

含核函数切换的风电功率短期预测新方法

欧阳庭辉, 查晓明, 秦亮, 熊一, 夏添, 黄鹤鸣

(武汉大学 电气工程学院, 湖北 武汉 430072)

摘要: 为了降低大规模风电接入对电网造成的潜在威胁, 提出基于核函数切换机制的混沌时间序列预测新方法, 以进一步提高短期风电功率预测性能。首先, 结合互信息法和虚假邻近点法实现原始风电功率序列的相空间重构, 通过递归图和最大 Lyapunov 指数验证了风电功率是来自含确定性和随机性的混沌系统, 说明了混沌预测方法的可行性。其次, 给出了使用核函数进行混沌时间序列预测的实现方法, 结合训练样本分析了该方法优于传统预测方法, 并结合训练结果提出了使用支持向量机(SVM)训练最优核函数的切换机制, 进一步提高了预测精度。最后, 以美国 BPA 数据为实例, 通过预测误差指标的对比分析, 说明了含切换机制的核函数预测法可有效地实现风电功率短期预测, 同时也证明了该方法可较好地提高风电预测性能。

关键词: 风电; 预测; 核函数; 支持向量机; 切换机制; 混沌时间序列; 风电功率预测

中图分类号: TM 614

文献标识码: A

DOI: 10.16081/j.issn.1006-6047.2016.09.012

0 引言

随着化石能源的不断消耗, 能源危机日益严重, 为此, 全世界范围内都在大力发展可再生能源^[1]。目前, 风电作为一种丰富的、可开发利用的资源, 在电网中的渗透率逐步提高。然而, 由于风能具有随机性和波动性, 特别是在大规模、高集中度的风电发展模式下^[2], 风电带来的挑战大于机遇, 如 2008 年美国德州发生了一场大规模、危害较大的大功率下坡事件^[3]。如果能够对风电功率进行提前预测, 不仅可以指导调度部门提前制定有效的发电计划, 而且可以通过分析未来风电变化, 有效预防爬坡等具有重大威胁的风电事件, 保证电力系统的安全、稳定运行, 因此, 对风电功率进行准确的预测意义重大。

目前针对风电功率的预测方法主要分为 2 类: 物理模型和统计模型^[4]。其中, 物理模型主要是利用数值天气预报的结果预测风速, 然后通过实际风电功率曲线求取未来风电功率^[5], 该方法可以预测较长时间的功率变化, 但预测精度较差。相对而言, 统计模型在较短期的风电功率预测时精度较高, 它主要通过大量历史实测数据进行模型训练, 寻找历史输入与输出的对应关系^[6], 如时间序列模型^[7]、神经网络模型^[8]、卡尔曼滤波模型^[9]和支持向量机(SVM)^[10]等方法。这些方法都是基于历史风电功率时间序列进行研究的, 而时间序列通常是所研究系统的外在行为表现, 包含了系统特征和运动规律, 因此通过时间序列的研究可把握系统的本质, 从而实

现预测。考虑到时间序列可以来自随机系统或确定系统, 确定系统又分为线性和非线性^[11]。而风作为大气动力学系统的一个重要组成因素^[12], 且大气动力学系统具有确定性, 因此由风产生的风电功率时间序列也具有确定性, 是可预测的。然而风电不仅有确定性特性, 也有随机性特性, 为了反映这种特性, 文献[13]中对实例风电功率数据进行了混沌性分析, 提出采用混沌预测方法实现短期风电功率预测。

在混沌时间序列预测方法中, 有基于一阶局部预测法的线性模型、基于最大 Lyapunov 指数和 Volterra 方法的非线性模型^[14-15], 其中局部预测法和 Lyapunov 法是基于历史相似性进行预测, 但由于混沌系统对初值的敏感性, 这种相似性是不完全的。同样, 低阶的 Volterra 法在预测高阶和时变混沌序列时也存在一定局限性^[13]。为了弥补混沌时间序列非线性预测精度不高的缺点, 文献[16]引入核函数将低维非线性空间映射到高维线性空间, 从而可使用高维线性模型提高预测精度。同样, 文献[17]分析了不同的核函数会有不同的预测精度, 从而提出采用组合核函数的预测法。另一方面, 为了提高预测精度, 文献[18]在风电功率预测时通过采用马尔科夫链切换机制选取不同的预测模型来实现最优预测。综上所述, 为了提高风电功率时间序列的预测精度, 本文提出基于核函数切换的风电功率短期预测方法, 一方面利用核函数将非线性预测转换成高维线性预测, 另一方面在预测过程中采用有效的模型切换机制来选取最优核函数, 从而实现含最优核函数的风电功率混沌时间序列预测新方法。

为此, 本文以美国能源部 BPA 控制区域 2013 年的风电功率数据为研究对象。首先, 对该实例序列进行混沌相空间重构, 并利用递归图法检验序列的确定性, 根据小数据量法求取最大 Lyapunov 指数检验该

收稿日期: 2015-05-30; 修回日期: 2016-06-18

基金项目: 国家重点基础研究发展计划(973 计划)资助项目(2012CB215101)

Project supported by the State Key Basic Research Development Program of China(973 Program)(2012CB215101)

