

# 基于随机过程自相关性的风速预测模型分析

史可琴<sup>1</sup>, 王方雨<sup>2</sup>, 梁琛<sup>3</sup>, 刘文颖<sup>2</sup>

(1. 国网陕西省电力公司, 陕西省 西安市 710048;

2. 新能源电力系统国家重点实验室(华北电力大学), 北京市 昌平区 102206;

3. 国网甘肃省电力公司, 甘肃省 兰州市 730050)

## A New Wind Speed Prediction Model Based on Random Process Considering Autocorrelation

SHI Keqin<sup>1</sup>, WANG Fangyu<sup>2</sup>, LIANG Chen<sup>3</sup>, LIU Wenying<sup>2</sup>

(1. State Grid Shaanxi Electric Power Company, Xi'an 710048, Shaanxi Province, China;

2. State Key Laboratory of Alternate Power System With Renewable Energy Resources (North China Electrical Power University), Changping District, Beijing 102206, China;

3. State Grid Gansu Electric Power Company, Lanzhou 730050, Gansu Province, China)

**ABSTRACT:** It is of increasing importance to predict wind speed in times of wind energy widely utilized. This paper proposed a new wind speed prediction model based on random process, taking autocorrelation of raw data into consideration. Concept of random process was introduced in this model in parallel with traditional probability model. This model took wind speed at every moment as a random variable and utilized multidimensional distribution function of random process to describe wind changing process. In order to feature wind's time autocorrelation, this paper introduced continuous Markov chain model and solved Markov chain state transfer function based on random process model to characterize evolution law of wind speeds between adjacent times. Simulation shows that the model can better simulate wind speed distribution and predict wind speed with higher confidence level.

**KEY WORDS:** random process; normal distribution process model; continuous Markov chain model; state transfer function

**摘要:** 提出一种基于随机过程自相关性的风速预测模型, 在传统概率模型分析的基础上引入了随机过程的概念, 将每个时刻的风速均看作是一个随机变量, 利用随机过程多维分布函数的统计特性描述风速过程。为了表征风速在时间上的自相关特性, 引入连续马尔科夫模型, 依据前述随机过程模型, 求解马尔科夫模型的状态转移函数, 从而表征风速相邻时刻间的演化规律, 并从理论上证明了该模型具有较好保持自相关特性的能力。仿真表明, 利用该模型能够更好地模拟风速分布, 预测风速大小, 并且具有良好的置信度。

**关键词:** 随机过程; 正态过程模型; 连续马尔科夫链模型; 状态转移函数

**DOI:** 10.13335/j.1000-3673.pst.2016.0949

## 0 引言

风速大小在时间和空间上不仅具有较大的波动性和随机性, 还具有时间上的概率分布特性以及延续性和自相关特性。因此可以采用相应的数学方法通过对历史数据进行分析从而对下一时刻风速进行预测。然而要准确的反应风速大小的变换规律, 通常需要对大量的历史“熟数据”进行统计分析。但在实际情况下, 由于硬件限制等原因, 大量的历史数据通常难以获得, 因此需要采用数学模型对风速进行模拟, 产生风速的时间序列, 以便准确反映风速的波动和变换特性。

当对风速进行模拟预测时, 产生的时间序列要和原始的数据保持相近的统计特性。传统研究方法主要有2种: 概率密度法和时间序列法。前者采用特定的概率密度函数描述某一地区风速分布, 分为单一概率模型和混合概率分布模型, 如文献[1]采用双参数威尔布分布模型对土耳其的风能资源以及风能可利用容量进行了研究分析, 文献[2]采用混合概率模型为基础对台湾地区风速分布进行了讨论。相对单一概率分布模型提高了精度, 文献[3-5]则从不同概率分布函数的角度分别研究了风能利用系数、风电场等效容量以及含风电场的充裕度评估等科学问题。该模型方法简单, 但精度和适应性较差。时序风速分析法主要包括时间序列分析和离散马尔科夫链模型, 如文献[6]运用离散马尔科夫链模型

**基金项目:** 国家科技支撑计划项目(2015BAA01B04); 国家电网公司科技项目(522727160002)。

Project Supported by the National Science and Technology Support Program (2015BAA01B04); Science and Technology Foundation of SGCC (522727160002).

分析了风速时间序列，研究了各阶模型的自相关特性，文献[7]研究了预测区间长度以及模型阶数对自相关特性的影响等。该方法考虑了风速的自相关特性，但存在“精度”和“速度”选择困难的问题。

针对上述问题，本文提出一种计及随机过程自相关性的风速预测模型，首先将风速的时间过程看作是一个随机过程，认为每个时刻的风速均是服从某个分布的随机变量，提高了该时刻风速预测的准确度，根据不同的分布特性将随机过程划分为正态随机过程和非正态随机过程。为了表征风速间的自相关特性，引入连续马尔科夫模型，基于上述随机过程模型计算马尔科夫模型的状态转移函数，从而体现风速在时间上的相关性。

