



# 东海大桥风电场短期风速序列特性及其预报

吕国钦<sup>①②</sup>, 张会琴<sup>①②\*</sup>, 李家春<sup>①②</sup>

① 中国科学院力学研究所, 流固耦合系统力学重点实验室, 北京 100190;

② 中国科学院大学工程科学学院, 北京 100049

\*联系人, E-mail: zhanghuiqin@imech.ac.cn

收稿日期: 2016-09-30; 接受日期: 2016-10-24; 网络出版日期: 2016-11-10

国家自然科学基金(编号: 11232012)和国家自然科学基金青年基金(编号: 11102212)资助项目

**摘要** 风能作为一种重要的可再生能源, 能够减少环境污染, 缓解能源短缺. 为减轻风电并网所带来的不利影响, 降低供电系统的运行成本, 精确的短期风速预报则显得十分必要. 本文基于对东海大桥风电场实测数据的分析, 指出其短期风速序列具有波动集群效应与非对称效应, 而波动反馈效应并不显著. 并进而从物理角度对产生这一现象可能的原因进行了详细阐述. 此外, 为选取适用于东海大桥风电场的预报模型, 我们以平均绝对误差(MAE)、平均绝对百分误差(MAPE)以及均方根误差(RMSE)这三个误差指标作为判别标准, 定量对比了五个模型的预报精度. 结果表明, 无论是单步预报还是多步预报, ARMA-EGARCH模型与ARMA-EGARCH-M模型的预报能力十分接近, 均优于其他模型, 且这两个模型随着提前预报步数的增加, 误差增长率较低.

**关键词** 海上风电场, 短期风速预报, 时间序列分析, 波动集群效应, 波动反馈效应, 非对称效应

**PACS:** 02.50.Fz, 89.30.Ee, 89.20.Kk

## 1 引言

风能是重要的清洁能源之一, 近些年呈现规模化发展趋势. 根据全球风能理事会(GWEC)的报告显示, 2015年, 我国新增装机容量超过30.8 GW, 居世界第一, 约占全球新增装机容量的48%. 其中海上风力发电机新增装机容量约360 MW, 排名世界第三, 仅次于德国与英国(详见GWEC报告: <http://www.gwec.net/global-figures/wind-energy-global-status/>). 然而, 与传统的火力或水力发电相比, 风电场电力输出与自然风速密切相关, 具有波动性与间歇性的特征. 这与电网对

电源平稳、持续的要求相矛盾, 是风电并网所面临的主要制约因素. 通过建立短期风速预报模型, 对风速进而对风电机组发电功率进行实时预报, 是解决这一问题, 确保电网安全运行的有效途径之一<sup>[1]</sup>.

在这一领域, 数值天气预报、统计类方法、人工智能算法以及空间关联算法是四种最为经典的预报方法<sup>[2]</sup>. 其中, 时间序列分析方法(属于统计类方法)因其节省计算资源等优势, 得到了较为广泛的研究与应用<sup>[3]</sup>. 这一方法通过对时间序列建模, 以预测未来事物的发展, 最早由George Box与Gwilym Jenkins两位学者提出并进行了系统性阐述<sup>[4]</sup>. 后来, Engle<sup>[5]</sup>突破

**引用格式:** 吕国钦, 张会琴, 李家春. 东海大桥风电场短期风速序列特性及其预报. 中国科学: 物理学 力学 天文学, 2016, 46: 124713  
Lyu G Q, Zhang H Q, Li J C. Characteristics and forecast of short-term wind speed series in the Donghai Bridge wind farm (in Chinese). Sci Sin-Phys Mech Astron, 2016, 46: 124713, doi: 10.1360/SSPMA2016-00407

了这一理论中随机误差项的方差为常数的基本假定, 提出了自回归条件异方差模型(ARCH: Autoregressive Conditional Heteroscedasticity). 不久, Bollerslev<sup>[6]</sup>再次对ARCH模型进行了推广, 建立了广义自回归条件异方差模型(GARCH: Generalized Autoregressive Conditional Heteroskedasticity). 基于这两位学者的研究, 相关扩展模型后续被一一提出, 现可以将其统称为ARCH族模型<sup>[7-9]</sup>.

将时间序列分析方法应用到风速预报, 已开展了一些工作. Kamal与Jafri<sup>[10]</sup>在分析了2年的风速数据的基础上, 建立了ARMA (Autoregressive Moving Average) 模型, 认为在短期与长期风速预报领域中时间序列分析方法均可以适用. Cadenas等人<sup>[11]</sup>则采用SARIMA (Seasonal Autoregressive Integrated Moving Average) 模型对墨西哥南部海岸的风速进行预报. 其算例表明, 当实测数据的数量有限时, 采用这一模型的预报精度优于采用人工智能网络的算法. 为进一步提高预报精度, 在经典预报模型的基础上, 新的组合建模方法被陆续提出. Liu等人<sup>[12]</sup>针对风速波动的特性, 建立时间序列与自回归条件异方差的组合模型. Wang等人<sup>[13]</sup>则建立了时间序列与模糊逻辑的组合模型, 并将这一模型用于中国海南的风电场.

风速序列特征具有区域差异, 同一模型在不同风电场的短期风速预报中的表现可能出现较大差异<sup>[3]</sup>. 由于不同区域的风速序列可能具有不同的特性, 当目标风电场发生改变时, 若预报模型无法刻画其特征,

则预报精度必然下降. 因此不能仅仅将风速作为数据进行处理, 而需深入分析风速序列的特性, 以便优选合适的预报模型.

本文通过选择不同的时间序列以及条件异方差的组合模型, 结合各个模型的特征参数, 详细讨论了我国东海大桥风电场的短期风速序列特性, 并从物理角度进行了解释. 最后, 基于风电场的实测风速数据, 分析并对比了五个短期风速预报模型在单步及多步预报情况下的精度, 优选出适用于东海大桥风电场的短期风速预报模型, 以为工程应用提供参考.