序列的混沌性。其次,分析含不同核函数的混沌时间序列预测方法的性能,结合 SVM,训练得到最优核函数的切换机制,并按照该机制指导风电功率时间序列预测。最后,给出 3 种误差指标,通过对算例数据的预测结果进行对比分析,验证了本文所提方法的有效性,并为有效地捕获爬坡特征的爬坡事件预测方法提供参考。

1 风电功率序列混沌性分析

Pakard 认为^[19],确定任一系统的状态所需要的全部动力学信息可反映在该系统任一变量的时间序列中,因此把单变量时间序列嵌入新的坐标空间,形成的轨迹可以保留原状态空间轨道的主要特征。即重构的相空间与原始系统动力学行为等价,2 个状态空间满足微分同胚^[20],即通过混沌相空间重构可以指导实现混沌时间序列的预测。根据 Takens 延迟嵌入定理^[20],在对风电功率序列进行相空间重构时,假设实测风电功率序列为 $\{x_n\}_{n=1}^N$,若给定延迟时间 τ 和嵌入维数 m 这 2 个变量,则按照式(1)可重构相空间。

$$\mathbf{x}_n = (x_n, x_{n+\tau}, \dots, x_{n+(m-1)\tau}) \in \mathbf{R}^m \quad n=1, 2, \dots, N_0 \quad (1)$$

其中, \mathbf{x}_n 为重构后的相空间相点; $N_0 = N - (m-1)\tau$ 。对无噪声、无限长的时间序列,可以任意选择延迟时间 τ ,且嵌入维数 m 只需满足 $m \geq 2d + 1$ 即可,其中 d 为系统动力学维数。然而实际的风电功率时间序列必然会受噪声影响,因此通常先对序列进行去噪,然后根据具体方法选取延迟时间和嵌入维数。

1.1 相空间重构参数选取

1.1.1 延迟时间选取

选取延迟时间 τ 的目标是为了保证间隔 τ 的 2 个独立又不完全无关的量作为重构空间中的独立坐标时,可有效地实现吸引子的重构。为了适应风电功率时间序列的分析,本文选用互信息法^[21]选取合适的延迟时间 τ 。

假设风电时间序列为 $\{x_n\}$,则相隔时间间隔 τ 的 2 个观测量 x_n 和 $x_{n+\tau}$,它们之间的互信息值可表示为:

$$\begin{aligned} I(x_n, x_{n+\tau}) &= \sum_{n=1}^N P(x_n, x_{n+\tau}) \ln \frac{P(x_n, x_{n+\tau})}{P(x_n)P(x_{n+\tau})} = \\ &= - \sum_{n=1}^N P(x_n) \ln P(x_n) - \sum_{n=1}^N P(x_{n+\tau}) \ln P(x_{n+\tau}) + \\ &= \sum_{n=1}^N P(x_n, x_{n+\tau}) \ln P(x_n, x_{n+\tau}) \end{aligned} \quad (2)$$

其中, $P(x_n)$ 、 $P(x_{n+\tau})$ 分别为观测量 x_n 、 $x_{n+\tau}$ 的概率分布, $P(x_n, x_{n+\tau})$ 为这 2 个观测量的联合概率分布,可通过对风电功率时间进行概率统计获得。通常用 $H(X) = - \sum_i P(x_i) \ln P(x_i)$ (x_i 为变量 X 中的每个取值元素)表示离散变量 X 的信息熵,则可将式(2)表示

成各变量信息熵间的关系。 τ 的选择既不能太大,也不能太小,当 τ 太小时,重构后吸引子十分靠近对角线,使相空间杂乱无章;当 τ 太大时,吸引子可能投影到完全不相关的方向上,不能反映系统的演化规律。通常选择 $I(x_n, x_{n+\tau})$ 第一次达到局部极小值时对应的时延作为重构相空间的延迟时间 τ ^[22],此时产生的冗余最小,具有最大的独立性。本文按式(2)计算样本数据的互信息值,计算结果如图 1 所示,可知 $I(x_n, x_{n+\tau})$ 第一次达到极小值时 $\tau=9$,即为所求延迟时间。

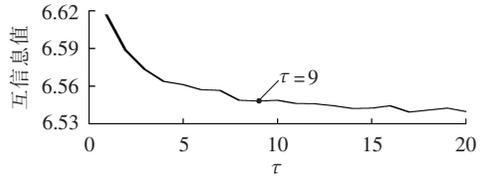


图 1 互信息法求相空间重构延迟时间

Fig.1 Delay time of phase-space reconstruction calculated by mutual information method

1.1.2 嵌入维数选取

在对风电功率序列进行相空间重构时,嵌入维数的选取必须合适,如果 m 取得太小,吸引子可能因为折叠而造成自相交;如果 m 取得太大,在对实际受噪声影响的序列进行分析时,计算工作量将增大。因此本文采用虚假邻近点法^[23]确定嵌入维数 m 。若风电功率时间序列重构后如式(1)所示,假设 \mathbf{x}_n 的最近邻点为 $\mathbf{x}_{\eta(n)}$,当嵌入维数从 m 增大到 $m+1$ 时,相点 \mathbf{x}_n 中增加分量 $\mathbf{x}_{n+m\tau}$,此时邻近点间的距离变为:

$$\|\mathbf{x}_{\eta(n)} - \mathbf{x}_n\|_2^{m+1} = \sqrt{(\|\mathbf{x}_{\eta(n)} - \mathbf{x}_n\|_2^m)^2 - (\mathbf{x}_{\eta(n)+m\tau} - \mathbf{x}_{n+m\tau})^2} \quad (3)$$