## 1 随机过程简介

随机过程是一族随机变量的有机组合，定义：随机过程为与参数  $t$  有关的一族随机变量，其中参数  $t \in T$ ， $T$  为无限实数集，记作  $\{X(t), t \in T\}$ 。在随机过程中，对于每个  $t \in T$ ， $X(t)$  均可看作是服从某一分布的随机变量。随机过程所有可能状态的集合构成该过程的状态空间。其可以用随机过程分布函数族的概念进行描述，定义：对任意  $n$  个不同的时刻，可引入  $n$  维随机变量  $[X(t_1), X(t_2), \dots, X(t_n)]$ ，它的分布函数可记作：

$$F_X(x_1, \dots, x_n; t_1, \dots, t_n) = P\{X(t_1) \leq x_1, \dots, X(t_n) \leq x_n\} \quad (1)$$

对于固定的  $n$ ，称上式为该随机过程的  $n$  维分布函数，该描述可以给出随机过程在不同时刻随机变量的统计联系。当  $n=1$  时，称为该随机过程在某时刻  $t$  的一维分布函数。为了方便分析，我们给出关于随机过程的两个定理：

**定理 1：**对于随机过程  $\{X(t), t \in T\}$ ，若该族的每一个随机变量都服从正态分布，则它的任一有限维分布均是正态分布，即对任意的  $n \geq 1$  及任意  $t_1, t_2, \dots, t_n \in T, [X(t_1), X(t_2), \dots, X(t_n)]$  服从  $n$  维正态分布<sup>[8]</sup>。

**定理 2：**科尔莫格罗夫定理指出， $n$  维分布函数族，即  $\{F(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n), n=1, 2, \dots, t_i \in T\}$  当  $n$  充分大时可以完全表征随机过程的统计特性<sup>[8]</sup>。

## 2 风速随机过程预测模型分析

若将风速看作随机过程，其依赖于时间参数  $t$  和方向参数  $d$ ，该随机过程可记作：

$$\{V(t, d), t \geq 0, 0 \leq d \leq 2\pi\} \quad (2)$$

对于风电机组而言其迎风方向基本确定，可调角度较小，因此上述随机变量中  $d$  为常数。

### 2.1 正态随机过程模型分析

假设每个时刻的风速  $v(t)$  均服从正态分布，其概率密度函数<sup>[8]</sup>为

$$f(v)_t = \frac{1}{\sqrt{2\pi}\sigma(t, d)} \exp\left(-\frac{[v - \mu(t, d)]^2}{2\sigma(t, d)^2}\right) \quad (3)$$

因此随机过程  $V(t)$  服从

$$\{V(t, d) \sim N[\mu(t, d), \sigma(t, d)], t \geq 0, 0 \leq d \leq 2\pi\} \quad (4)$$

任意时刻风速均满足正态分布的风速过程称为标准风过程，由于  $n$  维正态随机变量  $(X_1, X_2, \dots, X_n)$  的每一个分量  $X_i, i=1, 2, \dots, n$  都是正态随机变量，反之亦然，结合定理 1，该随机过程的全部统计值完全由它的均值函数和自协方差函数所确定，并且只要确定了  $\mu(t, d)$  和  $\sigma(t, d)$  曲线，便可方便求取均值函数和方差函数。

为了方便研究，根据定理 2 可以用  $n$  维分布函数族进行表示，对于正态风速过程其联合概率密度函数为

$$f(v(t_1), v(t_2), \dots, v(t_n)) = \frac{1}{(2\pi)^{n/2} (\det C)^{1/2}} \exp\left\{-\frac{1}{2} (V - \mu)^T C^{-1} (V - \mu)\right\} \quad (5)$$

式中  $C$  为  $V(t)$  的自相关函数矩阵，且有

$$V = [v(t_1), v(t_2), \dots, v(t_n)]^T \quad (6)$$

$$U = [E(v(t_1)), E(v(t_2)), \dots, E(v(t_n))]^T \quad (7)$$

由定理 2 可知：当  $n$  充分大时该  $n$  维分布函数族可以完全表征随机风过程。

对于不服从正态分布的风速过程，采用 3.3 节方法可转为正态随机过程描述，本文不再赘述。

### 2.2 随机过程参数估计方法分析

对于某时刻  $t_k$ ，对应风速随机变量为  $V_{t_k}$ ，该随机变量的概率分布按如下步骤获得：

1) 确定风速的测量周期  $T$  及样本数据  $\{V_{t_k}\}$ 。

对于以时间  $T_f$  为周期变化的风速过程， $T = T_f$ ，则该随机变量对应的样本数据为  $\{v_{t_k}, v_{t_k+T}, v_{t_k+2T}, \dots\}$ ；对于非周期性风速过程，其风速-时间曲线为  $v(t)$ ，对其进行傅里叶积分变换可得

$$F(v(t)) = a_1 \sin\left(\frac{2\pi}{T_1} t\right) + a_2 \sin\left(\frac{2\pi}{T_2} t\right) + \dots \quad (8)$$

