

王同亮,马绍休,高扬,等.小波包分解与多个机器学习模型耦合在风速预报中的对比[J].中国沙漠,2021,41(2):38-50.

小波包分解与多个机器学习模型耦合 在风速预报中的对比

王同亮^{1a,2},马绍休^{1a},高扬^{1a,2},宫毓来^{1a,2},安志山^{1ab}

(1.中国科学院西北生态环境资源研究院 a.沙漠与沙漠化重点实验室, b.公共技术服务中心,甘肃 兰州 730000;
2.中国科学院大学,北京 100049)

摘要:准确预报风速是提高风电利用率以及电力系统稳定性的有效方法。学者们提出了大量风速预报模型,但针对不同下垫面不同风速预报模型的对比研究较少。该研究主要探究小波包分解和12个机器学习模型耦合对3种下垫面(戈壁、绿洲和沙漠)风速预报能力,探索风速预报的优化耦合模型。设置3组模型实验进行对比:单一机器学习模型、小波包分解-机器学习混合模型和小波包分解-机器学习-卷积神经网络混合模型。结果表明:具有特征选择和记忆功能的深度学习模型(如卷积长短时记忆网络)以及极限学习机对风速具有较好的预报能力,小波包分解可以显著提高模型精度。小波包分解与卷积长短时记忆网络、卷积门控循环单元和极限学习机的耦合模型在风速预报中具有较好的表现。这表明信号分解和深度学习的耦合模型,能有效提高预报精度,值得推广。

关键词:风速预报;小波包分解;深度学习模型;干旱半干旱区;耦合模型

文章编号: 1000-694X(2021)02-038-13

DOI: 10.7522/j.issn.1000-694X.2020.00124

中图分类号: P425.63

文献标志码: A

0 引言

随着能源需求的增加,清洁且丰富的风能得到广泛关注^[1]。但是风具有间歇性和变异性,导致风电呈现出的不稳定性^[2]。风电的不稳定性给风能管理和电力系统稳定运行提出了巨大挑战,这是造成风电并网困难、利用率低的主要原因之一^[3]。风电规划和风电预测预报能有效减小这一问题影响^[4],准确的风速预报是提高风电利用率的最有效方式之一^[5]。

国内外学者提出了多种模型进行风速预测预报,主要分为物理模型和统计模型。物理模型基于物质、能量平衡和空气动力学机理,考虑温度、气压、地形等物理参数进行大空间、长时间风速预测,代表方法有数值天气预报等^[6]。但物理过程复杂,物理模型所需参数多,且需要长期稳定的气象资料作为驱动数据^[7],不适用于风电场所需要的短期小区域风速预报^[8]。统计模型主要使用历史风速,通过建立数据间的映射关系实现风速预报。统计模型原理简

单,计算量小,适应范围广,具有强大的非线性映射、泛化和自我学习能力。统计模型包括自回归时间序列模型,支持向量机、人工神经网络、极限学习机、卷积神经网络、长短时记忆网络等机器及深度学习模型。这些方法已经被应用于风速预报^[9-14]。

但单独使用上述模型时预报精度并不理想,需要结合信号分解方法形成混合模型以提高预报精度^[15]。信号分解方法包括小波分解(Wavelet Decomposition, WD)、小波包分解(Wavelet Packet Decomposition, WPD)、经验模式分解(Empirical Mode Decomposition, EMD)及其改进方法等^[16]。小波分解可将复杂的原始信号分解为高频信号和低频信号,但该方法仅对信号低频部分详细分解,忽略了信号的高频部分,所以小波分解不适用于包含大量高频信息的信号(如风速)^[17];经验模式分解及其改进方法虽然计算简单,但数学解释不明确^[18],且有残余子信号(噪声);小波包分解数学解释明确,能够自行设定分解层数和分解使用的小波函数,已被

收稿日期:2020-10-08; 改回日期:2020-12-04

资助项目:中国科学院“百人计划”项目(Y729G01001);国家重点研发计划项目(2017YFE0119100)