若满足 $m+1$ 维下邻近点距离 $\|\mathbf{x}_{\eta(n)} - \mathbf{x}_n\|_2^{m+1}$ 与 m 维下邻近点距离 $\|\mathbf{x}_{\eta(n)} - \mathbf{x}_n\|_2^m$ 相差不大,则说明在 \mathbf{R}^m 空间中吸引子完全打开,不再存在虚假邻近点,此时 m 即为所求嵌入维数。对按式(1)重构后的序列进行嵌入维数分析,结果如图 2 所示。

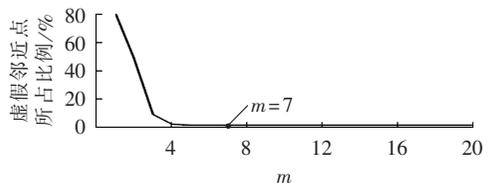


图 2 虚假邻近点法求嵌入维数

Fig.2 Embedding dimension calculated by false nearest neighbor method

图 2 给出了随着嵌入维数 m 增大,虚假邻近点所占比例的变化趋势。由图可知,当 $m=7$ 时,虚假邻近点数不再随维数 m 增加而减少,此时可认为系统的几何结构被完全打开,因此 $m=7$ 即为所求嵌入维数。根据上述求得的延迟时间和嵌入维数,结合式(1)可完成风电功率时间序列的重构。

1.2 混沌特性分析

对于上述分析给出的重构风电功率混沌时间序列,在进行混沌特性分析时,主要考虑确定性和混沌性 2 个特性,其中确定性检测是为了保证构建可预测的模型,混沌性检测是为了保证可使用混沌方法进行预测。

1.2.1 确定性检测

对于实例的风电功率时间序列,本文采用递归图法^[11]来检测风电功率是否来自确定系统。假设按照式(1)重构后的相空间序列为 $\{\mathbf{x}_n\}_{n=1}^{N_0}$, 设 i 和 j 时刻的相点间的距离为 $d_{ij} = \|\mathbf{x}_i - \mathbf{x}_j\|$, 若时间序列以 T 为周期, 则当 $|i-j| = kT (k=0, 1, \dots)$ 时, $d_{ij}=0$, 利用这种性质, 以 i 为横坐标, j 为纵坐标, 可画出 $d_{ij}=0$ 时的递归图。由于考虑到实际序列中的噪声影响, 通常以 $d_{ij} < r$ (r 为人为设定的距离阈值) 为标准画递归图, 则可得如图 3 所示的风电功率序列重构后相点的递归图。

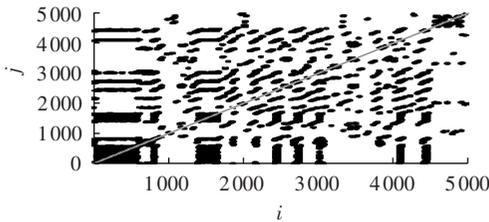


图 3 风电功率序列重构空间的递归图
Fig.3 Recurrence plot of reconstruction space of wind power series

从图中可以看出,风电功率时间序列重构后相点的递归图呈现复杂的轨迹迹象,但从一些与对角线平行的小带中可以看出一些周期特性,因此可认为风电功率时间序列具有确定性特性,可构建预测模型进行预测。

1.2.2 混沌性检测

在判断风电功率时间序列具有确定性后,需要进一步检测其是否具有混沌性,以便确定采用混沌方法进行预测。由于最大 Lyapunov 指数是混沌系统的一个重要几何不变量,因此,通常用它来判断系统是否具有混沌性。当 Lyapunov 指数值大于零时,说明系统具有混沌特性。为此,本文采用稳定性较强、对数据点要求较少的小数据量法^[24]来求取重构相空间的最大 Lyapunov 指数。

设重构的时间序列为 $\{\mathbf{x}_n\}_{n=1}^{N_0}$, 取一个相点作为初始点 $\mathbf{x}(n_1)$, 在相空间中寻找其最邻近点 $\mathbf{x}(n'_1)$, 则距离 L_1 为:

$$L_1 = \|\mathbf{x}(n_1) - \mathbf{x}(n'_1)\| \quad |n_1 - n'_1| > p \quad (4)$$

其中, p 为时间序列的平均周期。按时间 T 演化后, 上述 2 个相点演化成点 $\mathbf{x}(n_1+T)$ 和 $\mathbf{x}(n'_1+T)$, 此时的距离为 L'_1 。根据混沌相空间轨迹的指数分离规律 $L'_1 = L_1 e^{\lambda_1 T}$, 则其指数增长率为:

$$\lambda_1 = \frac{1}{T} \ln \frac{L'_1}{L_1} \quad (5)$$

同理, 以 $\mathbf{x}(n_1+T)$ 为新初始点 $\mathbf{x}(n_2)$, 寻找其邻近点 $\mathbf{x}(n'_2)$, 且满足 $\mathbf{x}(n_2)\mathbf{x}(n'_2)$ 与 $\mathbf{x}(n_2)\mathbf{x}(n'_1+T)$ 的夹角尽可能地小, 同样按照上述步骤, 经过时间 T 的演化, 求出指数增长率 λ_2 。重复上述过程直到演化到终点 \mathbf{x}_{N_0} 为止, 记演化总步数为 M , 则最大 Lyapunov 指数估计值可取其指数增长率的平均值:

$$\lambda_m = \frac{1}{M} \sum_{i=1}^M \frac{1}{T} \ln \frac{L'_i}{L_i} \quad (6)$$

