

短期风速时间序列混沌特性分析及预测

田中大 李树江 王艳红 高宪文

Chaotic characteristics analysis and prediction for short-term wind speed time series

Tian Zhong-Da Li Shu-Jiang Wang Yan-Hong Gao Xian-Wen

引用信息 Citation: *Acta Physica Sinica*, 64, 030506 (2015) DOI: 10.7498/aps.64.030506

在线阅读 View online: <http://dx.doi.org/10.7498/aps.64.030506>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn/CN/Y2015/V64/I3>

您可能感兴趣的其他文章

Articles you may be interested in

交通流突变点的无标度特征分析

[Analysis of scale-free characteristic on sharp variation point of traffic flow](#)

物理学报.2014, 63(24): 240509 <http://dx.doi.org/10.7498/aps.63.240509>

基于压缩感知的振动数据修复方法

[Vibration data recovery based on compressed sensing](#)

物理学报.2014, 63(20): 200506 <http://dx.doi.org/10.7498/aps.63.200506>

用于混沌时间序列预测的组合核函数最小二乘支持向量机

[Combination kernel function least squares support vector machine for chaotic time series prediction](#)

物理学报.2014, 63(16): 160508 <http://dx.doi.org/10.7498/aps.63.160508>

基于能力区域的交通状态预测方法

[Traffic state prediction based on competence region](#)

物理学报.2014, 63(14): 140504 <http://dx.doi.org/10.7498/aps.63.140504>

一种基于相关分析的局域最小二乘支持向量机小尺度网络流量预测算法

[A local least square support vector machine prediction algorithm of small scale network traffic based on correlation analysis](#)

物理学报.2014, 63(13): 130504 <http://dx.doi.org/10.7498/aps.63.130504>

短期风速时间序列混沌特性分析及预测*

田中大^{1)†} 李树江¹⁾ 王艳红¹⁾ 高宪文²⁾

1)(沈阳工业大学信息科学与工程学院, 沈阳 110870)

2)(东北大学信息科学与工程学院, 沈阳 110819)

(2014年6月30日收到; 2014年8月15日收到修改稿)

针对短期风速时间序列的预测问题进行了研究. 首先通过0-1混沌测试法确定短期风速时间序列具有混沌特性. 采用相空间重构技术, 利用C-C算法确定延迟时间, G-P算法确定嵌入维数. 然后提出一种参数在线修正的最小二乘支持向量机预测模型, 采用改进的粒子群算法进行预测模型中参数的优化. 最后通过仿真对比实验表明提出的预测方法在预测精度、预测误差、预测效果方面都要优于其他常见的预测方法, 证明该预测方法是有效的.

关键词: 短期风速, 时间序列, 混沌, 预测

PACS: 05.45.Tp

DOI: 10.7498/aps.64.030506

1 引言

随着化石燃料的日益枯竭, 以及环境污染的日益严重, 风能作为一种无污染、可再生能源, 得到世界各国的高度重视, 风力发电成为世界各国重点发展的可再生能源发电技术之一^[1]. 但是研究发现风速的随机性将导致风电并网后对电力系统的稳定性产生严重影响^[2]. 目前解决这一问题的方法有两种: 一是增加风电装机容量与常规机组的旋转备用容量, 来抑制风电并网对电网带来的冲击, 但这会增加系统的运行费用; 二是对风速或风电功率做出比较准确的预测, 这样可以及时对电网进行调度, 以提高风力发电的经济性. 风速预测分为长期、中期与短期预测, 其中以“小时”为预测单位的短期风速预测一般是提前1到48小时对每小时的风速进行预测, 有效的短期风速预测对风力发电行业具有重要的应用价值^[3,4].

目前, 风速预测方法可分为基于物理模型的方法和基于历史数据的预测方法两大类. 基于物理模型的预测方法一般采用数值天气预报 (numerical

weather prediction, NWP) 数据进行风速预测^[5,6], 但是我国很多情况下, 数值天气预报模型无法获得, 而且目前很少有专门服务于风电场的数值气象预报模型可以利用, 因此, 第二类的基于历史数据的风速预测研究具有广阔的应用前景.

国内外学者对于短期风速预测问题进行了广泛的研究, 其中包括自回归滑动平均 (auto regressive moving average, ARMA) 模型^[7]、差分自回归滑动平均 (auto regressive integrated moving average, ARIMA) 模型^[8,9] 等线性模型, 以及包括支持向量机 (support vector machine, SVM)^[10,11]、最小二乘支持向量机 (least square support vector machine, LSSVM)^[12,13]、卡尔曼滤波^[14]、人工神经网络^[15,16] 等非线性预测模型. 虽然这些预测模型都取得了广泛应用, 但是这些预测模型都存在着自身的局限性. 文献^[17, 18] 的研究表明短期风速是非线性、非平稳的时间序列, 而基于自回归的 ARMA、ARIMA 等模型虽然具有实现简单、运算量小等优点, 但是只适合平稳线性或者简单非平稳线性时间序列的预测^[19], 对于非平稳且具有较强非线性特

* 国家自然科学基金重点项目 (批准号: 61034005) 资助的课题.

† 通信作者. E-mail: tianzhongda@126.com

征的短期风速时间序列难以表达其复杂的演化特征, 只能取得有限的预测精度. 神经网络具有较好的非线性和自学习能力, 但是神经网络存在易振荡、收敛速度慢、易陷于局部最优值, 同时其包括隐含层数目等在内的网络结构难以确定 [20]; SVM 与 LSSVM 具有很好的稀疏性和泛化能力, 能很好的解决小样本、非线性、高维数和局部极小点等实际问题, 其性能显著优于传统的神经网络算法 [21], 但是其关键参数很难确定, 主要依靠人工选择 [22]; 卡尔曼滤波器可以处理多维、时变和非平稳随机过程, 但卡尔曼滤波器的缺点在于其要求精确已知系统的数学模型和噪声统计特性, 而包括短期风速预测在内的大部分实际应用问题中系统模型和噪声统计特性都是部分已知、近似已知或完全未知的, 此时卡尔曼滤波器的性能就会变坏甚至发散 [23].

本文针对短期风速的预测问题进行了相关研究. 首先对实际采集的短期风速时间序列进行混沌特性分析, 通过 0-1 混沌测试法确定短期风速时间序列具有混沌特性. 针对短期风速时间序列的混沌特性结合相空间重构技术展开预测, 利用 C-C 算法确定序列延迟时间, 利用 G-P 算法确定序列嵌入维数. 在上述基础上, 提出一种参数在线修正的 LSSVM 预测模型, 同时针对预测模型中的参数优化问题而提出一种改进的粒子群算法 (improved particle swarm optimization, IPSO), 与其他几种常见预测方法进行了仿真对比, 结果表明本文的方法具有更高的预测精度与更小的预测误差.

2 短期风速时间序列混沌特性判定

混沌现象是非线性动力系统中常见的现象. 混沌系统是一个确定系统, 它对系统初值具有较强的敏感性. 对于混沌系统的预测而言, 由于其存在初值敏感性, 因此不能长期预测, 但是同时由于其确定性, 混沌系统具有短期预测能力. 因此对于时间序列的预测问题, 首先需要确定其是否具有混沌特性.