则该随机样本数据对应的测量周期  $T$  为  $\{T_1, T_2, \dots, T_n\}$ ，其中， $T_i$  选取  $V_i$  值较大项的对应周期，则该随机变量的样本为 1 个数据族：

$$\left\{ \frac{a_1}{\sum a_i} \{v_{t_k}, v_{t_k+T_1}, v_{t_k+2T_1}, \dots\}, \frac{a_2}{\sum a_i} \{v_{t_k}, v_{t_k+T_2}, v_{t_k+2T_2}, \dots\}, \dots \right\} \quad (9)$$

2) 确定随机变量对应的概率分布类型。

采用 K-S 检验法对周期性变化的风速样本数据及非周期各态样本数据进行正态检验, 对不满足条件的各态样本数据按 3.3 节方法进行正态变换。则该时刻随机变量分布函数为各态分布函数之和。

$$F(V_{t_k}) = \sum F_i(a_i, T_i) \quad (10)$$

由于各态数据间相关性不强, 可认为相互独立, 对于正态分布, 由概率论知识可知:

$$\mu_{t_k} = \sum \mu_i(a_i, T_i), \sigma_{t_k}^2 = \sum \sigma_i^2(a_i, T_i) \quad (11)$$

式中:  $\mu_{t_k}$ 、 $\sigma_{t_k}^2$  为随机变量  $V_{t_k}$  (正态变换后) 的参数值;  $\mu_i(a_i, T_i)$ 、 $\sigma_i^2(a_i, T_i)$  为各态样本数据参数值。

3) 对该变量进行参数估计并进行检验。

采用极大似然估计正态分布参数数值。

### 3 计及风速自相关性的预测模型分析

上节所述风速预测模型相对传统概率风速模型, 考虑了风速在时间上的分布特性, 更加合理。但是仍然没有考虑风速各时刻间的相关特性, 下面将引入连续马尔科夫(continuous state Markov chain model, CSMC)技术, 结合上述随机过程模型, 建立综合考虑随机过程自相关特性的风速预测模型。

#### 3.1 连续马尔科夫模型分析

假设  $\{V(t)\}$  为描述某一风速序列的随机过程, 则该风速序列的  $k \sim$  阶 CSMC 模型<sup>[9-10]</sup>可用式(3)表示为

$$P[V(t+1) \leq v_{t+1} | V(t) = v_t, \dots, V(t-(k-1)) = v_{t-(k-1)}, \dots, V(0) = v_0] = P[V(t+1) \leq v_{t+1} | V(t) = v_t, \dots, V(t-(k-1)) = v_{t-(k-1)}] = k(v_{t+1} | v_t, \dots, v_{t-(k-1)}) \quad (12)$$

式中  $k(v_{t+1} | v_t, \dots, v_{t-(k-1)})$  被称为 CSMC 过程的一步状态转移核, 表示该风速过程在已知前  $k$  个时刻风速分布的前提下, 下一时刻风速变量  $V(t+1)$  的概率分布函数。马尔科夫链随机过程实际上是用条件概率分布来表达风速随机过程间的自相关特性, 且当已知状态转移核, 采用拒绝抽样法<sup>[11]</sup>便可以得到下一时刻的风速数值。因此表征条件概率分布的状态转移核的求取在此模型中最为重要。

#### 3.2 正态随机过程状态转移核的求取

对正态随机过程, 可以方便地通过联合概率密度方便的求取转移密度函数, 经过推导, 其公式最终结果<sup>[12]</sup>为

$$f(v_{t+1} | v_t) = \frac{\partial^2 H(v_{t+1}, v_t)}{\partial v_{t+1} \partial v_t} / \frac{\partial F(v_t)}{\partial v_t} \quad (13)$$

式中:  $H(\cdot)$  是随机过程中相邻时刻随机变量的二维

联合概率分布;  $F(\cdot)$  为风速随机变量的边缘分布。

#### 3.3 非正态随机过程状态转移核的求取

对于服从非正态随机过程的风速模型, 很难求取该随机过程相邻时刻随机变量的联合概率分布函数, 因此不能利用式(13)计算状态转移密度函数。本文采用一种基于累积概率分布的正态变换技术, 将非正态分布数据转为正态域进行研究<sup>[12-14]</sup>。具体原理如下:

对某风速随机过程  $V_t$ , 可采用公式(14)进行变换, 得到新的随机过程  $W_t$ , 由于公式(14)具有单调性, 因此得到  $W_t$  与  $V_t$  具有一一对应性, 且随机过程  $W_t$  的边缘分布函数为正态分布。变换原理图如图(1)所示。

$$W_t = \Phi^{-1}(F_t) \quad (14)$$

式中:  $F_t$  为该随机过程的边缘分布函数;  $\Phi^{-1}$  为标准正态分布函数的逆函数。

由于相邻  $k$  个时刻的随机变量均来自同一个随机过程, 因此可以认为它们的边缘分布均服从正态分布。因而, 它们可以构成的  $k$  维随机向量可近似的认为服从多元正态分布, 即

$$(W_{t-(k-1)}, \dots, W_t, W_{t+1}) \sim N(\mu, \chi) \quad (15)$$

式中:  $\mu = (\mu_{t-(k-1)}, \dots, \mu_t, \mu_{t+1})$  为均值向量;  $\chi$  为协方差矩阵, 表示为

$$\chi = \begin{bmatrix} \sigma_{t-(k-1), t-(k-1)} & \dots & \sigma_{t-(k-1), t} & \sigma_{t-(k-1), t+1} \\ \vdots & \ddots & \vdots & \vdots \\ \sigma_{t, t-(k-1)} & \dots & \sigma_{t, t} & \sigma_{t, t+1} \\ \sigma_{t+1, t-(k-1)} & \dots & \sigma_{t+1, t} & \sigma_{t+1, t+1} \end{bmatrix}$$