## 2 风速序列特性分析

### 2.1 数据来源

本文分析所采用的数据, 由上海勘测设计研究院提供, 为东海风电场一期某风电机组的实测风速数据. 数据共分为两组, 一组从2010年2月5日01:00至2月24日23:00, 时间间隔为60 min, 共计479个数据点; 另一组从2010年6月14日00:00至2010年6月20日23:50, 时间间隔为10 min, 共计1008个数据点, 是两组典型的海上实测风速数据, 见图1.

本文中, 我们将两组实测数据各自分为两个部分: 第一部分作为模型训练数据, 用于分析东海大桥风电场的风速序列特性, 并以此建立短期风速预报模型; 第二部分作为预报对比数据, 用于定量地讨论各个模

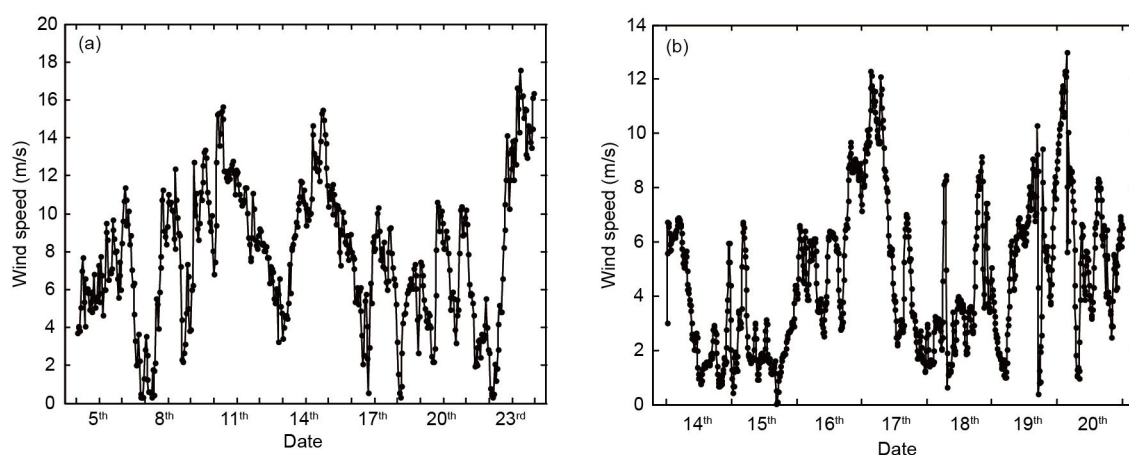


图1 东海大桥风电场某风电机组的实测风速序列. (a) 2010年2月5日至2月24日, 时间间隔为60 min; (b) 2010年6月14日至6月20日, 时间间隔为10 min

Figure 1 Measured wind speed series of a wind turbine in the Donghai Bridge wind farm. (a) 60 min interval, from February 5<sup>th</sup>, 2010 to February 24<sup>th</sup>, 2010; (b) 10 min interval, from June 14<sup>th</sup>, 2010 to June 20<sup>th</sup>, 2010.

型的实际预报能力. 其中, 第一组60 min间隔的数据, 第一部分由2月5日至23日, 共456个数据组成; 第二部分由2月24日01:00至23:00, 共23个数据组成. 第二组10 min间隔的数据, 由6月14日00:00至20日11:50, 共936个数据组成; 第二部分由6月20日12:00至23:50, 共72个数据组成.

### 2.2 预报模型的选择

本文采用时间序列分析与条件异方差模型的组合模型, 描述风速时间序列的特性. 这里, 我们选用了时间序列分析中经典的ARMA模型以及四个具有代表性的组合模型(ARMA-GARCH, ARMA-GARCH-M, ARMA-EGARCH以及ARMA-EGARCH-M), 其一般表达式见表1.

ARMA模型中, 不存在条件方差方程; 而四个组合模型均由均值方程与条件方差方程组成. 在均值方程中,  $x_t$ 为目标分析序列;  $\delta$ 为常系数;  $\varepsilon_t$ 为白噪声序列, 且满足 $\varepsilon_t \sim N(0, h_t^2)$ ;  $\varphi_i$ 与 $\theta_j$ 分别为自回归系数与滑动平均系数. 在条件方差方程中,  $h_t^2$ 为白噪声的方差,  $\alpha_i, \beta_j, \gamma$ 与 $\lambda_k$ 为模型参数. 其中,  $\gamma$ 称为波动反馈效应系数,  $\lambda_k$ 称为非对称效应系数.

为波动反馈效应系数.

理论上, 由于条件方差方程描述了随机误差项的方差随时间的变化规律, 因此, 四个组合模型均可以描述波动集群效应. 而在ARMA-GARCH-M模型中, 均值方程中的 $\gamma h_t^2$ 这一项可以反映方差对均值的影响, 因此这一模型可以对波动的反馈效应加以描述. 在ARMA-EGARCH模型中, 当 $\lambda_k \neq 0$ 时, 残差 $\varepsilon_t$ 的正负对下一时刻的条件方差影响不同, 因此这一模型可以对波动的非对称效应加以描述. 现将各个模型对序列特性的描述整理见表2, 后文将结合实际算例, 详细阐述这些效应.

### 2.3 风速序列特性

本文分别采用ADF(Augmented Dickey Fuller)检验<sup>[14]</sup>与KPSS (Kwiatkowski Phillips Schmidt Shin)检验<sup>[15]</sup>对风速序列进行平稳性检验, 以加强检验结果的可靠性. 当统计值小于各个显著性水平下的临界值时, 则认为该序列为平稳序列. 表3为具体的检验结果, 结果表明, 两组实测风速序列在5%的显著性水平下, 均可以视为平稳序列.