其中, λ_m 为所求最大 Lyapunov 指数。本文以重构后的风电功率时间序列为例, 按上述方法求最大 Lyapunov 指数。由于演化时间 T 太长会造成吸引子结构上的重叠、突变, 从而造成最大 Lyapunov 指数错误估计, 为此, 本文取 $T=20$, 并绘出最大 Lyapunov 指数随着演化时间的变化规律, 如图 4 所示。

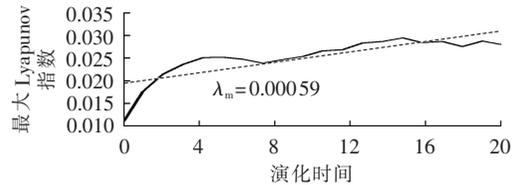


图 4 最大 Lyapunov 指数随演化时间的变化规律
Fig.4 Maximum Lyapunov value varying with evolution time

从图中可见,最大 Lyapunov 指数 $\lambda_m=0.00059 > 0$, 可知风电功率时间序列来自确定性的混沌系统, 因此在对风电功率进行预测时, 可选取合适的混沌时间序列预测方法。

2 含核函数切换的预测模型

目前,混沌预测方法有局部线性预测法、基于最大 Lyapunov 指数的预测模型以及基于非线性的 Volterra 预测方法^[14-15]。局部线性预测法和最大 Lyapunov 指数预测模型都是建立在混沌相空间中吸引子短期内的自相似性基础上, 但由于混沌系统具有初值敏感性、长期不可预测性和分叉结构, 从而造成这些方法预测时间较短, 预测精度不高。相对而言, Volterra 预测方法是一种非线性模型, 可以较好地提高短期内的预测精度。为了进一步提高混沌时间序列的预测精度, 需要更高阶的非线性预测模型, 然而高阶模型一般实现困难, 例如 Volterra 预测方法常采用二阶截断模型^[13]。为了实现非线性高阶预测模型, 本文提出如下方法。首先假设混沌时间序列为 $\{\mathbf{x}_n\}$, 则其预测模型可以表述为:

$$y_{n+1} = f(\mathbf{x}_n, \mathbf{x}_{n-1}, \dots, \mathbf{x}_{n-q}) + e \quad (7)$$

其中, y_{n+1} 为所求预测量; q 为历史相关数据个数; e 为误差波动项^[25]; $f(\ast)$ 为预测模型, 考虑到混沌时间序列的非线性, f 通常为非线性函数, 如 Volterra

函数。如果存在非线性函数 Φ , 可将序列 $\{x_n\}$ 映射到高维 Hilbert 空间 H 中, 如式(8)所示。

$$\Phi: \begin{matrix} \mathbf{R}^m \rightarrow H \\ x \rightarrow \Phi(x) \end{matrix} \quad (8)$$

通过合适的非线性变换, 在高维空间 H 中, 低维的非线性预测模型可变为高维线性模型, 即式(7)可变为式(9), 此时通过线性预测模型即可实现较高精度的预测性能。

$$y_{n+1} = \sum_{i=0}^q \omega^T \Phi(x_{n-i}) + e \quad (9)$$

其中, ω 为权重向量。然而考虑到非线性变换 Φ 的实现困难, 为了简化计算, 借助 SVM 中核函数不仅可将低维非线性模型投射成高维的线性模型, 同时能以内积形式来代替非线性变换的实现。因此, 本文提出利用核函数来实现非线性混沌时间序列的高维映射, 然后通过求解线性预测模型实现高性能预测。

2.1 含核函数的预测模型

核函数的选取必须满足 Mercer 条件^[26], 目前在 SVM 中常用的核函数有 3 种: 多项式核函数、径向基(RBF)核函数和 Sigmoid 核函数, 它们的表达式如式(10)所示。

$$\begin{cases} K_{\text{poly}}(\mathbf{x}, \mathbf{y}) = (\langle \mathbf{x}, \mathbf{y} \rangle + 1)^b \\ K_{\text{RBF}}(\mathbf{x}, \mathbf{y}) = \exp(-\|\mathbf{x} - \mathbf{y}\|^2 / \sigma^2) \\ K_{\text{Sigmoid}}(\mathbf{x}, \mathbf{y}) = \tanh(v \langle \mathbf{x}, \mathbf{y} \rangle + a) \end{cases} \quad (10)$$