本文利用 0-1 混沌测试方法 [24] 进行短期风速时间序列混沌特性的判定, 设短期风速时间序列为 $w(i)$, $i = 1, 2, \dots, n$, 选择一取值范围为 $(0, 2\pi)$ 的随机常数 c , 定义如下 2 个函数:

$$p(n) = \sum_{i=1}^n w(i) \cos(\theta(i)), \quad n = 1, 2, \dots, L, \quad (1)$$

$$q(n) = \sum_{i=1}^n w(i) \sin(\theta(i)), \quad n = 1, 2, \dots, L, \quad (2)$$

上面两式中 $\theta(i)$ 定义为

$$\theta(i) = ic + \sum_{k=1}^i w(k), \quad i = 1, 2, \dots, n. \quad (3)$$

基于函数 $p(n)$ 与 $q(n)$, 定义均方位移 $M(n)$ 为

$$M(n) = M_c(n) - (E(\varphi))^2 \frac{1 - \cos nc}{1 - \cos c}, \quad (4)$$

上式中,

$$M_c(n) = \lim_{N \rightarrow \infty} \sum_{j=1}^L [(p(j+n) - p(j))^2 - (q(j+n) - q(j))^2], \quad (5)$$

$$E(\varphi) = \lim_{L \rightarrow \infty} \frac{1}{L} \sum_{j=1}^L w(j). \quad (6)$$

如果函数 $p(n)$ 与 $q(n)$ 的轨迹为布朗运动, 则上面的函数 $M(n)$ 随着时间呈现线性增长趋势, 如果函数 $p(n)$ 与 $q(n)$ 的轨迹有界, 则函数 $M(n)$ 也是有界. 定义如下的增长率 K_c :

$$K_c = \lim_{n \rightarrow \infty} \lg M(n) / \lg n. \quad (7)$$

当 $K_c \approx 1$ 时则表明时间序列均有混沌特性, 如果 $K_c \approx 0$, 则表明时间序列不具有混沌特性 [25]. 采集东北某电厂 12 月份测得的 300 组采样间隔 1 h 的风速数据, 如图 1 所示.

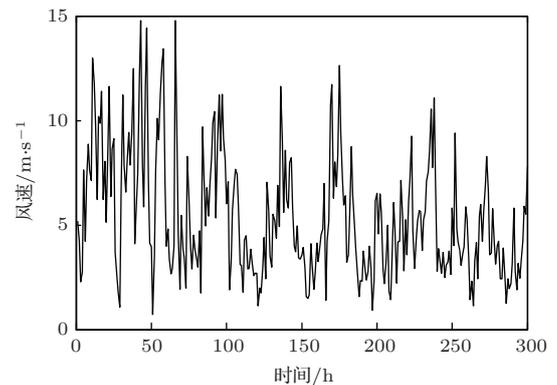


图 1 短期风速时间序列

利用 0-1 混沌测试法对 300 组短期风速时间序列进行分析, 得到图 2 所示的 $p(n)$ 与 $q(n)$ 的相图, 图 3 所示的均方位移 $M(n)$ 的时间变化图, 图 4 所示的渐进增长率 K_c 的变化图.

从图 2 到图 4 可看出, $p(n)$ 与 $q(n)$ 的相图呈现布朗特性, 而均方位移 $M(n)$ 随时间线性增长, 计

算渐进增长率 K_c 为 0.9978, 其趋近于 1, 因此可证明短期风速时间序列具有混沌特性.

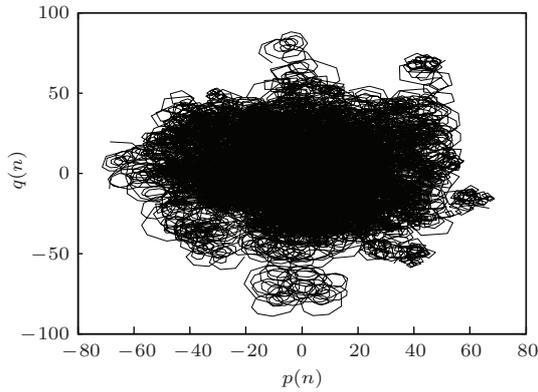


图2 $p(n)$ 与 $q(n)$ 的相图

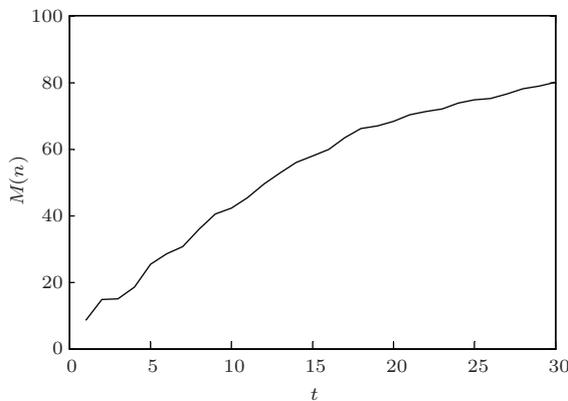


图3 均方位移 $M(n)$ 变化图

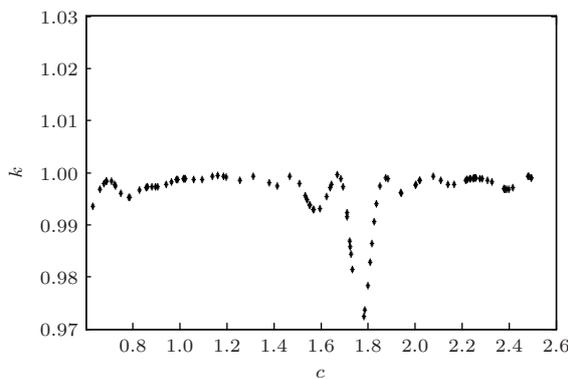


图4 渐进增长率 K_c 变化图

3 短期风速时间序列相空间重构与可预测尺度

在确定短期风速时间序列具有混沌特性之后, 就可以通过相空间重构技术来进行预测. 相空间重构是混沌时间序列预测的基础, 相空间重构用原始

系统中某个变量的延迟坐标来重构相空间, Takens 从数学上为其奠定了可靠的基础. 他的基本观点是: 相空间重构法虽然是用一个变量在不同时刻的值构成相空间, 但动力系统的变量的变化自然跟此变量与系统的其他变量的相互作用有关, 即此变量随时间的变化隐含着整个系统的动力学规律. 因此, 重构的相空间的轨迹也反映系统状态的演化规律.

设短期风速时间序列为 $w(1), w(2), \dots, w(n)$, 由此序列嵌入 m 维相空间, 得到一系列 m 维相空间的相点为

$$\begin{aligned} W_1 &= [w(1), w(1 + \tau), \dots, w(1 + (m - 1)\tau)], \\ W_2 &= [w(2), w(2 + \tau), \dots, w(2 + (m - 1)\tau)], \\ &\vdots \\ W_N &= [w(N), w(N + \tau), \dots, \\ &\quad w(N + (m - 1)\tau)], \end{aligned} \quad (8)$$

上式中, τ 为延迟时间, $N = n - (m - 1)\tau$ 表示由时间序列的 n 个点构成的 m 维相点的个数.