**表1** 经典预报模型与四个组合预报模型的一般表达式  
**Table 1** General expression of classical forecasting model and four combined forecasting models

模型	均值方程	条件方差方程
ARMA	$x_t = \delta + \sum_{i=1}^p \varphi_i x_{t-i} + \varepsilon_t + \sum_{j=1}^q \theta_j \varepsilon_{t-j}$	-
ARMA-GARCH	$x_t = \delta + \sum_{i=1}^p \varphi_i x_{t-i} + \varepsilon_t + \sum_{j=1}^q \theta_j \varepsilon_{t-j}$	$h_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i \varepsilon_{t-i} + \sum_{j=1}^q \beta_j h_{t-j}^2$
ARMA-GARCH-M	$x_t = \delta + \sum_{i=1}^p \varphi_i x_{t-i} + \varepsilon_t + \sum_{j=1}^q \theta_j \varepsilon_{t-j} + \gamma h_t$	$h_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i \varepsilon_{t-i} + \sum_{j=1}^q \beta_j h_{t-j}^2$
ARMA-EGARCH	$x_t = \delta + \sum_{i=1}^p \varphi_i x_{t-i} + \varepsilon_t + \sum_{j=1}^q \theta_j \varepsilon_{t-j}$	$\log(h_t^2) = \alpha_0 + \sum_{i=1}^p \alpha_i \cdot \left  \frac{\varepsilon_{t-i}}{h_{t-i}} \right  + \sum_{k=1}^r \lambda_k \cdot \frac{\varepsilon_{t-k}}{h_{t-k}} + \sum_{j=1}^q \beta_j \log(h_{t-j}^2)$
ARMA-EGARCH-M	$x_t = \delta + \sum_{i=1}^p \varphi_i x_{t-i} + \varepsilon_t + \sum_{j=1}^q \theta_j \varepsilon_{t-j} + \gamma h_t$	$\log(h_t^2) = \alpha_0 + \sum_{i=1}^p \alpha_i \cdot \left  \frac{\varepsilon_{t-i}}{h_{t-i}} \right  + \sum_{k=1}^r \lambda_k \cdot \frac{\varepsilon_{t-k}}{h_{t-k}} + \sum_{j=1}^q \beta_j \log(h_{t-j}^2)$

**表2** 各个模型对风速序列特性的描述能力  
**Table 2** The ability of each model to describe the characteristics of wind speed series

模型	波动集群效应	波动反馈效应	非对称效应
ARMA	×	×	×
ARMA-GARCH	√	×	×
ARMA-GARCH-M	√	√	×
ARMA-EGARCH	√	×	√
ARMA-EGARCH-M	√	√	√

**表 3** ADF检验与KPSS检验各自统计值以及1%, 5%, 10%的显著性水平下的临界值  
**Table 3** ADF test and KPSS test and the critical value of 1%, 5%, 10% significance level

风速数据	检验方法	检验统计量	1%临界值	5%临界值	10%临界值
10 min间隔	ADF	-3.248	-3.438	-2.865	-2.569
10 min间隔	KPSS	0.443	0.739	0.463	0.347
60 min间隔	ADF	-4.323	-3.445	-2.868	-2.570
60 min间隔	KPSS	0.278	0.739	0.463	0.347

本文从低阶开始, 逐步增加回归阶数 $p$ 与平滑阶数 $q$ , 综合考虑AIC准则与BIC准则进行模型阶数及参数的预估, 最终得到风速预报模型. 两组风速序列模型参数估计的具体结果见表4与表5. 其中,  $\varphi_i$ 与 $\theta_j$ 为均值方程中自回归系数与滑动平均系数, 若缺少某些阶数的自回归项或滑动平均项, 则对应项系数用0表示; 在条件方差方程中, 由于不同模型包含的项各不相同, 故用“-”表示这一模型不存在对应项. 下面从各个效应出发, 对模型各项系数进行详细分析.

(1) 波动集群效应

以第二组10 min间隔的风速数据为例, 对本文选用的ARMA模型的残差进行LM (Lagrange Multiplier) 检验, 在滞后阶数 $q=15$ 时检验的 $P$ 值为0.042. 因此, 可在5%的显著性水平下拒绝随机误差项的方差为常数这一假设, 即认为风速序列存在波动集群效应. 直观起见, 我们做出其残差序列的平方图, 见图2.

由图2可知, 整体而言, 残差平方并不在一个范围内相对平稳的波动. 为更形象地描述这一现象, 我们借助波浪领域中波群因子(Groupiness)的概念<sup>[16]</sup>, 计算了残差序列的波能时间序列(SIWEH: Smoothed Instantaneous Wave Energy History)与波群因子, 具体计算公式如下:

$$GF = \sqrt{\frac{\frac{1}{T_R} \int_0^{T_R} [E_s(t) - \overline{E_s(t)}]^2 dt}{\overline{E_s}}}, \quad (1)$$

$$E_s(t) = \frac{1}{T_R} \int_{-\infty}^{\infty} \eta^2(t + \tau) W(\tau) d\tau, \quad (2)$$

$$W(\tau) = \begin{cases} 1 - \frac{\tau}{T_p}, & |\tau| < T_p, \\ 0, & |\tau| > T_p, \end{cases} \quad (3)$$

**表 4** 10 min时间间隔的五个短期风速预报模型的参数估计结果  
**Table 4** The results of parameter estimation of five short-term wind speed forecasting models at 10-minute intervals