假设给定混沌时间序列 $\{x_n\}$, 可根据当前状态 x_n 预测未来的演化状态 x_{n+1} , 即 $x_{n+1} = f(x_n)$ 。按照前文分析, 核函数可利用内积完成高维映射, 因此式(9)中的非线性变换可通过核函数实现。然而在训练高维线性模型过程中, 需要若干具有相同特性的训练样本, 根据混沌系统的几何不变量可知, 结构完全打开的混沌系统, 其邻近点空间通过核函数映射到高维空间中仍是邻近空间, 同时考虑邻近点具有近似的演化规律, 因此可在相空间中取 k 个当前状态 x_n 的最邻近点用于训练预测模型。假设核函数为 $K(\mathbf{x}, \mathbf{y}) = \langle \Phi(\mathbf{x}), \Phi(\mathbf{y}) \rangle$, 则每个相点可分别与其 h 个历史点构成核函数矩阵 K_i , 其中 $K_{ij} = K(\mathbf{x}_i, \mathbf{x}_{i-j})$ ($i=1, 2, \dots, k; j=1, 2, \dots, h$)。相应 H 空间中的线性预测模型可表述为 $\hat{y}_i = K_i \cdot \omega$, 则其预测误差为:

$$e_i = y_i - \hat{y}_i = y_i - K_i \cdot \omega \quad (11)$$

其中, y_i, \hat{y}_i 分别为观测值和预测值; ω 为权重向量。在求解该线性预测模型时, 由最小二乘法(LMS)可知, 以均方误差最小为目标, 其最优解可由式(12)求得, 即最优解 $\omega^* = P^{-1}Q$ 。

$$\begin{cases} \Delta \omega = -2Q + 2P \cdot \omega = 0 \\ Q = K^T \cdot y, \quad P = K^T \cdot K \end{cases} \quad (12)$$

根据式(12)求得的最优参数 ω^* , 即可完成预测模型的构建。本文取 $k=5, h=3$, 对历史样本数据进

行训练可得到含具体参数的预测模型, 并由预测模型对历史样本进行拟合, 结果如图 5 所示, 图中给出了历史样本中 2013 年 5 月 18 日至 21 日的 1000 组拟合结果。同时为了对比分析, 可以以自回归-滑动平均模型 ARMA(1,2)作为参考标准。

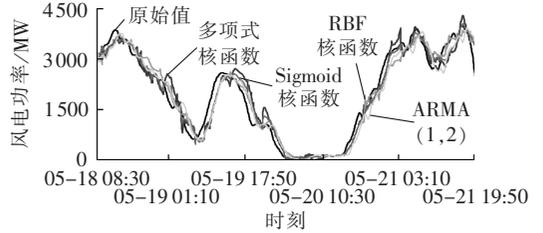


图 5 核函数模型对历史观测值的拟合结果
Fig.5 Results of historical data fitting by kernel function models

2.2 模型切换机制

根据图 5 的仿真结果可知, 各个模型在时间序列的不同时刻具有不同的表达能力, 为了达到更好的预测效果, 可以在不同时刻选择最合适的核函数进行预测。为了得到不同时刻核函数的切换机制, 本文将 3 类核函数分别对应为 3 个类别, 考虑 SVM 的泛化能力较强, 只需有限样本数据即可较好地训练分类器^[27], 因此, 在寻找切换机制时本文采用支持向量分类机。

假设重构后的混沌时间序列为 $\{x_n\}_{n=1}^{N_0}$, 并取 x_n 为输入量, 根据图 5 的仿真结果, 取误差最小的核函数作为输出量 c_n , 因此得到训练集为:

$$P = \{(\mathbf{x}_1, c_1), (\mathbf{x}_2, c_2), \dots, (\mathbf{x}_{N_0}, c_{N_0})\} \in (\mathbf{R}^m \times S)^{N_0} \quad (13)$$

其中, $S = \{K_{\text{poly}}, K_{\text{RBF}}, K_{\text{Sigmoid}}\}$ 。由于 S 是包含 3 个元素的集合, 而基本的支持向量机(BSVM)是基于二分类的模型, 为了实现多类别的分类, 本文采用分级分类方法^[28], 即每次只区分一个最主要的类别, 对该类别的样本定义 $y_n = 1$, 余下的样本定义 $y_n = -1$, 为使模型适用性更好, 通过引入松弛变量 $\xi_i \geq 0$ 可构造 SVM 如式(14)所示。

$$\begin{aligned} \min_{\omega, b} & \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^{N_0} \xi_i \\ \text{s.t.} & y_i (\langle \omega, \mathbf{x}_i \rangle + b) \geq 1 - \xi_i \\ & \xi_i \geq 0 \quad i=1, 2, \dots, N_0 \end{aligned} \quad (14)$$

其中, ω, b 为 SVM 模型参数; N_0 为训练样本长度。为了寻找模型最优解, 通过引入 Lagrange 乘子向量 α 和 β , 可得到 Lagrange 函数:

$$L(\omega, b, \alpha) = \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^{N_0} \xi_i - \sum_{i=1}^{N_0} \alpha_i [y_i (\langle \omega, \mathbf{x}_i \rangle + b) - 1 + \xi_i] - \sum_{i=1}^{N_0} \beta_i \xi_i \quad (15)$$

其中, C 为惩罚因子, 为常数。通过式(15)可构造原问题的对偶问题如式(16)所示。

$$\begin{aligned} \max_{\alpha, \beta} & -\frac{1}{2} \sum_{i=1}^{N_0} \sum_{j=1}^{N_0} y_i y_j \alpha_i \alpha_j \langle \mathbf{x}_i, \mathbf{x}_j \rangle + \sum_{i=1}^{N_0} \alpha_i \\ \text{s.t.} & \begin{cases} \sum_{i=1}^{N_0} y_i \alpha_i = 0 \\ C - \alpha_i - \beta_i = 0 \quad i=1, 2, \dots, N_0 \\ \alpha_i \geq 0 \quad i=1, 2, \dots, N_0 \\ \beta_i \geq 0 \quad i=1, 2, \dots, N_0 \end{cases} \end{aligned} \quad (16)$$