在相空间重构中, 延迟时间 τ 与嵌入维数 m 的选取十分重要, 目前很多学者提出了很多的方法. 本文采用 C-C 方法^[26] 进行 τ 的选取, 该方法利用统计结果获得, 因此运算量少, 实现简单. C-C 方法计算过程如下^[27]:

取 $m = 2, 3, 4, 5, r_i = i \times 0.5\varepsilon, i = 1, 2, 3, 4, \varepsilon$ 为时间序列标准差, 计算下列各式:

$$\bar{S}(t) = \frac{1}{16} \sum_{m=2}^5 \sum_{i=1}^4 S(m, r_i, t), \quad (9)$$

$$\Delta \bar{S}(t) = \frac{1}{4} \sum_{m=2}^5 \Delta S(m, t), \quad (10)$$

$$S_{\text{cor}}(t) = \Delta \bar{S}(t) + |\bar{S}(t)|, \quad (11)$$

其中的 $\bar{S}(t)$ 与 $\Delta \bar{S}(t)$ 均可反映短期风速序列的自相关性, 寻找 $\bar{S}(t)$ 的第一次过零点或者 $\Delta \bar{S}(t)$ 的第一个局部极小值所对应的 t 即为最佳时延 τ .

对于嵌入维数 m 的选取利用 Grassberger 和 Procaccia 提出的 G-P 算法^[28], 该方法依据吸引子关联指数 D_m 随着 m 的增加逐渐达到饱和的原理求取 m . 对于以计算得到的 τ 选择较小的 m 进行相空间重构, 对不同的领域半径 r , 计算相应的关联

积分

$$C_n(r) = \frac{1}{M^2} \sum_{i,j=1}^M \theta[r - |X(i) - X(j)|], \quad (12)$$

式中, $\theta(\cdot)$ 为 Heaviside 单位函数:

$$\theta(x) = \begin{cases} 0, & x \leq 0, \\ 1, & x > 0. \end{cases} \quad (13)$$

对于 r 的恰当范围, D_m 与 $C_n(r)$ 存在下列关系: 让 m 由小增大, 作出不同的 m 值对应的双对数关系曲线 $\ln C_n(r) - \ln r$, 曲线中直线段部分斜率即为关联指数 D_m . 如果随着 m 增大 D_m 出现饱和状态, 该饱和值即为关联维数 D , 根据 $m \geq 2D + 1$ 原则即可确定嵌入维数 m .

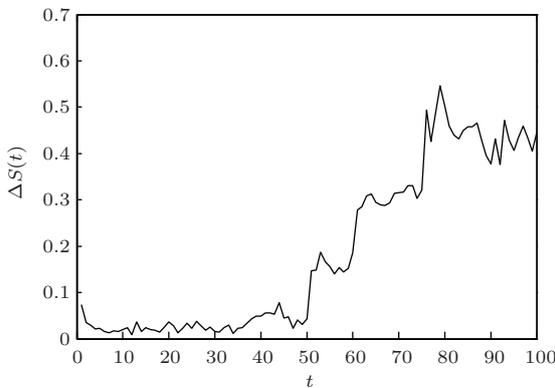


图5 短期风速时间序列 $\Delta S(t)$ 的变化曲线

利用 C-C 算法得到短期风速时间序列 $\Delta S(t)$ 的变化曲线如图 5 所示. 从图 5 中可看出 $\Delta S(t)$ 的第一个局部最小值发生在 $t = 4$, 因此确定短期风速时间序列的时延 τ 为 4. 按照 G-P 算法得到图 6 所示的随着 m 的增大, $\ln C_n(r) - \ln r$ 的变化曲线, 图 7 的 m 与关联维数 D 的曲线. 从图 6 与图 7 可知, 随着嵌入维数 m 的增加, 线性区域内出现了趋于饱和的现象, 继续增加嵌入维数 m 几乎

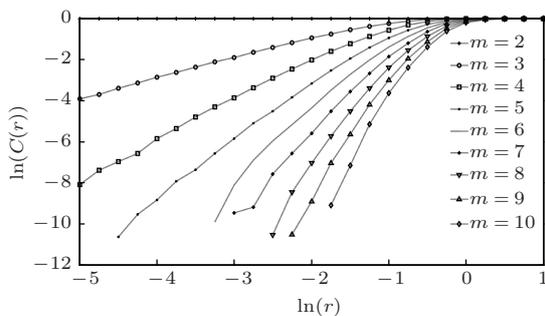


图6 $\ln C_n(r) - \ln r$ 的变化曲线

不影响关联积分的值. 从图 7 中可看出关联维数 D 的收敛值为 5.012 左右, 依据 $m \geq 2D + 1$ 原则, 本文确定短期风速时间序列的嵌入维数 m 为 11.

图 8 与图 9 为延迟时间 τ 为 4 时短期风速时间序列的二维相图与三维相图.

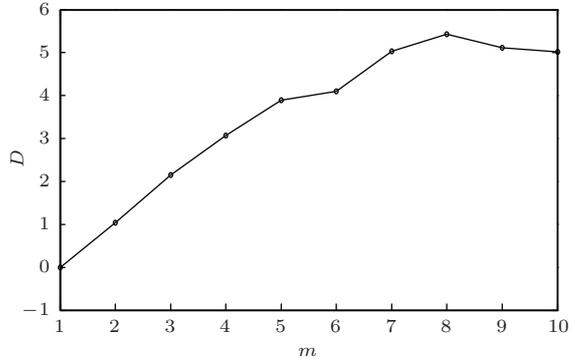


图7 关联维数 D 的变化曲线

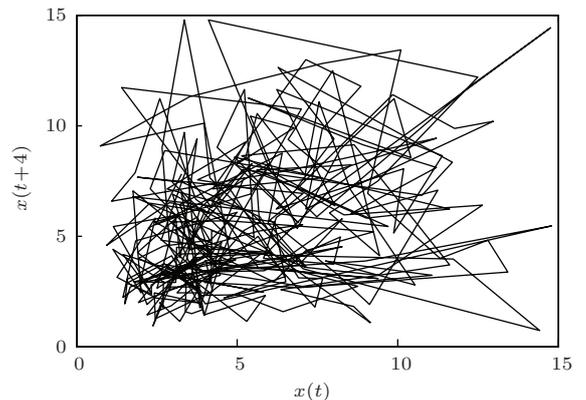


图8 短期风速时间序列的二维相图

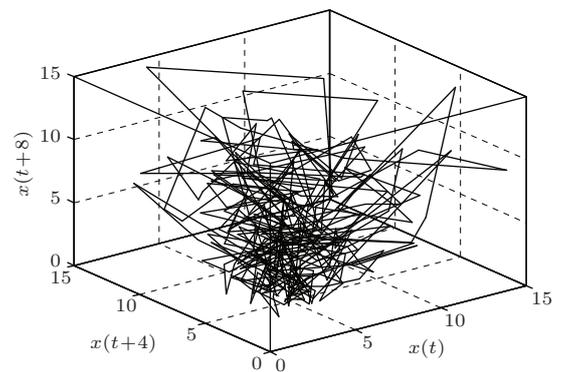


图9 短期风速时间序列的三维相图

对于一个时间序列, 如果延迟时间取值不合理, 混沌吸引子的相轨线被压缩或被折叠, 空间点对的距离分布不均匀. 而从图 8 与图 9 的相图可看

出在 τ 为 4 时, 空间点对点间的距离趋于均匀分布, 吸引子所包含信息量较大, 二维与三维相图表明其对吸引子有序结构的展开相对较为充分, 因此确定延迟时间 τ 为 4 较为合理.