模型参数	ARMA	ARMA-GARCH	ARMA-GARCH-M	ARMA-EGARCH	ARMA-EGARCH-M
$\delta$	4.530	3.482	1.200	2.666	4.087
$\varphi_1$	0.000	0.525	0.314	0.000	1.162
$\varphi_2$	1.109	0.510	0.709	0.875	-0.110
$\varphi_3$	0.000	0.000	0.000	0.000	-0.186
$\varphi_4$	-0.418	-0.188	-0.192	-0.240	0.000
$\varphi_5$	0.273	0.000	0.000	0.097	0.167
$\varphi_6$	-0.107	-0.249	-0.127	0.000	-0.097
$\varphi_7$	0.000	0.000	0.000	0.000	0.031
$\varphi_8$	0.000	0.373	0.212	0.000	0.000
$\varphi_{10}$	0.084	0.000	0.000	0.153	0.000
$\theta_1$	1.931	1.856	2.012	1.827	0.586
$\theta_2$	1.190	1.527	1.724	1.159	0.000
$\theta_3$	0.108	0.378	0.525	0.249	0.000
$\theta_4$	0.000	0.000	0.000	0.000	0.147
$\alpha_0$	-	0.008	0.004	-0.392	-0.767
$\alpha_1$	-	0.307	0.356	0.759	0.441
$\beta_1$	-	0.546	0.620	-0.406	-0.745
$\gamma$	-	-	0.085	-	0.074
$\lambda$	-	-	-	0.101	0.058

表5 60 min时间间隔的五个短期风速预报模型的参数估计结果

Table 5 The results of parameter estimation of five short-term wind speed forecasting models at 60-minute intervals

模型参数	ARMA	ARMA-GARCH	ARMA-GARCH-M	ARMA-EGARCH	ARMA-EGARCH-M
$\delta$	7.653	7.702	6.623	7.667	6.796
$\varphi_1$	0.000	-0.158	-0.147	0.000	0.585
$\varphi_2$	1.021	1.298	1.229	1.027	0.366
$\varphi_3$	-0.198	0.000	0.000	-0.184	0.000
$\varphi_4$	0.000	-0.433	-0.331	0.000	-0.094
$\varphi_5$	0.197	0.000	0.000	0.155	0.000
$\varphi_6$	-0.110	0.000	0.000	-0.119	0.000
$\varphi_7$	0.089	0.184	0.130	0.118	0.085
$\varphi_{10}$	-0.140	0.000	0.000	-0.130	-0.085
$\theta_1$	1.283	1.391	1.331	1.176	0.614
$\theta_3$	0.000	-0.312	-0.218	0.000	0.000
$\theta_4$	0.000	0.126	0.136	0.000	0.000
$\theta_5$	-0.362	0.000	0.000	-0.267	-0.058
$\alpha_0$	-	0.470	0.420	-0.110	-0.221
$\alpha_1$	-	0.161	0.145	0.281	0.411
$\beta_1$	-	0.498	0.546	0.537	0.637
$\gamma$	-	-	0.052	-	0.053
$\lambda$	-	-	-	0.081	0.074

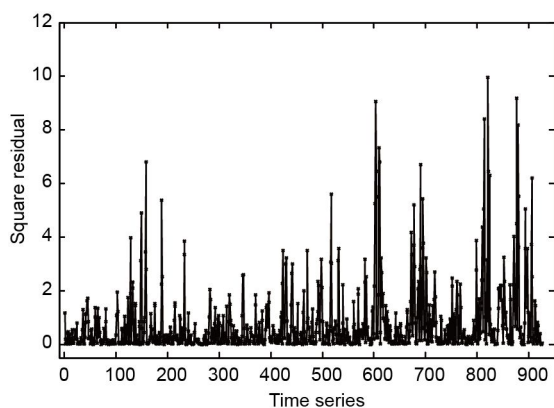


图2 ARMA模型残差序列平方. 横坐标表示6月14日00:00至6月20日11:50各个时刻, 此处用序列号表示; 纵坐标为风速序列各个时刻采用ARMA模型拟合值与实测值的残差平方

Figure 2 Residual squared of ARMA model. Horizontal axis represents the June 14<sup>th</sup> 00:00 to June 20<sup>th</sup> 11:50, here we use the serial number to describe; vertical axis represents the residual squared which is calculated by ARMA model for each time.

其中,  $E_s(t)$ 是谱窗平均(瞬时光滑)的波能时间序列,  $W(\tau)$ 是Bartle谱窗光滑函数,  $T_R$ 为时间序列记录长度,  $T_P$ 为谱峰周期.

波群因子可以定量地描述波能时间序列偏离平均波能的谱窗平均方差, 以反映各种成分出现的先后

次序<sup>[17]</sup>. 波群因子的数值越大, 表明波能时间序列偏离平均波能的程度越大, 波能时间序列显示出较大震荡; 反之, 波能时间序列的变化则较为平稳. 图3为10 min间隔的风速残差序列的SIWEH, 显然, 其幅值的波动在整体范围内并不均匀. 当显著高于均值的波动出现后, 往往伴随着另一些较大的波动; 当显著低于均值的波动出现后, 往往伴随着另一些较小的波动. 因此, 风速残差序列具有典型的波动集群现象. 我们分别计算了第一组60 min间隔与第二组10 min间隔的风速残差序列的波群因子, 前者为0.62, 后者为0.93. 说明时间间隔较短的实测风速序列的波动集群效应更为显著.

风的本质是大气流动. 作为自然界的一种流动现象, 受大气环境的影响, 对流场的扰动时刻发生着, 故风速大小必然会出现波动. 由于扰动的猝发具有随机性与间歇性, 故风速的波动也必然具有随机性与间歇性. 而风场的湍流强度直接决定了风速的波动强弱, 因此, 波动强弱随时间变化. 然而, ARMA模型则无法描述这一现象, 当选用ARMA模型描述风速序列时, 即假设风速的波动强弱与时间无关, 这是建模后风速的残差序列呈现波动集群现象的直接原因.