通过消去变量 β 求解该凸规划模型, 得到解 $\alpha^* = (\alpha_1^*, \dots, \alpha_m^*)^T$, 则一个 BSVM 分类器的最优参数可通过式(17)给出。

$$\begin{cases} \omega^* = \sum_{i=1}^m \alpha_i^* y_i \mathbf{x}_i \\ b^* = y_i - \sum_{i=1}^m \alpha_i^* y_i \langle \mathbf{x}_i, \mathbf{x}_j \rangle \end{cases} \quad (17)$$

根据前文分析, 由 BSVM 的结果 $y = \langle \omega^*, \mathbf{x} \rangle + b^*$ 可判断出样本属于哪一类。而对于本文重构的混沌时间序列, 按照分级分类方法需构造 2 个 BSVM, 根据式(14)–(17)的计算可得到 2 个分类超平面, 它们共同组成了风电功率混沌预测模型的核函数切换机制。

3 算例分析

本文以美国 BPA(Bonneville Power Administration) 控制区域 2013 年采样间隔为 5 min 的风电功率数据为数据样本集, 取 1—5 月份共 43476 组数据为仿真算例, 首先根据相空间重构的基本思想重构原始风电功率序列, 并通过最大 Lyapunov 指数验证了重构的序列具有混沌性, 如图 4 所示。根据前文分析, 取 1—4 月份的数据进行预测模型训练, 取 5 月份的数据作为测试集用于预测和评估。为了评估本文基于核函数切换机制的风电功率预测方法的性能, 取测试集中 5 月 27 日到 5 月 28 日共 500 组数据进行预测仿真, 其预测结果如图 6 所示。图 6 中分别给出了使用 3 类核函数进行单独预测的结果和使用切换核函数机制的混沌时间序列预测结果。

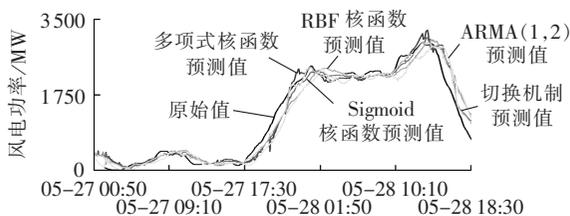


图 6 风电功率混沌时间序列预测结果

Fig.6 Results of chaotic time series prediction for wind power

为了说明本文提出的基于核函数切换机制的风电功率预测方法的有效性, 需引入误差指标对比分析该方法的预测性能。文献[29]中将误差指标分为纵向和横向 2 类, 其中纵向误差从宏观角度描述预测系

统的长期运行状态, 横向误差则从时间角度研究预测的变化。因此, 在对短期风电功率预测结果进行分析时, 本文取纵向误差如绝对值平均误差 MAE (Mean Absolute Error)、均方根误差 RMSE(Root Mean Squared Error)和横向误差相关系数 CC(Correlation Coefficient)进行分析, 其表达式如式(18)所示。

$$\begin{cases} e_{\text{MAE}} = \frac{1}{s} \sum_{i=1}^s |x_i - \hat{x}_i| / s \\ e_{\text{RMSE}} = \sqrt{\frac{1}{s} \sum_{i=1}^s (x_i - \hat{x}_i)^2} / s \\ e_{\text{CC}} = \frac{\text{cov}(x_i - \hat{x}_i)}{\sqrt{Dx_i} \sqrt{D\hat{x}_i}} \end{cases} \quad (18)$$

其中, x_i 为实测数据点; \hat{x}_i 为预测数据点; s 为预测样本数; Dx_i 表示求变量 x_i 的方差。

首先, 为了说明含核函数的模型对风电功率时间序列的表达力, 对图 5 中的仿真结果进行分析, 结果如表 1 所示。

表 1 训练样本的误差指标值

Table 1 Error indexes of training samples

预测模型	误差指标		
	MAE/MW	RMSE/MW	CC
ARMA(1,2)模型	230.5211	301.0524	0.9717
多项式核函数模型	227.4322	300.2698	0.9713
RBF 核函数模型	201.5869	256.8109	0.9787
Sigmoid 核函数模型	185.9728	244.9562	0.9805

对比 ARMA(1,2)模型可以看出, 含核函数的混沌预测模型无论是从纵向幅值表达能力, 还是横向时间延迟, 对历史风电功率时间序列均具有较好的表达效果, 因此说明采用含核函数预测模型的可行性。同样, 为了进一步说明含核函数切换模型在实际预测过程中的效果, 图 6 中还给出了 ARMA(1,2)模型的预测结果, 通过分析计算, 表 2 中给出了含核函数切换机制和不含核函数切换的混沌序列预测, 以及 ARMA(1,2)模型的预测误差指标性能。

表 2 含核函数预测模型的预测性能

Table 2 Performances of prediction models with kernel functions

预测模型	误差指标		
	MAE/MW	RMSE/MW	CC
ARMA(1,2)模型	190.4950	263.3488	0.9657
多项式核函数模型	126.8251	176.8874	0.9834
RBF 核函数模型	161.1548	217.6228	0.9751
Sigmoid 核函数模型	120.1812	163.1319	0.9852
含切换机制模型	91.5177	141.5394	0.9884