将短期风速时间序列按照得到的嵌入维数与延迟时间进行相空间重构, 计算得到其递归图 (recurrence plot, RP), 递归图计算方法如下:

$$s_{i,j} = \|D_i - D_j\|,$$

$$i = 1, 2, \dots, N - (m - 1)\tau,$$

$$j = 1, 2, \dots, N - (m - 1)\tau, \quad (14)$$

$$R(i, j) = H(\varepsilon_i - s_{i,j}), \quad (15)$$

ε_i 为介质距离, 可取定值, $H(r)$ 为 (13) 式所示的 Heaviside 单位函数, 计算后的递归图如图 10.

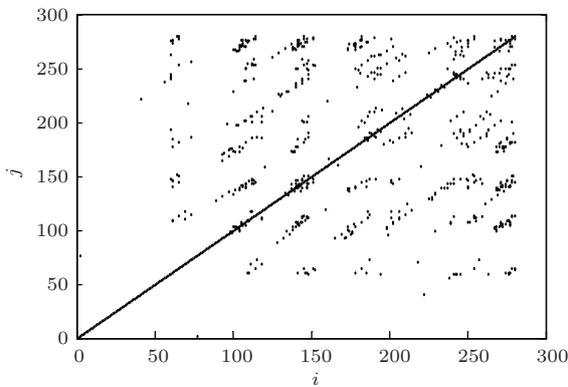


图 10 短期风速时间序列递归图

图 10 表明, 短期风速时间序列不同于随机信号, 其递归图并非均匀分布整个递归平面, 而且存在与主对角线平行的直线段, 因此短期风速时间序列为典型的混沌序列, 在局部范围内具有可预测性.

在确定了相空间重构参数之后, 需要确定短期风速时间序列的可预测尺度, 可通过最大 Lyapunov 指数来得到. 混沌系统的 Lyapunov 指数是反映系统混沌水平的一个变量. 对于 k 维混沌系统, 存在着从大到小的 Lyapunov 指数 $\lambda_1, \lambda_2, \dots, \lambda_k$, 本文采用 Wolf 方法 [29] 计算最大 Lyapunov 指数, 方法如下:

相空间内两个相邻相点 $Y(i)$ 与 $Y(j)$ 的 k 次迭代后, 计算相点 $Y(i+k)$ 与 $Y(j+k)$ 距离分离的平均指数

$$|Y(i+k) - Y(j+k)| = d(0) \cdot e^{k\lambda_1}, \quad (16)$$

$d(0)$ 为其初始距离, 将上式转换为

$$\frac{|Y(i+k) - Y(j+k)|}{d(0)} = e^{\lambda_1 t_0}. \quad (17)$$

假设上式超过临界值 C 时即可认为序列轨道发散无法预测.

$$C = \frac{|Y(i+k) - Y(j+k)|}{d(0)} = e^{\lambda_1 t_0}. \quad (18)$$

因此有 $t_0 = (1/\lambda_1) \ln C$, 取 $C = e$, 则时间序列最大预测步数 $t_0 = 1/\lambda_1$, λ_1 为最大 Lyapunov 指数. 混沌理论研究表明: 对于混沌系统, 在最大预测时间尺度 t_0 范围内时, 系统预测误差随预测步长变化比较平稳, 反之则误差会成倍的增加. 对本文的短期风速时间序列进行最大 Lyapunov 指数的计算, 得到最大 Lyapunov 指数 λ_1 为 0.2452, 由此可知短期风速时间序列的最大可预测时间尺度为 4. 因此本文在下文的预测中采用提前 4 步预测, 也即对未来 4 个采样时刻即 4 小时的风速进行预测, 再大的预测步长将会降低预测精度.

4 短期风速时间序列的 IPSO-LSSVM 预测模型

从本文以上的分析可知, 短期风速时间序列具有混沌特性, 因此对其只能进行短期预测, 同时应该选用具有非线性预测能力的预测模型. LSSVM 用解线性方程组来代替标准 SVM 的二次规划问题, 算法能够解决大规模问题, 需要较少的训练样本, 同时能够提高标准 SVM 的训练速度. 因此本文的短期风速时间序列的预测采用 LSSVM. LSSVM 的建模与预测可详见文献 [30].

4.1 LSSVM 预测模型离线建模

假设采集到得短期风速时间序列为 $w(1), w(2), \dots, w(n)$, 首先对原始短期风速时间序列进行相空间重构. 设当前采样时刻为 i , 对于 LSSVM 单步预测模型输入即为 $w(i+1-m\tau), w(i+1-m\tau+\tau), \dots, w(i+1-\tau)$. 输出为 $w(i+1)$. 按照上述思想, 对于长度为 n 的短期风速时间序列, LSSVM 作未来 4 步预测时的建模输入与输出矩阵分别如下:

$$\mathbf{X} = \begin{bmatrix} w(i+1-m\tau) & w(i+1-(m-1)\tau) & \cdots & w(i+1-\tau) \\ w(i+2-m\tau) & w(i+2-(m-1)\tau) & \cdots & w(i+2-\tau) \\ \vdots & \vdots & \cdots & \vdots \\ w(i+4-m\tau) & w(i+4-(m-1)\tau) & \cdots & w(i+4-\tau) \end{bmatrix}, \quad (19)$$

$$\mathbf{Y} = \begin{bmatrix} w(i+1) \\ w(i+2) \\ w(i+3) \\ w(i+4) \end{bmatrix}. \quad (20)$$

将矩阵 \mathbf{X} 作为输入, 矩阵 \mathbf{Y} 作为输出, 按照标准 LSSVM 算法即可完成预测模型的建立.

4.2 LSSVM 预测模型参数优化

LSSVM 回归模型的预测精度与泛化能力受预测模型中正则化参数 γ 和核函数参数 (本文使用径向基核函数, 对应的参数为 σ^2) 影响, 为了定性说明参数 γ 与 σ^2 对于 LSSVM 短期风速预测模型预测精度的影响, 将采集得到短期风速数据前 250 组作为训练集, 后 50 组作为测试集, 图 11 给出了 γ 与 σ^2 取值范围均为 $[0.1, 10]$, 增长步长为 0.1 情况下, 50 组短期风速实际值与预测值的均方误差 (mean square error, MSE) 的变化曲线图.

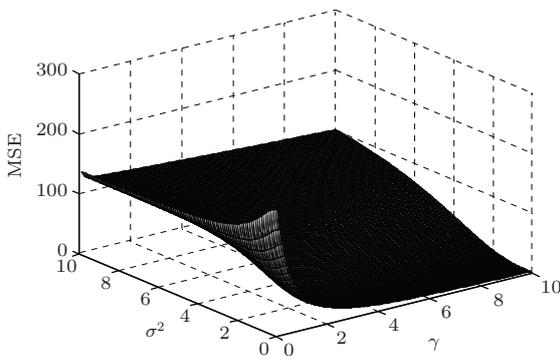


图 11 不同 γ 与 σ^2 时预测误差的均方差

从图 11 可观察到不同的 γ 与 σ^2 组合对 LSSVM 预测性能有着重要的影响, 目前通用的 LSSVM 预测模型参数选择方法是交叉验证法, 即在一定参数范围内, 对每个参数按照一定的间隔取值, 然后进行组合优化. 该方法存在算法耗时长, 同时对参数空间的划分有限, 很难找到最优解. 而粒子群算法 (PSO) 是由 Kennedy 和 Eberhart [31]

提出的一种演化计算技术, 算法概念简单、容易实现, 目前已广泛应用于函数优化 [32], 神经网络训练 [33], 模糊系统控制 [34] 等应用领域. 但是标准 PSO 算法仅仅是对自然生物群集智能的一种简单模拟, 即粒子向种群最优和个体最优两个点的趋近, 而忽略了存在于种群中的大量有用信息, 因此算法在进化过程中容易失去种群多样性, 最终导致早熟收敛. 本文尝试对这些潜在于粒子中的有用信息进行挖掘, 提出一种 IPSO 算法.