## (2) 波动反馈效应

ARMA-GARCH-M模型与ARMA-EGARCH-M模

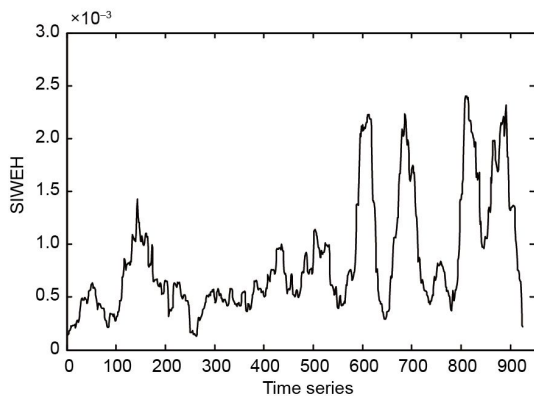


图3 ARMA模型残差序列的波能时间序列. 横坐标表示6月14日00:00至6月20日11:50各个时刻, 此处用序列号表示; 纵坐标风速残差序列的波能时间序列

Figure 3 SIWEH of residual sequence of ARMA model. Horizontal axis represents the June 14<sup>th</sup> 00:00 to June 20<sup>th</sup> 11:50, here we use the serial number to describe; vertical axis represents the SIWEH of residual sequence of ARMA model.

型中, 反馈效应系数 $\gamma$ 的数值大小表征了反馈效应的强弱. 对于本文所采用的算例, 60 min时间间隔的风速数据,  $\gamma$ 与模型中其他参数相比小一个数量级, 更接近于零. 表明这组风速序列数据的反馈效应并不显著, 即风速波动的强弱程度对平均风速的大小无显著影响. 而10 min时间间隔的风速数据, 两个组合模型中参数 $\gamma$ 均大于60 min的参数 $\gamma$ 的数值, 表明这组风速序列数据的反馈效应强于60 min时间间隔的风速序列.

为更好地分析风速序列的波动性, 我们分别计算了两组风速数据的峰度值. 峰度, 是表征统计数据分布形态陡缓程度的统计量. 数值越大, 则分布越陡峭, 即序列整体大部分取值接近平均值, 数据整体波动越弱; 相反地, 若数值越小, 则分布越平坦, 即序列整体大部分取值远离平均值, 数据整体波动越强. 对于两组风速数据, 10 min间隔的数据峰度值为2.47; 60 min时间间隔的数据峰度值为2.58, 均小于正态分布的3.0. 故与正态分布相比, 两组风速数据的峰度值较小, 较正

态分布而言, 波动性较强. 且10 min间隔的风速数据整体的波动水平高于60 min间隔的数据. 而波动反馈效应正是对风速序列随机波动项对平均风速的影响的描述, 某种意义上而言, 峰度值与反馈效应系数均可以描述序列的波动水平, 两者的含义具有一定的相似性.

其他学者也曾采用ARMA-GARCH-M模型对实测风速进行建模, 表6列出了其模型的反馈效应系数、风电场选址类型以及实测风速的时间间隔.

对比表6中的结果, 当风速数据时间间隔接近时, 不同的风电场所在地也会导致模型中反馈效应系数的数值大小出现差异. 我们认为, 表面粗糙度的不同可能是造成这一现象的主要原因. 具体而言, 湍流强度表征了风速随时间和空间变化情况, 是一个反应脉动风速相对强度的物理量, 受地表面的地貌特征的强烈影响. 已有研究指出<sup>[19]</sup>, 湍流强度随表面粗糙长度的增加而增大. 由于本文的风速数据来源于海上风电场的实际测量, 与陆地相比, 海洋的表面粗糙长度远小于陆地的表面粗糙长度<sup>[20]</sup>, 所以其湍流强度较小. 而湍流强度越小则风速的波动越弱, 故对于本文的算例, 风速序列的反馈效应并不显著.

当风电场所在地的类型相同时, 风速数据时间间隔的不同也会导致模型中反馈效应系数的数值出现差异. 风速的波动是由流场中的湍流引起的, 其时间尺度一般为秒或者分钟量级. 所以, 当实测风速数据的时间间隔越长, 则包含风速波动的信息越少. 在模型中, 即表现为反馈效应项的系数更趋于零, 统计参数峰度值更小. 综上所述, 我们认为风电场所在地与采集数据的时间间隔是决定风速序列反馈效应强弱的两大关键因素.

### (3) 非对称效应

表4与表5中, 两组风速数据的建模结果表明, 风速预报模型中的非对称项系数 $\lambda$ 均不为零, 故东海大桥风电场的短期风速序列存在非对称效应, 即随机扰动

表6 ARMA-GARCH-M模型不同算例中的应用

Table 6 Application of ARMA-GARCH-M model in different case

算例	反馈效应系数	风电场所在地	数据时间间隔(min)
Ewing et al. <sup>[18]</sup>	0.182	陆上	15
Liu et al. <sup>[12]</sup>	0.080	陆上	60
本文	0.085	海上	10
本文	0.052	海上	60

项的正负对下一时刻风速波动的强弱影响不同. 具体而言, 结合表1中ARMA-EGRACH模型条件方差方程的表达式, 由于本文风速模型中非对称项系数为正值, 故在当前时刻波动大小相同的情况下(即 $|\varepsilon_{t-1}|/|h_{t-1}|$ 为某一定值), 未来条件方差( $h_{t+\tau}^2$ )在正波动下( $\varepsilon_t > 0$ )的增幅大于在负波动下( $\varepsilon_t < 0$ )的增幅. 换言之, 使风速增大的脉动对下一时刻可能引发的风速波动大于使风速减弱脉动的情况. 风速的波动由湍流引起, 因此, 风速序列的非对称效应表明湍流对流场的扰动是非对称的.