分析可知, MAE 和 RMSE 的值越小, 说明预测的幅值误差越小; 指标 CC 的值越接近于 1, 说明预测值和实际观测值的时间偏移越小。对比表 2 中几种预测方法可知, 含核函数的预测方法比 ARMA 模型性能更优; 当核函数单独预测时, 使用 Sigmoid 核函数的

预测性能较好,而通过加入了核函数切换机制,风电功率短期预测性能整体得到提高,因此可说明基于核函数切换机制的混沌时间序列预测法不仅对风电功率短期预测有效,同时可较好地提高预测性能。

4 结论

本文为了改善风电功率短期预测性能,提出基于核函数切换机制的混沌时间序列预测方法。一方面,通过递归图和最大 Lyapunov 指数验证了重构的风电功率序列具有混沌性。其次,对比含核函数映射的混沌时间序列预测法与传统 Volterra 预测法,说明了使用核函数预测的优越性。最后,通过训练核函数切换机制,基于实例数据的预测误差分析验证了本文所提方法的高预测性能。综合上述结果可知,基于核函数切换机制的混沌预测方法对风电功率短期预测有效,并可较好地提高混沌时间序列的预测性能。

参考文献:

- [1] 李俊峰,蔡丰波,乔黎明,等. 2013 中国风电发展报告[M]. 北京:中国资源综合利用专业委员会,2013:9-12.
- [2] 戚永志,刘玉田. 风电高风险爬坡有限度控制[J]. 中国电机工程学报,2013,33(13):69-75.
QI Yongzhi,LIU Yutian. Finite control of high risk wind power ramping[J]. Proceedings of the CSEE,2013,33(13):69-75.
- [3] FRANCIS N. Predicting sudden changes in wind power generation [J]. North American Wind Power,2008,5:58-60.
- [4] 张露,卢继平,梅亦蕾,等. 基于不同优化准则的风电功率预测[J]. 电力自动化设备,2015,35(5):139-145.
ZHANG Lu,LU Jiping,MEI Yilei,et al. Wind power forecasting based on different optimization criterions[J]. Electric Power Automation Equipment,2015,35(5):139-145.
- [5] 冯双磊,王伟胜,刘纯,等. 风电场功率预测物理方法研究[J]. 中国电机工程学报,2010,30(2):1-6.
FENG Shuanglei,WANG Weisheng,LIU Chun,et al. Study on the physical approach to wind power prediction[J]. Proceedings of the CSEE,2010,30(2):1-6.
- [6] ZHENG H,KUSIAK A. Prediction of wind farm power ramp rates;a data-mining approach[J]. Journal of Solar Energy Engineering,2009,131(3):376-385.
- [7] 潘迪夫,刘辉,李燕飞. 风电场风速短期多步预测改进算法[J]. 中国电机工程学报,2008,28(26):87-91.
PAN Difu,LIU Hui,LI Yanfei. Optimization algorithm of short-term multi-step wind speed forecast[J]. Proceedings of the CSEE,2008,28(26):87-91.
- [8] 李文良,卫志农,孙国强,等. 基于改进空间相关法和径向神经网络的风电场短期风速分时预测模型[J]. 电力自动化设备,2009,29(6):89-92.
LI Wenliang,WEI Zhinong,SUN Guoqiang,et al. Multi-interval wind speed forecast model based on improved spatial correlation and RBF neural network[J]. Electric Power Automation Equipment,2009,29(6):89-92.
- [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.
QING Xiangyun,YANG Fuwen,WANG Xingyu. Short-term wind speed forecasting for multiple wind farms using Bayesian Kriged Kalman mode[J]. Proceedings of the CSEE,2012,32(35):107-114.
- [10] 凌武能,杭乃善,李如琦. 基于云支持向量机模型的短期风电功率预测[J]. 电力自动化设备,2013,33(7):34-38.
LING Wuneng,HANG Naishan,LI Ruqi. Short term wind power forecasting based on cloud SVM model[J]. Electric Power Automation Equipment,2013,33(7):34-38.
- [11] 王海燕,卢山. 非线性时间序列分析及其应用[M]. 北京:科学出版社,2006:40-44.
- [12] CHEN N,QIAN Z,NABNEY I,et al. Wind power forecasts using Gaussian processes and numerical weather prediction[J]. IEEE Transactions on Power Systems,2014,29(2):656-665.
- [13] 张学清,梁军. 风电功率时间序列混沌特性分析及预测模型研究[J]. 物理学报,2012,61(19):70-81.
ZHANG Xueqing,LIANG Jun. Chaotic characteristics analysis and prediction model study on wind power time series[J]. Acta Phys Sin,2012,61(19):70-81.
- [14] 任丽佳,江秀臣,盛戈峰,等. 输电线路允许输送容量的混沌预测[J]. 中国电机工程学报,2009,29(25):86-91.
REN Lijia,JIANG Xiuchen,SHENG Gehao,et al. Prediction of transmission line rating based on chaotic time series analysis [J]. Proceedings of the CSEE,2009,29(25):86-91.
- [15] 彭春华. 电力市场短期边际电价的分时重构混沌相空间预测[J]. 中国电机工程学报,2005,25(23):80-85.
PENG Chunhua. Electricity market short-term marginal price forecasting based on period clustering restructuring chaotic phase space[J]. Proceedings of the CSEE,2005,25(23):80-85.
- [16] 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. Multivariate local prediction method for short-term wind speed of wind farm [J]. Proceedings of the CSEE,2012,32(1):24-31.
- [17] 田中大,高宪文,石彤. 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[J]. 物理学报,2014,63(16):66-76.
TIAN Zhongda,GAO Xianwen,SHI Tong. Combination kernel function least squares support vector machine for chaotic time series prediction[J]. Acta Phys Sin,2014,63(16):66-76.
- [18] PINSON P,MADSEN H. Adaptive modelling and forecasting of offshore wind power fluctuations with Markov-switching autoregressive models[J]. Journal of Forecasting,2012,31(4):281-313.
- [19] PAKARD H,CRUTCHFISLD J P,FARMER J D,et al. Geometry from a time series[J]. Physical Review Letters,1980,45(9):712-716.
- [20] TAKENS F. Detecting strange attractors in turbulence,dynamical systems and turbulence,Warwick 1980[M]. Berlin,Germany: Springer Berlin Heidelberg,1981:366-381.
- [21] 安学利,蒋东翔. 风力发电机组运行状态的混沌特性识别及其趋势预测[J]. 电力自动化设备,2010,30(3):15-19.
AN Xueli,JIANG Dongxiang. Identification of chaos characteristics and prediction of wind turbine state[J]. Electric Power Automation Equipment,2010,30(3):15-19.
- [22] 冬雷,王丽婕,高爽,等. 基于混沌时间序列的大型风电场发电功率预测建模与研究[J]. 电工技术学报,2008,23(12):125-129.
DONG Lei,WANG Lijie,GAO Shuang,et al. Modeling and