式中  $\sigma_{i,j} = E[(v_i - \mu_i)(v_j - \mu_j)]$ , 写成分块形式为

$$\chi = \begin{bmatrix} \chi_{11} & \chi_{12} \\ \chi_{21} & \chi_{22} \end{bmatrix} \begin{matrix} k \times k \\ k \times 1 \\ 1 \times k \\ 1 \times 1 \end{matrix} \quad (16)$$

由概率论知识可知, 若一个随机向量服从多元正态分布, 则该向量间的条件分布概率也服从正态分布, 表示为

$$(W_{t+1} | W_t, \dots, W_{t-(k-1)}) \sim N(\mu, \sigma) \quad (17)$$

式中  $\mu$ 、 $\sigma$  为定常数, 可分别根据均值向量和协方

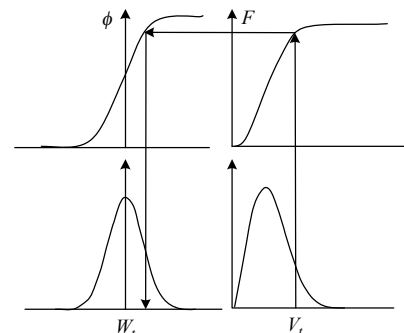


图1 正态变换原理图

Fig. 1 Schematic diagram of normal transformation



差矩阵求得，具体根据式(18)和式(19)求解。

$$\mu = \mu_{t+1} + \chi_{21} \times \chi_{11}^{-1} \times \begin{bmatrix} w_{t-(k-1)} - \mu_{t-(k-1)} \\ \vdots \\ w_t - \mu_t \end{bmatrix} \quad (18)$$

$$\sigma^2 = \chi_{22} - \chi_{21} \times \chi_{11}^{-1} \times \chi_{12} \quad (19)$$

因此随机过程  $W_t$  的  $k \sim$  阶连续马尔科夫模型的转台转移函数为

$$f(w_{t+1} | w_t, \dots, w_{t-(k-1)}) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(w_{t+1}-\mu)^2}{2\sigma^2}} \quad (20)$$

然后经过式(11)进行反变换便可得到原始风速过程  $V_t$  的马尔科夫链模型。

$$V_t = F^{-1}(\Phi(W_t)) \quad (21)$$

### 4 计及自相关性风速预测模型分析的具体步骤

根据上面分析，现将基于随机过程自相关性的风速预测模型的具体过程陈述如下：

1) 对于正态随机过程略过此步骤，对于非正态随机过程先将其转换至正态域。

①首先通过拟合的方法得到原始风速数据序列  $V_t$  的累积概率分布函数  $F(\cdot)$ 。

②利用 CDF 技术对该原始数据  $V_t$  进行变换，得到服从[0,1]分布的均匀序列  $U_t$ 。

③利用公式(14)对原始序列  $U_t$  进行正态变换，得到服从标准正态分布的风速过程序列  $W_t$ 。

2) 根据马尔科夫转移函数识别风速过程间的相关特性<sup>[15-16]</sup>，产生得到一条时间序列。

①根据  $W_t$  各时刻的边缘分布，产生随机向量  $(W_1, \dots, W_k)$  的初始值，不妨记作  $(w'_1, \dots, w'_k)$ 。

②依据公式(15)~(20),采用拒绝抽样法产生下一时刻的状态  $w'_{k+1}$ 。

③将产生的  $w'_{k+1}$  作为已知状态，更新初始状态为  $(w'_2, \dots, w'_{k+1})$ ，依据公式(15)~(20)，抽样产生下一时刻的状态  $w'_{k+2}$ ，并以此类推，得到  $w'_t(t=k+1, k+2, \dots, n)$ ，进行多点预测。

3) 按公式(21)将模拟产生的变换后的数据序列转换为原始风速过程序列，从而预测之后时间的风速大小。

### 5 计及风速自相关性预测模型模型的适应性分析

为了分析本文所提预测模型的适应性，首先对传统离散马尔科夫链模型(discrete state Markov chain model, DSMC)进行分析，研究本文模型与 DSMC 模型在多点预测时自相关性变化规律，从而说明本文所提模型在各种风速情况下都可以保持

原始数据的自相关性，具有良好的适应性。

### 5.1 离散马尔科夫模型自相关性分析

假设  $\{V(t)\}$  为描述某一风速序列的随机过程，其状态空间  $I = \{v_1, v_2, \dots, v_m\}$  存在有限个风速数值，则其  $k \sim$  阶 DSMC 模型可表示为

$$P[V(t+1)=v_j | V(t)=v_s, \dots, V[t-(k-1)]=v_{s-(k-1)}] = P_{v_{s-(k-1)}, v_{s-1}, v_s, v_j} \quad (22)$$