综上所述, 对于东海大桥风电场而言, 短期风速序列具有波动集群效应以及非对称效应, 而反馈效应并不显著. 就预报模型而言, ARMA-EGARCH-M模型可以同时对上述三个风速序列特性加以定量描述, 其余预报模型均无法实现, 故仅从理论上而言, 这一模型的预报能力可能优于其他预报模型.

### 3 短期风速预报

前文对适用于东海大桥风电场的短期风速预报模型进行了定性的判定. 为定量地分析能够刻画风速序列各个特性的ARMA-EGARCH-M模型的预报能力, 并与其他模型的进行对比, 我们结合具体算例, 分别计算平均绝对误差(MAE: Mean Absolute Error)、平均绝对百分误差(MAPE: Mean Absolute Percentage Error)与均方根误差(RMSE: Root Mean Square Error)以度量模型的预报精度. 具体计算公式如下:

$$\text{MAE} = \frac{1}{N} \sum_{t=1}^N |x_t - \hat{x}_t|, \quad (4)$$

$$\text{MAPE} = \frac{1}{N} \sum_{t=1}^N \left| \frac{x_t - \hat{x}_t}{x_t} \right| \times 100\%, \quad (5)$$

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{t=1}^N (x_t - \hat{x}_t)^2}, \quad (6)$$

其中,  $x_t$ 为风速实测值,  $\hat{x}_t$ 为风速预报值,  $N$ 为预报个数.

#### 3.1 单步预报结果分析

所谓单步预报, 即在预报的过程中, 把每一时刻的风速实测值更新为历史数据, 代替当前时刻的预报值而进行下一时刻风速的预报, 以减少多次迭代计算产生的累计误差. 这里以10 min间隔的风速数据为例, 使用前文已建立的五个预报模型, 对6月20日12:00至23:50的风速进行预报, 并计算各个模型的MAE, MAPE与RMSE, 误差结果见表7.

对比表7中各个模型的误差, 我们可以得到如下结论: 1)与任一组合模型的预报结果相比, ARMA模型的三个误差指标均最大. 这表明组合模型通过条件方差方程描述风速序列的波动集群效应, 预报精度得到了显著提升. 2)对比ARMA-GARCH与ARMA-EGARCH模型的预报结果: 在考虑波动集群效应的基础上, 修改条件方差方程, 增加非对称效应项, 能够进一步提高模型预报精度. 因此, 能够刻画风速序列的波动集聚效应与非对称效应的模型比忽略这两个效应的模型有着更高的预报精度. 3)对比ARMA-GARCH与ARMA-GARCH-M模型的预报结果以及ARMA-EGARCH与ARMA-EGARCH-M模型的预报结果: 在均值方程中添加波动反馈效应项与不添加此项, 两者的预报结果十分接近. 因此, 就提升模型预报精度而言, 波动反馈效应项仅提供极其有限的帮助. 预报结果再次表明对于东海大桥10 min间隔的风速序列, 其波动反馈效应并不显著. 4)整体而言, 在组合模型中, ARMA-EGARCH与ARMA-EGARCH-M模型的各项误差指标均小于其他模型. 综上所述, 对于本文的算例, 就单步预报而言, 较ARMA-EGARCH-M模型而言, ARMA-EGARCH模型的预报能力同样出色. 这里, 我们给出了ARMA-EGARCH-M模型的各个时刻的预报值, 见图4. 显然, 这一模型的预报结果与实测风速及其变化趋势较为吻合.

表7 6月20日12:00至23:50经典模型与组合模型的预报误差

Table 7 Prediction error of classical model and combined model from June 20<sup>th</sup> 12:00 to 23:50

模型	MAE (m/s)	MAP (%)	RMSE (m/s)
ARMA	0.70	13.80	0.96
ARMA-GARCH	0.52	9.98	0.65
ARMA-GARCH-M	0.51	9.86	0.63
AEMA-EGARCH	0.40	7.61	0.51
AEMA-EGARCH-M	0.40	7.56	0.53

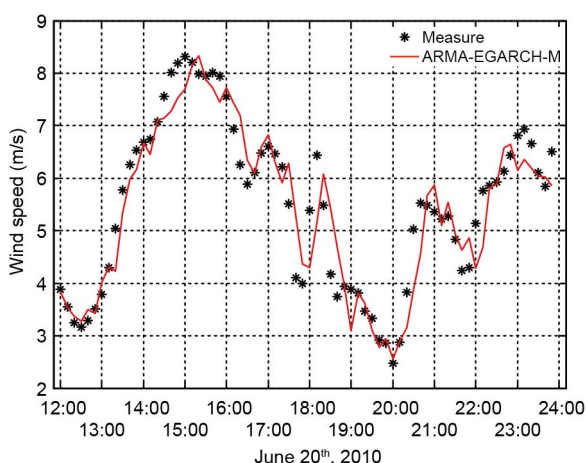


图4 (网络版彩图)ARMA-EGARCH-M模型在6月20日12:00至23:50内的预报值与实测风速.散点为实测值;折线为模型的预报值

Figure 4 (Color online) The ARMA-EGARCH-M model was used to predict the wind speed from 12:00 to 23:50 on June 20<sup>th</sup>. The scattered points are the measured wind speed and the broken line is the predicted value of the model.