- analysis of prediction of wind power generation in the large wind farm based on chaotic time series[J]. Transactions of China Electrotechnical Society, 2008, 23(12):125-129.
- [23] KENNEL M B, BROWN R, ABBANEL H D I. Determining embedding dimension for phase-space reconstruction using a geometrical construction[J]. Physical Review A, 1992, 45(6):3403.
- [24] ROSENSTEIN M T, COLLINS J J, LUCA C J D. A practical method for calculating largest Lyapunov exponents from small data sets[J]. Physica D-Nonlinear Phenomena, 1993, 65(1-2): 117-134.
- [25] 林卫星, 文劲宇, 艾小猛, 等. 风电功率波动特性的概率分布研究[J]. 中国电机工程学报, 2012, 32(1):38-46.
LIN Weixing, WEN Jinyu, AI Xiaomeng, et al. Probability density function of wind power variations[J]. Proceedings of the CSEE, 2012, 32(1):38-46.
- [26] 李慧杰, 刘亚南, 卫志农, 等. 基于相关向量机的短期风速预测模型[J]. 电力自动化设备, 2013, 33(10):28-32.
LI Huijie, LIU Yanan, WEI Zhinong, et al. Short-term wind speed forecasting model based on relevance vector machine [J]. Electric Power Automation Equipment, 2013, 33(10):28-32.
- [27] CRISTIANINI N, SHAWE-TAYLOR J. An introduction to support vector machines: and other kernel-based learning methods [M]. Cambridge, UK: Cambridge University Press, 2000:103-112.
- [28] HSU C W, LIN C J. A comparison of methods for multiclass support vector machines[J]. IEEE Transactions on Neural Networks, 2002, 13(2):415-425.
- [29] 徐曼, 乔颖, 鲁宗相. 短期风电功率预测误差综合评价方法[J]. 电力系统自动化, 2011, 35(12):20-26.
XU Man, QIAO Ying, LU Zongxiang. A comprehensive error evaluation method for short-term wind power prediction[J]. Automation of Electric Power Systems, 2011, 35(12):20-26.

作者简介:



欧阳庭辉

欧阳庭辉(1989—),男,湖南岳阳人,博士研究生,主要研究方向为新能源发电及微电网技术(**E-mail**:thouyang@whu.edu.cn);

查晓明(1967—),男,湖北武汉人,教授,博士研究生导师,主要研究方向为电力电子功率变换、智能电网及新能源发电中的电力电子技术应用;

秦亮(1980—),男,湖北武汉人,讲师,博士,通信作者,主要研究方向为风力发电技术、多端柔性直流输电技术(**E-mail**:qinliang@whu.edu.cn)。

Short-term wind power prediction based on kernel function switching

OUYANG Tinghui, ZHA Xiaoming, QIN Liang, XIONG Yi, XIA Tian, HUANG Heming

(School of Electrical Engineering, Wuhan University, Wuhan 430072, China)

Abstract: A method of chaotic time series prediction based on the switching regime of kernel functions is proposed to further improve the performance of short-term wind power prediction for reducing the potential risk of power grid caused by the large-scale wind power integration. The mutual information method and false nearest neighbor method are applied to reconstruct the phase-space of original wind power series. The recurrence plot and the maximum Lyapunov value are used to verify that, the wind power series are from a chaotic system with certainty and randomness and the chaotic prediction is applicable. The implementation of chaotic time series prediction based on kernel functions is given and the training sample analysis proves it is better than the traditional prediction method. According to the training results, the support vector machine is proposed to train the switching regime of optimal kernel functions for future improving the prediction accuracy. As an example, the comparison among the error indexes based on the data from BPA website proves that, the prediction based on the kernel functions with switching regime can effectively realize the short-term wind power prediction with better performances.

Key words: wind power; prediction; kernel functions; support vector machines; switching regime; chaotic time series; wind power prediction