1) 精英集的建立与更新

由于粒子在每一维上都向种群最优位置和个体最优位置的方向靠近, 在适应度值整体提高的同时, 会发生某些分维上的退化现象. 由于适应度值整体的提高很难注意到这种退化, 但却影响到进化后期的精度. 考虑到具有较好适应度值的精英粒子往往包含部分优秀的分维信息, 因此将一定规模的精英粒子集结成一个精英集, 并通过该集影响种群内部粒子的寻优来提高优化精度.

精英集是一个非空集合, 元素为种群中适值较高的精英粒子. 在初始化精英集时, 将种群中适应度值最好的 K (K 为精英集规模) 个粒子入选精英集. 更新精英集时, 入选精英集的粒子需要具备两个条件: 一是高适应度值; 二是与精英集中已有的粒子形态差异大.

定义 1 粒子距离.

粒子距离是指两个粒子间各维上的欧式距离和, 即

$$\text{Dis}(x_i, x_j) = \sqrt{\sum_{h=1}^D (x_{i,h} - x_{j,h})^2}. \quad (21)$$

定义 2 粒子与精英集的距离.

粒子与精英集的距离指某一特定粒子与精英集中所有粒子间的距离的平均值.

$$\begin{aligned} & \text{Dis}(x_i, ES) \\ &= \sum_{j=1}^K \text{Dis}(x_i, ES_j) / K \end{aligned}$$

$$= \sum_{j=1}^K \sqrt{\sum_{h=1}^D (x_{i,h} - ES_{j,h})^2 / K}, \quad (22)$$

其中, D 为优化问题维数, K 为精英集规模, 即精英集包含的粒子个数.

定义 3 精英集中心.

精英集中心即精英集中心点位置. 它是由精英集中各粒子的位置平均值计算而得.

$$C_j = \sum_{i=1}^K ES_{i,j} / K. \quad (23)$$

定义精英集中心与集的距离为

$$\begin{aligned} \text{Dis}(C, ES) &= \sum_{i=1}^K \text{Dis}(C, ES_i) / K \\ &= \sum_{i=1}^K \sqrt{\sum_{h=1}^D (C_h - ES_{i,h})^2 / K}. \end{aligned} \quad (24)$$

由上述定义, 若某一粒子的适值大于精英集中适值最差的粒子, 且该粒子与精英集的距离大于精英集中心点与该集的距离, 则该粒子替换原集中适值最差的粒子, 反之丢弃此粒子, 保持精英集不变.

2) 粒子速度更新方式

在通常的 PSO 算法中, 粒子的每一维均向种群最优点与个体最优点两个位置的合力方向直线飞行, 而在本文的 IPSO 算法中, 粒子各维的飞行方向并不一致, 这是为了试图获得更优的分维信息组合. 本文采用的粒子寻优飞行方式有两种, 一种是朝着基本 PSO 算法普遍采用的种群最优点与个体最优点的合力方向飞行; 另一种则是朝着从精英集中随机抽取一个粒子的方向飞行. 即

$$\begin{aligned} v_{t+1} &= w \cdot v_t + c_1 \cdot \text{rand}() \cdot (p_{i,j} - x_{i,j}) \\ &\quad + c_2 \cdot \text{rand}() \cdot (g_j - x_{i,j}), \\ r &> 0.5, \end{aligned} \quad (25)$$

$$\begin{aligned} v_{t+1} &= w \cdot v_t + c_1 \cdot \text{rand}() \cdot (p_{i,j} - x_{i,j}) \\ &\quad + c_2 \cdot \text{rand}() \cdot (ES_{h,j} - x_{i,j}), \\ r &\leq 0.5, \end{aligned} \quad (26)$$

其中, r 为 $[0, 1]$ 内均匀分布的随机数, h 为 $[1, K]$ 内均匀分布的随机整数. 这样有利于粒子各维的独立性和多样性的保持. IPSO 算法中的粒子由于在进化中不以概率 1 向全局最优点靠近, 在种群进化前期会减慢收敛速度, 但是这一代价将获得

进化后期精度的明显提高. IPSO 算法具体流程如图 12 所示.

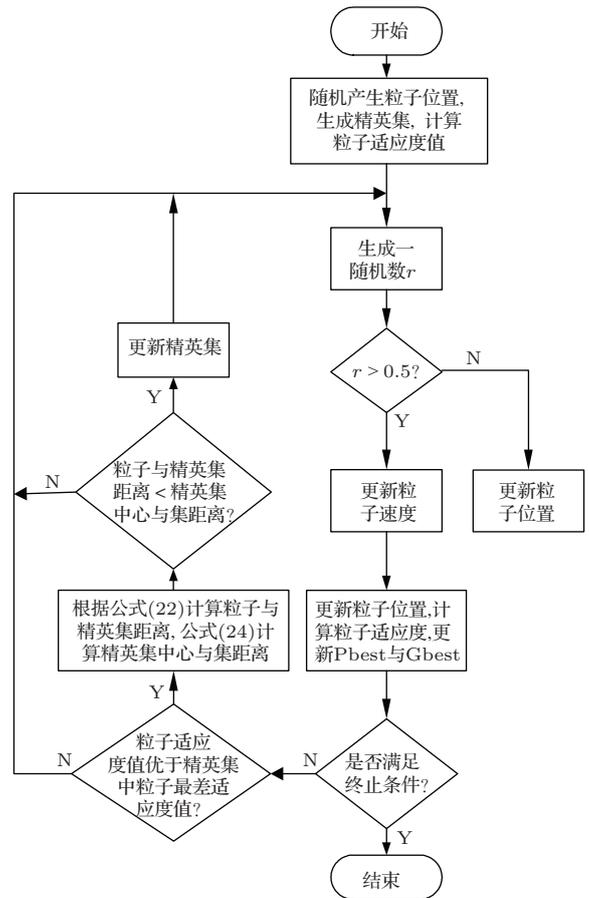


图 12 IPSO 算法流程图

本文将利用 IPSO 算法进行 LSSVM 预测模型中参数 γ 与 σ^2 的优化.

4.3 参数在线修正的 LSSVM 预测模型

LSSVM 离线建模时, 往往从测量数据中选出具有代表性的数据, 建立系统输入与输出之间的映射关系, 模型建立后一般不再变化. 但将这种模型应用于时变的系统时, 存在记录数据在大多数情况下并不能覆盖整个数据的论域. 因此, 模型应随系统的新的采样数据的变化不断地更新, 以确保模型的准确性. 本文采用建立一大大小为 N 的建模数据集与长度为 m 的预测数据集, 当获得新的风速采样值后将其放入建模数据集中, 同时删除最旧的数据, 按照 4.1 节生成预测数据集. 同时设置一预测误差累积阈值, 当预测模型预测误差的累计值大于该阈值时, 则进行模型的重新建模训练, 否则模型不变. 其流程可描述为图 13.