对于一阶 DSMC 模型，式(22)可简化为

$$P_{v_{s-(k-1)}, v_{s-1}, v_s, v_j} = P_{v_s, v_j} = P_{s,j}(1) \quad (23)$$

式中  $P_{s,j}(1)$  为状态  $v_s$  和  $v_j$  间的一步转移概率。

由于状态转移概率只与起始状态和时间间隔有关，因此它们可以构成  $m \times m^k$  维状态转移矩阵<sup>[17]</sup>，对于常用的一阶模型，其状态转移矩阵记为  $P(1)$ ，其每行元素之和为 1，即

$$\sum_{j=1}^m P_{ij}(1) = 1, i=1, 2, \dots, m \quad (24)$$

对于多点预测问题，为了研究相关性变化性质，需要求取其  $n$  步转移概率，先给出 C-K 方程：

$$P_{ij}(u+v) = \sum_{k=1}^{+∞} P_{ik}(u)P_{kj}(v), i, j=1, 2, \dots \quad (25)$$

令  $n = u + v, u=1, v=n-1$ ，可得

$$P(n) = P(1)P(n-1) \quad (26)$$

进而可得

$$P(n) = P^n(1) \quad (27)$$

由式(24)可知

$$\sum_{j=1, j \neq i}^m P_{ij}(1) \leq 1, i=1, 2, \dots, m \quad (28)$$

根据圆盘定理<sup>[18]</sup>可知  $P(1)$  矩阵的特征值均小于 1，故  $P(n)$  随着  $n$  的增大趋于一个常矩阵  $P_{constant}$ ，由于  $P(n) \neq P(1)$ ， $P_{constant} \neq P(1)$ ，故在多点预测时与原始数据自相关性差别较大。且  $P(n) \rightarrow P_{constant}$ ，说明该模型产生随机数据的自相关性会以保持一定的稳定性，但与原始数据不同。

### 5.2 连续马尔科夫模型自相关性分析

当原始数据已知时，根据式(13)或式(21)求得连续马尔科夫模型的核密度函数，由式(12)可知，该核密度函数即是下一时刻风速的概率分布函数，其转移概率为 1，对于  $k \sim$  阶正态 CSMC 模型，其转移过程如图 2 所示。

当由数据序列  $(w'_{t-k+1}, \dots, w'_t)$  得到下一时刻数据  $w'_{t+1}$ ，并由新的数据序列  $(w'_{t-k+2}, \dots, w'_{t+1})$  得到  $w'_{t+2}$ ，

$$\{(\mu_{t-k+1}, \delta_{t-k+1}), (\dots), (\mu_t, \delta_t)\} \xrightarrow{p=1} \{(\mu_{t+1}, \delta_{t+1})\} \xrightarrow{p=1} \{(\mu_{t+2}, \delta_{t+2})\} \xrightarrow{p=1} \{\dots\}$$

图 2  $k \sim$  阶正态 CSMC 模型状态转移图

Fig. 2 State transition diagram of normal CSMC model

以此类推，其核密度转移概率均为 1，并不随着预测时间间隔的延长而发生变化，且正态随机过程相关参数由原始数据统计而来，可以充分反映原始风速数据特性，故连续马尔科夫模型可以很好的保持原始数据的自相关特性。

由以上分析可知，随机过程的参数选择对本文模型多点预测的自相关特性影响很大，并不受预测时间间隔的影响，因此本文模型在多点预测时具有良好的适应性。

## 6 仿真分析

### 6.1 仿真环境及仿真数据预处理

本文选择新西兰水与大气国家研究院国家气象数据库中某一地区风速数据对上述所提模型进行仿真分析，该地区历史风速的统计量如表 1 所示。概率分布特性如图 3 所示。

该图所示概率密度分布并不服从正态分布，因此需对原始风速数据进行正态变换，首先采用核密度<sup>[19-20]</sup>估计的方法求取风速序列的边缘累积概率分布函数：

$$\hat{F}(v) = \int_{-\infty}^v \hat{f}(x) dx \quad (29)$$

当核密度函数采用正态核函数时，

$$\hat{F}(v) = \frac{1}{n} \sum_{i=1}^n \Phi\left(\frac{v-v_i}{h}\right) \quad (30)$$

式中： $n$  为样本容量； $h$  为窗宽，也称为光滑参数。

利用上式可将原始风速数据序列变换为服从 $[0,1]$ 均匀分布的时间序列 $U_i$ ；再根据公式(14)利用正态分布的逆函数 $\Phi^{-1}(\cdot)$ 将 $U_i$ 转换为服从正态分布的随机变量 $W_i$ 。本文采用 K-S 法<sup>[21-22]</sup>对上述变