### 3.2 多步预报结果分析

单步预报的提前时间即为实测风速采集间隔的时长. 而对于实际工程而言, 更多的预报提前量意味着电网调度具备更充分的准备时间<sup>[21]</sup>. 因此, 除了分析预报模型的单步预报能力, 对模型多步预报能力的研究也具有突出的工程意义. 这里, 记风速实测序列为 $\{x_t\}$ , 风速预报序列为 $\{\hat{x}_t\}$ . 当 $x_1, x_2, \dots, x_t$ 确定后, 求解 $\hat{x}_{t+1}, \hat{x}_{t+2}, \dots, \hat{x}_{t+n}$ 的过程, 称为在提前 $n$ 步的要求下进行的风速预报, 函数表达式如下:

$$\begin{cases} \hat{x}_{t+1} = f_k(x_{t-k}, x_{t+1-k}, x_{t+2-k}, \dots, x_{t-1}, x_t), \\ \hat{x}_{t+2} = f_k(x_{t+1-k}, x_{t+2-k}, x_{t+3-k}, \dots, x_t, \hat{x}_{t+1}), \\ \dots \\ \hat{x}_{t+n} = f_k(x_{t+n-1-k}, x_{t+n-k}, x_{t+n+1-k}, \dots, \hat{x}_{t+n-2}, \hat{x}_{t+n-1}), \end{cases} \quad (7)$$

这里, 我们分别在提前1步、3步、5步与10步的要求下, 采用ARMA-EGARCH-M模型对6月20日12:00至23:50的风速进行预报, 并将结果与实测风速进行对比, 结果如图5所示.

图5表明, 对于ARMA-EGARCH-M模型而言, 随着提前预报步数的增加, 20日12:00至23:50的风速预报值的波动将增大, 整体误差水平上升. 由于无法将下一时刻的预估值及时替换为最新的实测值, 故多步

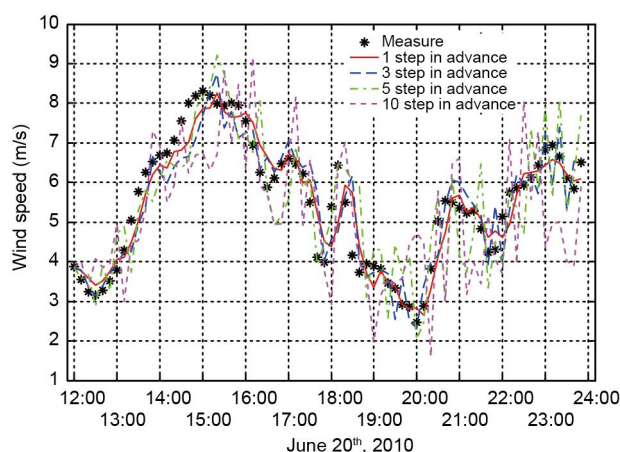


图5 (网络版彩图)ARMA-EGARCH-M模型在提前1步、3步、5步与10步的要求下对6月20日12:00至23:50的风速预报. 散点为实测值;折线为模型在各个提前步数情况下的预报值

Figure 5 (Color online) The ARMA-EGARCH-M model was used to predict the wind speed from 12:00 to 23:50 on June 20<sup>th</sup> at the request of 1 step, 3 steps, 5 steps and 10 steps in advance. The scattered points are the measured wind speed and the broken line is the predicted value of the model in each step in advance.

预报中的误差将逐步累计, 这是造成这一结果最主要的原因.

为进一步定量地分析这一模型在不同提前预报步数下的预报误差, 并与ARMA, ARMA-GARCH, ARMA-GARCH-M以及ARMA-EGARCH模型进行对比, 我们计算并绘制了五个模型从提前1步至10步的MAPE与RMSE, 结果见图6.

对比图6可以发现: 1)对于任意一个确定的模型, 当提前预报步数增加时, 整体而言, 预报误差均增大. 2)在相同的提前预报步数的情况下, 五个模型的预报精确度不同. 与经典ARMA模型相比, 组合预报模型的整体预报误差更小. 其中, ARMA-EGARCH模型与ARMA-EGARCH-M模型预报精度较为接近, 且均优于其他模型. 3)随着提前预报步数增加, 五个模型的预报误差增长率不同. 以均方根误差为例, ARMA模型提前10步预报的RMSE为1.53 m/s, 较单步预报的RMSE增长了0.59 m/s; 而ARMA-EGARCH-M模型提前10步预报的RMSE为0.99 m/s, 较单步预报的RMSE仅增长了0.45 m/s. 综上所述, 对于东海大桥风电场而言, ARMA-EGARCH模型与ARMA-EGARCH-M模型都具有较为出色的多步预报能力.



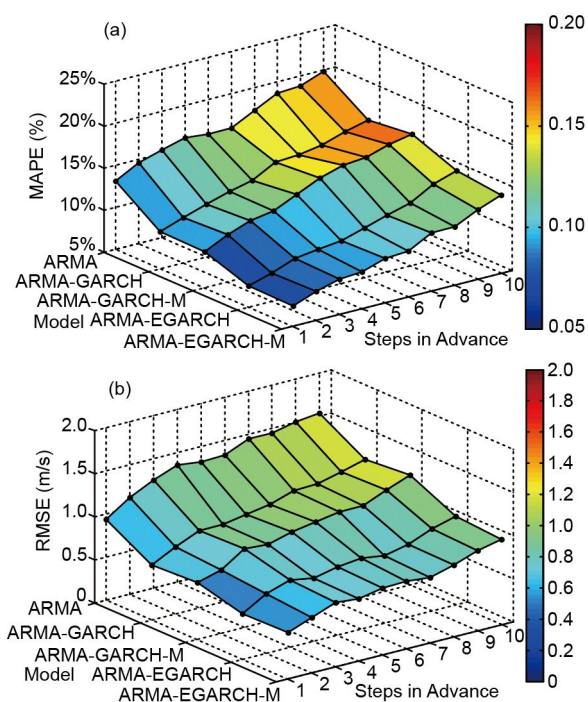


图6 (网络版彩图)各个模型在提前1步至10步的情况下, 26日12:00至23:50风速预报值与实测值的误差. (a) MAPE; (b) RMSE

Figure 6 (Color online) In the case of 1 to 10 steps ahead, the error between the wind speed prediction value and the measured value for each model, from 12:00 to 23:50 on the 26<sup>th</sup> day. (a) MAPE; (b) RMSE.