4.4 短期风速的IPSO-LSSVE 预测

结合本文上面的叙述, 给出短期风速的IPSO-LSSVM 预测方法步骤.

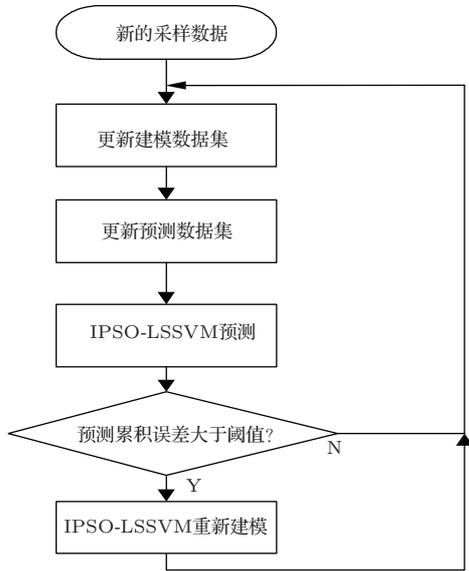


图 13 参数在线修正的 LSSVM 预测流程图

建模过程:

步骤 1 采集短期风速数据, 生成建模数据集.

步骤 2 对建模数据集进行相空间重构处理, 利用 C-C 算法计算短期风速时间序列的延迟时间 τ , 利用 G-P 算法计算嵌入维数 m .

步骤 3 相空间重构短期风速时间序列, 生成 (19) 式的输入矩阵, (20) 式的输出矩阵, 按照 LSSVM 算法进行建模.

步骤 4 利用 IPSO 算法对 LSSVM 预测模型参数进行优化, 输出初始最佳的 γ 与 σ^2 .

在线预测过程:

步骤 1 采集当前采样时刻实际风速值, 按照滑动窗口方式更新建模数据集与预测数据集.

步骤 2 将样本数据按照 (19) 式生成输入数据, 利用优化后的 LSSVM 预测模型预测未来 4 个采样时刻的风速.

步骤 3 计算预测误差累计值, 进行阈值判断, 如果大于阈值则转到步骤 4, 否则转入步骤 1.

步骤 4 相空间重构短期风速时间序列, 生成 (19) 式的输入矩阵, (20) 式的输出矩阵, 按照 LSSVM 算法进行建模.

步骤 5 利用 IPSO 算法对 LSSVM 预测模型参数进行优化, 输出当前最佳的 γ 与 σ^2 , 返回

步骤 1.

5 仿真与实验分析

利用第 2 节中的东北某电厂 12 月份测得的 300 组采样间隔 1 h 的短期风速数据, 前 250 组数据作为训练集用于建模, 后 50 组数据作为测试集用于建模精度的验证.

按照本文 4.4 节中的描述对 50 组短期风速测试样本进行了预测效果的对比, 首先完成 LSSVM 的离线建模. 为了对比 IPSO 算法的优越性, 给出 PSO 与 IPSO 的适应度曲线, 适应度函数取为风速实际值与预测值的均方误差, 即

$$\text{fitness}(x) = \frac{1}{T} \sum_{j=1}^T (w(j) - \hat{w}(j))^2, \quad (27)$$

式中, x 为迭代次数, $w(j)$ 为风速实际值, $\hat{w}(j)$ 为风速预测值. 其中 PSO 与 IPSO 算法的基本参数为: 粒子维数为 2, 参数 γ 取值范围 [0.01, 1000], σ^2 取值范围 [0.01, 1000], 学习因子 c_1 为 1.5, c_2 为 1.7, 最大速度 V_{\max} 为 2, 种群最大数量为 20, 迭代次数为 100, 权重 w 为 0.9. 图 14 为适应度变化对比曲线, 从图 14 中可看出, 本文的 IPSO 算法在收敛速度与收敛精度上都要优于标准的 PSO 算法, 对于短期风速时间序列优化具有更小的适应度值. PSO 优化后 LSSVM 预测模型最佳参数为 $\gamma = 13.24$, $\sigma^2 = 71.25$, IPSO 优化后 LSSVM 预测模型最佳参数为 $\gamma = 3.85$, $\sigma^2 = 265.31$.

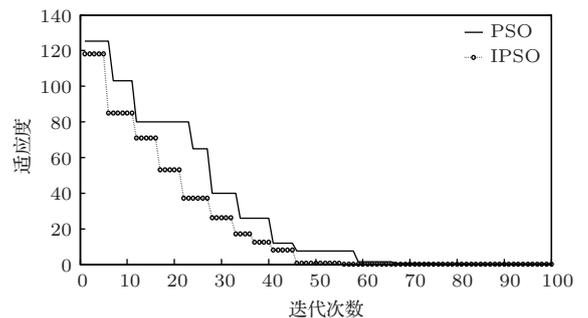


图 14 适应度对比曲线

利用最佳优化参数, 对 50 组测试集风速数据进行了验证. 图 15 为风速预测值与实际值的对比曲线, 仿真中预测误差累积阈值为 10, 在仿真过程中发生了 1 次 LSSVM 的重新建模, 新的模型参数为 $\gamma = 15.14$, $\sigma^2 = 78.65$.

为了衡量本文方法预测精度的提高,与 PSO 优化的 LSSVM(优化参数 $\gamma = 13.24, \sigma^2 = 71.25$)、文献 [8] 中的 ARIMA(根据 AIC 准则确定参数为 $p = 2, d = 1, q = 1$, 对应的 AIC 为 1.638)、文献 [11] 中的 SVM(利用网格法交叉验证得到参数为 $c = 12.32, g = 22.05$)、文献 [12] 中的 LSSVM(利用网格法交叉验证得到参数为 $\gamma = 0.91, \sigma^2 = 23.32$)、文献 [16] 中的 Elman 神经网络(输入层为 20, 中间层为 30, 输出层为 4, 最大迭代次数为 5000) 进行了预测效果与预测精度的对比. 图 16 到图 20 分别给出 50 组测试样本的 PSO-LSSVM、ARIMA、SVM、LSSVM、Elman 神经网络的风速预测值与实际值的对比曲线.

图 21 为几种预测方法预测误差的分布图, 从该图中可观察到本文的方法的预测误差更小、预测误差分布更加平均.