表 1 CLIDB 地区风速基本统计量表

Tab. 1 Basic statistics of wind speed in CLIDB area

统计量	均值/(m/s)	方差/(m <sup>2</sup> /s)	最大值	最小值
CLIDB	9.51	4.26	29.81	0.11

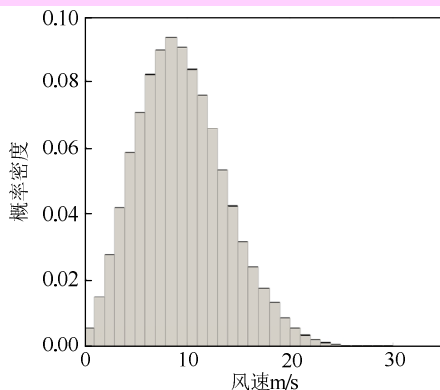


图 3 CLIDB 地区风速概率分布统计图

Fig. 3 Wind speed probability distribution graph in CLIDB area

换进行正态检验，结果如表 2 所示。由于检验统计量值小于临界值，因此在显著性水平  $\alpha = 0.05$  下，该地区经过变换后的风速序列服从标准正态分布，证明了变换的合理性。

表 2 正态变换检验表

Tab. 2 Normal transformation test table

测风点	检验统计量 D	临界值 CV
CLIDB	0.003 8	0.008 37

### 6.2 CMSC 模型概率统计特性仿真验证

采用连续自相关马尔科夫链模型模拟产生了 200 a 的风速序列，并采用 K-S 方法检验模拟生成的数据能否保持原数据的概率分布特性。表 3 给出了显著水平  $\alpha = 0.05$  下的检验结果。

表 3 结果表明，各阶模型的检验统计量值均小于临界值且各阶的检验统计量数值基本相似，变换不大，这说明：

1) 采用自相关性的连续马尔科夫链随机过程模型可以很好的保持原有风速序列的概率分布特性，能够较为准确描述某一地区的风速过程。

2) 风速模拟的精度对 CSMC 模型的阶数变换不敏感，采用任何阶数模型拟合产生的数据结果基本类似，因此可以采用较为简单的一阶模型进行分析计算，可以大大简化计算量。

将模拟时间延长至 500 a 甚至更长时，各阶模型的检验统计量数值更趋于一致，拟合精度对模型阶数的敏感度更低，表明在风速序列的长期预测中，出于简单可行的角度，采用一阶模型更为合理。

为了进一步说明本文所提模型的概率分布特性，给出 CLIDB 地区模拟风速基本统计量表，如表 4 所示，以及 CLIDB 地区模拟风速概率分布统计图，如图 4 所示。

由表 4 和图 4 可知：

1) 本文所提各阶预测模型都可以较好的保持原始数据的概率分布特性，二阶预测模型各项统计

表 3 模拟风速检验统计表

Tab. 3 Normal test table of simulating wind speed

模型阶数	检验统计量 D	CV
一阶	0.001 99	
二阶	0.001 96	0.008 45
三阶	0.001 97	

表 4 CLIDB 地区模拟风速基本统计量表

Tab. 4 Basic statistics of simulated wind speed in CLIDB area

统计量	均值/(m/s)	方差/(m <sup>2</sup> /s)	最大值/(m/s)	最小值/(m/s)
一阶模型	9.44	4.32	31.23	0.97
二阶模型	9.50	4.28	30.07	0.14
三阶模型	9.38	4.66	27.31	0.09
威尔布模型	11.57	6.72	42.30	6.41
DSMC 法	8.91	5.34	36.51	3.44

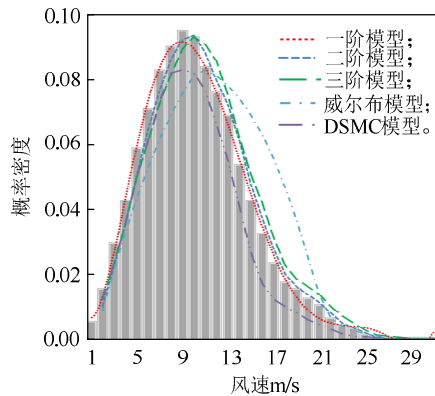


图4 模拟风速数据概率分布统计对比图

Fig. 4 Comparison graph of probability distribution of simulated wind speed data

特性与原始数据最为接近，总体相差不大。

2) 本文所提预测模型各项统计指标相对于传统模型更接近实际数据，具有较大的准确性优势。

### 6.3 CMSC 模型自相关特性仿真验证

为了对比分析模拟风速与原始风速间自相关特性，给出该地区原始风速与模拟风速的自相关曲线图，如图5所示，以及不同模型预测结果自相关性对比图，如图6所示。

图5表明，对于该区域的风速序列，二阶模型与其他阶数模型相比能够在较长时间内保持数据的自相关特性，这是因为该地区相邻时刻的风速数据具有近似正态的相关结构，这一点可从相邻时刻数据(经过正态变换后)构成的二维向量 $(W_t, W_{t+1})$ 散