**致谢** 作者对上海勘测设计研究院在国家自然科学基金研究中的合作, 和聂冰川博士、王晓亮博士的有益讨论表示衷心感谢.

## 参考文献

- 1 Kavasseri R G, Seetharaman K. Day-ahead wind speed forecasting using f-ARIMA models. *Renew Energy*, 2009, 34: 1388–1393
- 2 Tascikaraoglu A, Uzunoglu M. A review of combined approaches for prediction of short-term wind speed and power. *Renew Sustain Energy Rev*, 2014, 34: 243–254
- 3 Lei M, Shiyao L, Chuanwen J, et al. A review on the forecasting of wind speed and generated power. *Renew Sustain Energy Rev*, 2009, 13: 915–920
- 4 安鸿志. 时间序列分析. 上海: 华东师范大学出版社, 1989
- 5 Engle R F. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 1982, 50: 987
- 6 Bollerslev T. Generalized autoregressive conditional heteroskedasticity. *J Econom*, 1986, 31: 307–327
- 7 Engle R F, Lilien D M, Robins R P. Estimating time varying risk premia in the term structure: The Arch-M model. *Econometrica*, 1987, 55: 391
- 8 Nelson D B. Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, 1991, 59: 347
- 9 Harvey A, Ruiz E, Sentana E. Unobserved component time series models with Arch disturbances. *J Econom*, 1992, 52: 129–157
- 10 Kamal L, Jafri Y Z. Time series models to simulate and forecast hourly averaged wind speed in Quetta, Pakistan. *Solar Energy*, 1997, 61: 23–32
- 11 Cadenas E, Rivera W. Wind speed forecasting in the South Coast of Oaxaca, México. *Renew Energy*, 2007, 32: 2116–2128
- 12 Liu H, Erdem E, Shi J. Comprehensive evaluation of ARMA-GARCH(-M) approaches for modeling the mean and volatility of wind speed. *Appl*

- [Energy](#), 2011, 88: 724–732
- 13 Wang J, Xiong S. A hybrid forecasting model based on outlier detection and fuzzy time series-A case study on Hainan wind farm of China. [Energy](#), 2014, 76: 526–541
- 14 Dickey D A, Fuller W A. Likelihood ratio statistics for autoregressive time series with a unit root. [Econometrica](#), 1981, 49: 1057
- 15 Kwiatkowski D, Phillips P C B, Schmidt P, et al. Testing the null hypothesis of stationarity against the alternative of a unit root. [J Econom](#), 1992, 54: 159–178
- 16 Kao J C, Lin J M. Computer simulation of ocean wave grouping. In: Proceedings of the 7th Congress of APRD-IAHR. Beijing: Institute of Water Conservancy and Hydroelectric Power Research, 1990. 339–343
- 17 Huang M C. Wave parameters and functions in wavelet analysis. [Ocean Eng](#), 2004, 31: 111–125
- 18 Ewing B T, Kruse J B, Schroeder J L. Time series analysis of wind speed with time-varying turbulence. [Environmetrics](#), 2006, 17: 119–127
- 19 贺德馨. 风工程与工业空气动力学. 北京: 国防工业出版社, 2006
- 20 Bañuelos-Ruedas F, Angeles-Camacho C, Rios-Marcuello S. Analysis and validation of the methodology used in the extrapolation of wind speed data at different heights. [Renew Sustain Energy Rev](#), 2010, 14: 2383–2391
- 21 Lydia M, Suresh K S, Immanuel S A, et al. Linear and non-linear autoregressive models for short-term wind speed forecasting. [Energy Conv Manag](#), 2016, 112: 115–124

## Characteristics and forecast of short-term wind speed series in the Donghai Bridge wind farm

LYU GuoQin<sup>1,2</sup>, ZHANG HuiQin<sup>1,2\*</sup> & LI JiaChun<sup>1,2</sup>

<sup>1</sup> Key Laboratory for Mechanics in Fluid Solid Coupling Systems, Institute of Mechanics, Chinese Academy of Sciences, Beijing 100190, China;

<sup>2</sup> School of Engineering Science, University of Chinese Academy of Sciences, Beijing 100049, China

As a sort of worldwide renewable energy, wind energy can reduce environmental pollution and relieve the energy shortage. In order to reduce the adverse effect with the integration of wind energy into electricity grids and the operating cost of power supply system, it is becoming increasingly significant to acquire accurate short-term wind speed forecasts. In this paper, based on the analysis of the measured wind speed data in the Donghai Bridge wind farm, we suggest that the short-term wind speed series has volatility clustering effect and asymmetric effect, and the volatility feed-back effect is not significant. And then the possible causes for this phenomenon are elucidated in detail from the viewpoint of physics. In addition, in order to select the forecast model which is appropriate to the Donghai Bridge wind farm, we use these indexes of error: mean absolute error (MAE), mean absolute percentage error (MAPE) and root mean square error (RMSE) as the criterion, and compare the prediction accuracy of the five models by quantitative analysis. The results show that ARMA-EGARCH model and ARMA-EGARCH-M model are very close to each other in both single-step and multi-step forecasting, and they are superior to other models. What's more, with the increase of the number of advance forecasting steps, error growth rate of these two models is low.

**offshore wind farm, short-term wind speed forecast, time series analysis, volatility clustering effect, volatility feed-back effect, asymmetric effect**

**PACS:** 02.50.Fz, 89.30.Ee, 89.20.Kk

**doi:** [10.1360/SSPMA2016-00407](https://doi.org/10.1360/SSPMA2016-00407)