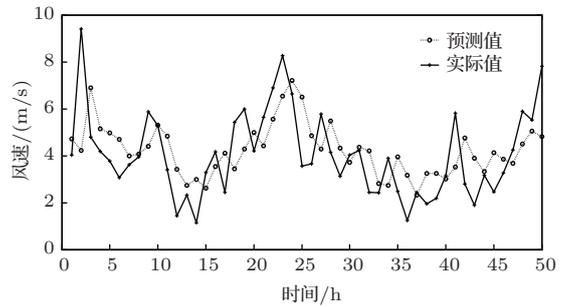


图 18 SVM 风速预测值与实际值对比曲线

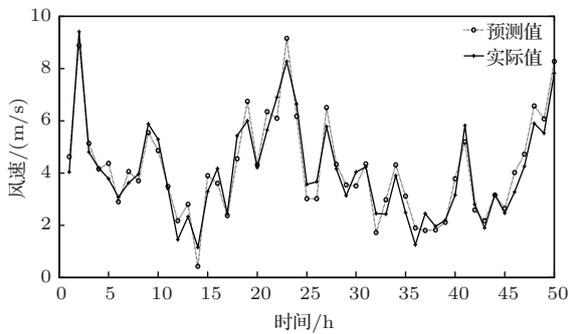


图 15 本文方法风速预测值与实际值对比曲线

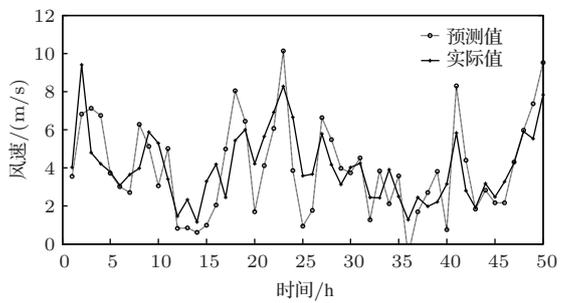


图 19 ARIMA 风速预测值与实际值对比曲线

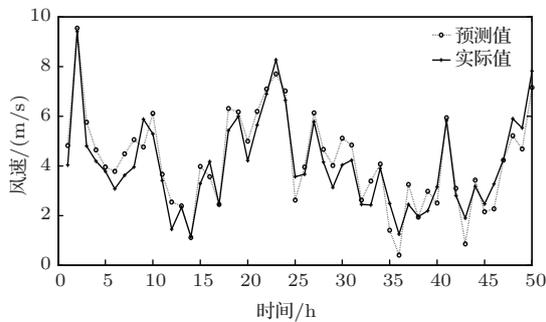


图 16 PSO-LSSVM 风速预测值与实际值对比曲线

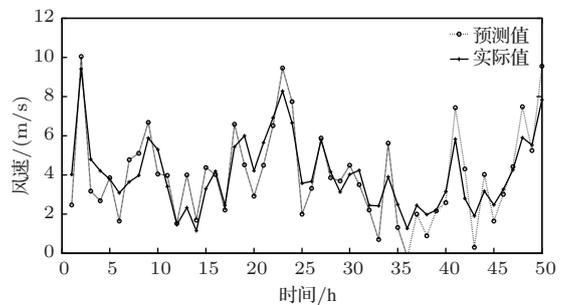


图 20 Elman 神经网络风速预测值与实际值对比曲线

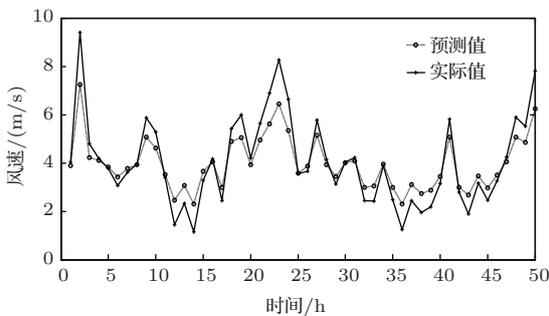


图 17 LSSVM 风速预测值与实际值对比曲线

表 1 给出了文中几种预测方法的风速预测值与实际值误差的 MSE 与平均绝对误差 (MAE) 的对比, 其中, MSE 与 MAE 都是衡量预测精度的指标, 其定义如下:

$$e_{mse} = \frac{1}{N} \sum_{k=1}^N (w(k) - \hat{w}(k))^2,$$

$$e_{mae} = \frac{1}{N} \sum_{k=1}^N |w(k) - \hat{w}(k)|.$$

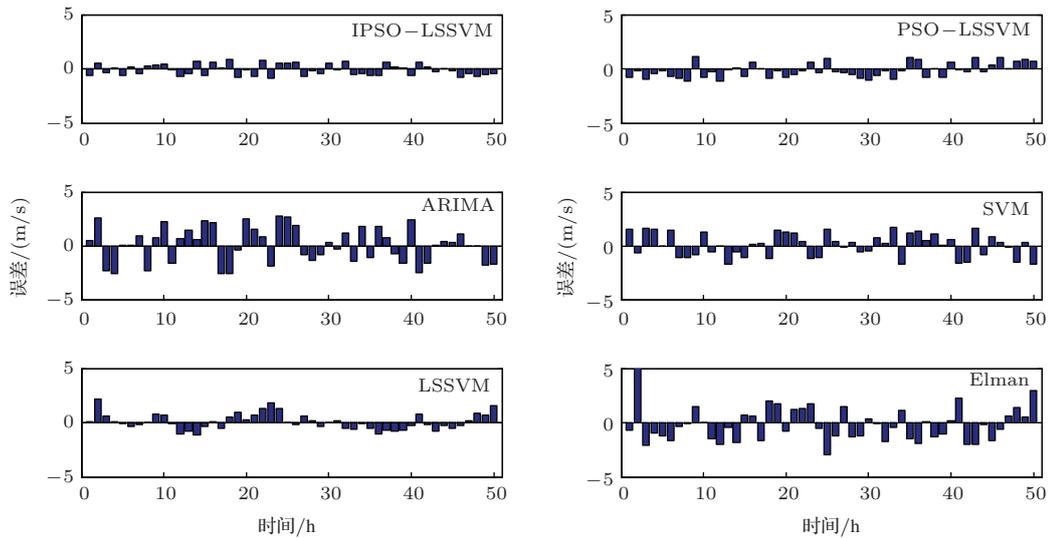


图 21 预测误差分布图

表 1 几种预测方法预测指标对比

预测方法	ϵ_{mse}	ϵ_{mae}
本文方法	0.2761	0.4681
PSO-LSSVM	0.5530	0.7754
ARIMA	2.6252	1.3735
SVM	1.1567	0.9231
LSSVM	0.6521	0.8674
Elman 神经网络	2.4613	1.2748

从仿真结果中图 15 至图 20 的风速预测值与实际值的对比曲线、图 21 的误差分布图以及表 1 的预测指标对比可知, 本文的基于相空间重构与参数在线修正 IPSO-LSSVM 的短期风速预测方法在预测精度与预测效果上都要优于其他常见的预测方法。

6 结 论

在短期风速的预测研究中, 由于风速变化的随机性以及非平稳性, 导致短期风速的预测效果不理想. 本文通过 0-1 混沌测试法确定短期风速时间序列具有混沌特性. 在此基础上, 结合相空间重构技术, 通过最大 Lyapunov 指数确定短期风速时间序列的最大可预测尺度. 提出一种参数在线修正的 IPSO-LSSVM 预测模型, 通过现场采集的采样间隔为一小时的短期风速数据进行了预测效果的研究, 可得出如下结论:

1) 短期风速时间序列具有混沌特性, 说明短期风速时间序列具有初值敏感性, 对其长期预测是不

可能的, 但是其确定性结构决定了系统的短期预测是可能的, 过大的预测步长会降低预测的精度.

2) 通过实际采集的短期风速数据进行最大 Lyapunov 指数分析, 确定风速时间序列最大可预测尺度为 4, 对于不同的风速数据其最大 Lyapunov 指数可能并不相同, 因此可预测尺度也不相同, 在利用本文方法预测时只需要修改 LSSVM 模型的输入输出矩阵维数即可.