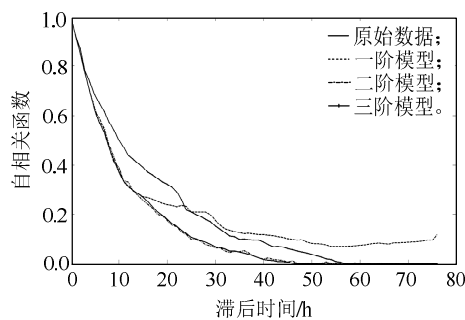


图5 模型阶数对自相关性分析图

Fig. 5 Self-correlation analysis of different model order

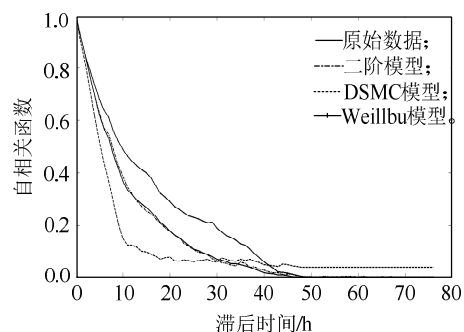


图6 不同模型预测结果自相关性对比图

Fig. 6 Comparison graph of self-correlation of different model prediction results

点图7看出，对于这种结构的数据采用正态状态转移核可以使模拟数据很好的保持原有数据的自相关特性。因此在对某地区风速进行模拟时，要结合概率分布特性和自相关特性合理选择模型阶数，以便达到更好的模拟效果，由上述的仿真分析可知，由于概率分布特性对模型阶数敏感度较低，以自相关特性为依准进行选择是合理的。图6表明，相对于传统预测模型，本文所提模型预测结果的自相关特性与原始数据更为接近，预测结果也更为准确。

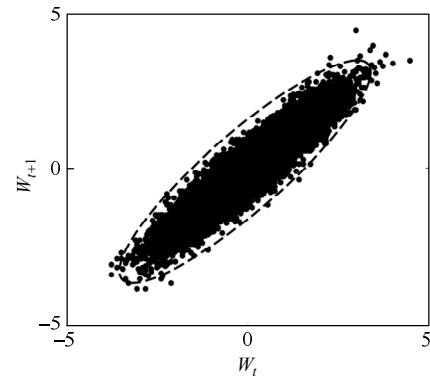


图7 相邻时刻风速数据构成二维向量散点图

Fig. 7 Two-dimensional vector scatter plot of the adjacent wind velocity data

## 7 结论

本文针对传统模型预测风速精度不准的缺陷，提出了基于随机过程的风速预测模型。以随机过程模型代替传统概率分布模型，以连续马尔科夫分模型代替离散马尔科夫链模型，以状态转移函数表征随机过程的自相关性质，同时兼顾了风速的概率分布特性和时序上的自相关特性。经过理论和仿真，得到如下结论：

