, 69: 365—374.
- [10] 王勃, 冯双磊, 刘纯. 基于天气分型的风电功率预测方法[J]. 电网技术, 2014, 38(1): 93—98.
- [10] Wang Bo, Feng Shuanglei, Liu Chun. Study on weather typing based wind power prediction [J]. Power System Technology, 2014, 38(1): 93—98.
- [11] Friedman J H, Tukey J W. A projection pursuit algorithm for exploratory data analysis[J]. IEEE Transactions on Computer, 1974, C-23(9): 881—890.
- [12] Ferraty F, Goia A, Salinelli E, et al. Functional projection pursuit regression [J]. Test, 2013, 22(2): 293—320.
- [13] 杨洪耕, 秦东, 张正书, 等. 用Laguerre多项式描述谐波随机求和问题[J]. 电网技术, 2005, 29(14): 26—29.
- [13] Yang Honggeng, Qin Dong, Zhang Zhengshu. A study on summation of harmonics with random phases angles based on Laguerre polynomials[J]. Power System Technology, 2005, 29(14): 26—29.
- [14] Zhou Xiaojun, Yang Chunhua, Gui Weihua. A new transformation into state transition algorithm for finding the global minimum [A]. International Conference on Intelligent Control and Information Processing (ICICIP) [C], Harbin, China, 2011, 674—678.
- [15] 范高峰, 王伟胜, 刘纯, 等. 基于人工神经网络的风电功率预测[J]. 中国电机工程学报, 2008, 28(34): 118—123.
- [15] Fan Gaofeng, Wang Weisheng, Liu Chun, et al. Wind Power prediction based on artificial neural network [J]. Proceedings of the CSEE, 2008, 28(34): 118—123.

WIND POWER PREDICTION BASED ON PROJECTION PURSUIT PRINCIPAL COMPONENT ANALYSIS AND COUPLING MODEL

Wang Cong¹, Zhang Hongli¹, Fan Wenhui²

(1. College of Electrical Engineering, Xinjiang University, Urumqi 830047, China;

2. Department of Automation, Tsinghua University, Beijing 100084, China)

Abstract: In order to improve prediction accuracy and solve the problem of high dimensional sample data, a wind power predicting model based on projection pursuit and improved state transition algorithm was proposed based on numerical weather forecast. Firstly, the numerical weather information of multiple heights and locations around the wind farm were chosen, the projection pursuit principal component analysis method was used to project high dimensional sample data into low dimensional space, extract principal component, then establish the projection pursuit coupled model. Meanwhile, the state transition algorithm added orthogonal transformation was used to optimize the best projection direction, polynomial coefficient and threshold, and determine the network structure to ensure the best model. The case study of certain wind farm shows that the model based on projection pursuit and state transition algorithm has high reliability and can effectively solve the practical problems of low prediction accuracy, data nonlinearity and high dimensionality in wind power prediction.

Keywords: wind power; prediction; projection pursuit; components analysis; coupling error back propagation model; state transition algorithm