3) LSSVM 是针对小样本情况下研究统计学习规律的理论, 因此本文的预测方法特别适用于短期风速数据样本量比较小的情况, 对于数据样本精度并无过多要求. 仿真结果表明本文的参数在线修正的 IPSO-LSSVM 预测方法对于短期风速时间序列取得了较好的预测效果, 该预测方法对于其他时间序列的预测也应该是适用的, 本文未来将要进行该方面的研究工作.

参考文献

- [1] Xiu C B, Liu X T, Zhang X, Yu T T 2013 *Power System Protection and Control* **41** 14 (in Chinese) [修春波, 刘新婷, 张欣, 于婷婷 2013 电力系统保护与控制 **41** 14]
- [2] Yang X Y, Sun B J, Zhang X F, Li L X 2012 *Proc. the CSEE* **32** 35 (in Chinese) [杨锡运, 孙宝君, 张新房, 李利霞 2012 中国电机工程学报 **32** 35]
- [3] Ma L, Benoudjit N 2011 *Appl. Energy*. **88** 2463
- [4] Ma L, Luan S Y, Jiang C W, Liu H L, Zhang Y 2009 *Renew. Sustain. Energy Rev.* **13** 915
- [5] Pelikán E, Eben K, Resler J, Juru P, Krc P, Brabec M, Brabec T, Musilek P 2010 *9th Conference on En-*

- vironment and Electrical Engineering (Piscataway, NJ: IEEE) p45
- [6] Cuo L, Zhang Y X, Wang Q C 2013 *J. Climate* **26** 85
- [7] Erdem E, Shi J 2011 *Appl. Energy* **88** 1405
- [8] Jiang J L, Lin G M 2008 *Control Theory Appl.* **25** 374 (in Chinese) [蒋金良, 林广明 2008 控制理论与应用 **25** 374]
- [9] Liu H, Tian H Q, Li Y F 2012 *Appl. Energy* **98** 415
- [10] Li H J, Liu Y N, Wei Z N, Li X L, Cheung K W, Sun Y H, Sun G Q 2013 *Electr. Power Autom. Equip.* **33** 28 (in Chinese) [李慧杰, 刘亚男, 卫志农, 李晓露, Kwok W Cheung, 孙永辉, 孙国强 2013 电力自动化设备 **33** 28]
- [11] Wang Y, Wu D L, Guo C X, Wu Q H, Qian W Z, Yang J 2010 *IEEE PES General Meeting* (Piscataway, NJ: IEEE) p1
- [12] Li X, Wang X, Zheng Y H, Li L X, Zhou L D, Sheng X K 2014 *International Conference on Renewable Energy and Environmental Technology* (Zurich-Durnten: Trans Tech) p1825
- [13] Zeng J, Zhang H 2011 *Acta Energetica Solaris Sin.* **32** 296 (in Chinese) [曾杰, 张华 2011 太阳能学报 **32** 296]
- [14] Louka P, Galanis G, Siebert N, Kariniotakis G, Katsafados P, Pytharoulis I, Kallos G 2008 *J. Wind Eng. Ind. Aerodyn.* **96** 2348
- [15] Liu J B, Ding T 2012 *Acta Energetica Solaris Sin.* **33** 1131 (in Chinese) [刘进宝, 丁涛 2012 太阳能学报 **33** 1131]
- [16] Fourati F, Chtourou M 2007 *Simul. Model. Pract. Theory* **15** 1016
- [17] Xiao H F, Ding T 2011 *Proceedings of the 2011 International Conference on Informatics, Cybernetics, Computer Engineering* (Heidelberg: Springer) p479
- [18] Soman S S, Zareipour H, Malik O, Mandal P 2010 *North American Power Symposium 2010* (Piscataway, NJ: IEEE) p1
- [19] Sheng Z 2012 *Acta Phys. Sin.* **61** 219401 (in Chinese) [盛峥 2012 物理学报 **61** 219401]
- [20] Tahat A, Marti J, Khwaldeh A, Tahat K 2014 *Chin. Phys.* **23** 046101
- [21] Li H C, Zhao J S 2005 *Chin. Phys. Lett.* **22** 2776
- [22] Tang Z J, Ren F, Peng T, Wang W B 2014 *Acta Phys. Sin.* **63** 050505 (in Chinese) [唐舟进, 任峰, 彭涛, 王文博 2014 物理学报 **63** 050505]
- [23] Wu X D, Wang Y N, Liu W T, Zhu Z Y 2011 *Chin. Phys.* **20** 069201
- [24] Gottald G A, Melbourne I 2009 *SIAM J. Appl. Dyn. Syst.* **8** 129
- [25] Gottald G A, Melbourne I 2009 *Nonlinearity* **22** 1367
- [26] Kim H S, Eykholt R, Salas J D 1999 *Physica D* **127** 48
- [27] Zhang H B, Sun X D, He Y L 2014 *Acta Phys. Sin.* **63** 040505 (in Chinese) [张洪宾, 孙小端, 贺玉龙 2014 物理学报 **63** 040505]
- [28] Grassberger P, Procaccia I 1983 *Physica D* **9** 189
- [29] Wolf A, Swift J B, Swinney H L, Vastano J A 1985 *Physica D* **16** 285
- [30] Suykens J A K, Vandevale J 1999 *Neural Process. Lett.* **9** 293
- [31] Kennedy J, Eberhart R 1995 *Proceedings of the 1995 IEEE International Conference on Neural Networks* (Piscataway, NJ: IEEE) p1942
- [32] Regis R G 2014 *J. Comput. Sci.* **5** 12
- [33] Dong Z S, Zhang X Y, Zeng J C 2013 *Trans. Can. Soc. Mech. Eng.* **37** 1189
- [34] Letting L K, Munda J L, Hamam Y 2012 *Solar Energy* **86** 1689

Chaotic characteristics analysis and prediction for short-term wind speed time series*

Tian Zhong-Da^{1)†} Li Shu-Jiang¹⁾ Wang Yan-Hong¹⁾ Gao Xian-Wen²⁾

1) (*College of Information Science and Engineering, Shenyang University of Technology, Shenyang 110870, China*)

2) (*College of Information Science and Engineering, Northeastern University, Shenyang 110819, China*)

(Received 30 June 2014; revised manuscript received 15 August 2014)

Abstract

A short-term wind speed time series prediction is studied. First, 0-1 test method for chaos is used to identify the short-term wind speed time series that has chaotic characteristics. Through phase space reconstruction, the delay time is determined by using C-C algorithm; and the embedding dimension is determined by using G-P algorithm. Then a least square support vector machine with parameters online modified is proposed, so that an improved particle swarm optimization algorithm may be used for the prediction of parameters optimization. Simulation experiment shows that the present method for its prediction accuracy, prediction error, and prediction effect is better than other prediction methods. Thus the proposed prediction method is effective, and feasible.

Keywords: short-term wind speed, time series, chaotic, prediction

PACS: 05.45.Tp

DOI: [10.7498/aps.64.030506](https://doi.org/10.7498/aps.64.030506)

* Project supported by the Key Program of the National Natural Science Foundation of China (Grant No. 61034005).

† Corresponding author. E-mail: tianzhongda@126.com