1) 该模型可以很好地保持原始风速数据的概率分布特性，且模型阶数对模拟结果统计参数影响不明显，尤其在长时间预测时，该结论更加明显。

2) 该模型可以很好地保持原始数据的自相关特性，但模型阶数对相关特性的影响较大，应以自相关特性为主选择模型阶数进行预测。

3) 相对传统离散马尔科夫模型，本文所提连续马尔科夫模型的自相关特性不会随着预测时间间隔的增加而偏离原始数据，更适用于长时间多点预测。

由于本文所提模型将随机过程中每一时刻均服从同一分布，而实际风速可能存在不同的情况，针对这种情况将做进一步研究。

## 参考文献

- [1] Ucar A, Balo F. Evaluation of wind energy potential and electricity

- generation at six locations in Turkey[J]. Applied Energy, 2013, 86(10): 1864-1872.
- [2] Tian P C. Estimation of wind energy potential using different probability density functions[J]. Applied Energy, 2011, 88(5): 1848-1856.
- [3] Xie K, Billinton R. Energy and reliability benefits of wind energy conversion systems[J]. Renewable Energy, 2011, 36(7): 1983-1988.
- [4] 姚国平, 余岳峰, 王志征. 如东沿海地区风速数据分析及风力发电量计算[J]. 电力自动化设备, 2004, 24(4): 12-14.  
Yao Guoping, Yu Yuefeng, Wang Zhizheng. Wind speed data analysis and calculation of wind power generation in coastal areas of Rudong [J]. Electric Power Automation Equipment, 2004, 24 (4): 12-14(in Chinese).
- [5] Jaramillo O A, Borja M A. Wind speed analysis in La Ventosa, Mexico: a bimodal probability distribution case[J]. Renewable Energy, 2014, 29(10): 1613-1630.
- [6] Sahin A D, Sen Z. First order Markov chain approach to wind speed modeling[J]. Journal of Wind Engineering & Industrial Aerodynamics, 2011, 89(3-4): 263-269.
- [7] 邓威, 李欣然, 徐振华, 等. 考虑风速相关性的概率潮流计算及影响分析[J]. 电网技术, 2012, 36(4): 45-50.  
Deng Wei, Li Xinran, Xu Zhenhua, et al. Probabilistic load flow calculation considering wind speed correlation and its influence analysis[J]. Power System Technology, 2012, 36(4): 45-50(in Chinese).
- [8] 范玉妹. 概率论与数理统计[M]. 北京: 机械工业出版社, 2012.
- [9] 卿湘运, 杨富文, 王行愚. 采用贝叶斯-克里金-卡尔曼模型的多风电场风速短期预测[J]. 中国电机工程学报, 2012, 32(35): 107-114.  
Xiang Yun Qing, Yang Fuwen, Wang Hangyu. Using Bayesian - Kriging-Kalman model of wind speed short-term prediction[J]. Proceedings of the CSEE, 2012, 32(35): 107-114(in Chinese).
- [10] 郭创新, 王扬, 沈勇, 等. 风电场短期风速的多变量局域预测法[J]. 中国电机工程学报, 2012, 32(1): 24-31.  
Guo Chuangxin, Wang Yang, Shen Yong, et al. Multi variable local forecasting method for short term wind speed of wind farm[J]. Chinese Journal of Electrical Engineering, 2012, 32(1): 24-31(in Chinese).
- [11] 李玉敦. 计及相关性的风速模型及其在发电系统可靠性评估中的应用[D]. 重庆: 重庆大学, 2012.
- [12] 曲翀, 王秀丽, 谢绍宇, 等. 不同风速模型和可靠性指标对风电可信容量评估的影响[J]. 电网技术, 2013, 37(10): 2896-2903.  
Qu Chong, Wang Xiuli, Xie Shaoyu. Different wind speed model and the reliability index of wind power credible capacity evaluation[J]. Power System Technology, 2013, 37(10): 2896-2903(in Chinese).
- [13] 吴俊利, 张步涵, 王魁. 基于 Ada Boost 的 BP 神经网络改进算法在短期风速预测中的应用[J]. 电网技术, 2012, 36(9): 221-225.  
Wu Junli, Zhang Buhuan, Wang Kui. BP neural network based on Ada Boost algorithm in short term wind speed forecasting application[J]. Power System Technology, 2012, 36 (9): 221-225(in Chinese).
- [14] 王松岩, 李碧君, 于继来, 等. 风速与风电功率预测误差概率分布的时变特性分析[J]. 电网技术, 2013, 37(4): 967-973.  
Wang Songyan, Li Bijun, Yu Jilai, et al. Analysis of time-varying characteristics of wind speed and wind power prediction error probability distribution[J]. Power System Technology, 2013, 37(4): 967-973(in Chinese).
- [15] 高爽, 冬雷, 高阳, 等. 基于粗糙集理论的中长期风速预测[J]. 中国电机工程学报, 2012, 32(1): 32-37.  
Gao Shuang, Dong Lei, Gao Yang, et al. Based on rough set theory in long term wind speed prediction[J]. The Chinese Journal of Electrical Engineering, 2012, 32 (1): 32-37(in Chinese).
- [16] 孙斌, 姚海涛, 刘婷. 基于高斯过程回归的短期风速预测[J]. 中国电机工程学报, 2012, 32(29): 104-109.  
Sun Bin, Yao Haitao, Liu Ting. Short term wind speed forecasting based on Gauss process regression[J]. Chinese Journal of Electrical Engineering, 2012, 32(29): 104-109(in Chinese).
- [17] 甘特马赫尔. 矩阵论[M]. 哈尔滨: 哈尔滨工业大学出版社, 2013.
- [18] 曾杰, 张华. 基于最小二乘支持向量机的风速预测模型[J]. 电网技术, 2009, 33(18): 144-147.  
Zeng Jie, Zhang Hua. Wind speed forecasting model based on least square support vector machine[J]. Power System Technology, 2009, 33 (18): 144-147(in Chinese).
- [19] 肖迁, 李文华, 李志刚, 等. 基于改进的小波-BP 神经网络的风速和风电功率预测[J]. 电力系统保护与控制, 2014, 42(15): 80-86.  
Xiao Qian, Li Wenhua, Li Zhigang, et al. Wind speed and wind power prediction based on Improved Wavelet -BP neural network[J]. Power System Protection and Control, 2014, 42(15): 80-86(in Chinese).
- [20] 王辉, 刘达, 王继龙. 基于谱聚类和优化极端学习机的超短期风速预测[J]. 电网技术, 2015, 39(5): 1307-1314.  
Wang Hui, Liu Da, Wang Jilong. Super short term wind speed forecasting based on spectral clustering and optimization of extreme learning machine[J]. Power System Technology, 2015, 39(5): 1307-1314(in Chinese).
- [21] 王松岩, 于继来. 风速与风电功率的联合条件概率预测方法[J]. 中国电机工程学报, 2011, 31(7): 7-15.  
Wang Songyan, Yu Jilai. Joint conditional probability of wind speed and wind power prediction method[J]. Proceedings of the CSEE, 2011, 31(7): 7-15(in Chinese).
- [22] 丁明, 张立军, 吴义纯. 基于时间序列分析的风电场风速预测模型[J]. 电力自动化设备, 2005, 25(8): 32-34.  
Ding Ming, Zhang Lijun, Wu Yichun. Wind speed forecasting model for wind farm based on time series analysis[J]. Electric Power Automation Equipment, 2005, 25(8): 32-34(in Chinese).



史可琴

收稿日期: 2016-04-29.

作者简介:

史可琴(1963), 女, 硕士, 高级工程师, 从事电网运行与控制工作;

王方雨(1992), 男, 硕士研究生, 研究方向为电力系统分析与控制, E-mail: 18001146577@163.com.

(责任编辑 王金芝)