作者简介:王同亮(1995—),男,河南驻马店人,硕士研究生,主要从事深度学习模型在风速预报中的应用研究。

E-mail: wangtongliang@niger.ac.cn

通信作者:马绍休(E-mail: shaoxiuma586@163.com)

广泛应用在风速预报中^[19-21]。

目前,大量研究通过改进已有模型^[1]、应用新方法^[22]以及耦合不同信号分解方法^[23]等来提高模型预报精度。随着深度学习方法快速发展,在混合模型中加入深度学习模型进行风速预报的研究日益增多。但这些模型预报机理差异大,多种单一模型或与信号分解方法耦合形成的混合模型之间的差异性对比研究不足。

鉴于此,本研究通过对比多种统计模型以及这些模型与小波包分解耦合模型在预报3种下垫面风速的表现,探索较好的风速预报模型。研究结果可为不同下垫面风速预报模型的选择与优化提供参考,为风电场准确预报风速,保证稳定供电提供理论支持。

1 研究方法 with 数据

1.1 基础预报模型

基于最新文献研究,本研究选取广泛应用的12种机器学习和深度学习模型。

1.1.1 支持向量回归

支持向量回归属于支持向量机,主要应用于回归问题。其核心思想是寻找一条最佳拟合曲线,在容忍偏差范围内,曲线包含尽可能多的数据点。回归方程的数学描述为^[24]:

$$f(x) = w^T \varphi(x) + b \quad (1)$$

式中: φ 为非线性映射函数; w 为权重; b 为偏差。

支持向量回归的优化问题描述为^[24]:

$$\min_{w, b, \xi, \xi^*} \frac{1}{2} w^T w + C \sum_{i=1}^n (\xi_i + \xi_i^*) \quad (2)$$

$$\text{s.t.} \begin{cases} y_i - (w, x_i + b) \leq \varepsilon + \xi_i \\ (w, x_i + b) - y_i \leq \varepsilon + \xi_i^* \quad (i = 1, 2, \dots, l) \\ \xi_i, \xi_i^* \geq 0 \end{cases} \quad (3)$$

式中: l 为样本数; x_i 和 y_i 分别为训练集的输入和输出数据; n 为总样本数; ξ 和 ξ^* 分别为训练误差上界和下界; ε 为不敏感损失因子; C 为正则常量。

支持向量回归的预报可以定义为^[24]:

$$f(x, a_i, a_i^*) = \sum_{i=1}^n (a_i - a_i^*) K(x, x_i) + b \quad (4)$$

式中: $f(x, a_i, a_i^*)$ 为预报值; a 为拉格朗日乘数; a^* 为拉格朗日乘数; $K(x, x_i)$ 为核函数。

1.1.2 梯度提升回归

梯度提升回归是一种集成学习算法^[25],该算法基于分类回归树(Classification and regression tree,

CART)和序列提升算法构建。算法目的是最小化损失函数,减小残差。为实现此目的,首先计算初始模型残差,在梯度下降方向使用残差拟合基础学习器,再使用拟合后的基础学习器更新模型,不断重复上述过程获得目标模型。

1.1.3 随机森林

随机森林也是一种基于分类回归树构建的集成学习算法,它基于装袋算法实现。随机森林在训练时,先从原始数据中有放回、平均地抽取多个子数据集,在子数据集上拟合分类回归树,最终结果由所有树结果的平均值决定,所有分类回归树构成随机森林模型。装袋算法和分类回归树构建过程中都存在随机性,可有效避免随机森林过拟合。

1.1.4 人工神经网络

人工神经网络受人类大脑结构启发得到,由一系列相互连接的神经元构成。网络内部不同神经元之间通过权重相连,神经元内设有激活函数和偏差。根据激活函数和学习算法的不同,人工神经网络可以分为不同类别。其数学描述为^[25]:

$$y_i^{(n)} = f \left(\sum_{j=1}^N w_{ij}^{(n, n-1)} y_j^{n-1} + \theta_i^n \right) \quad (5)$$

式中: i 为第 n 层的神经元; w_{ij} 为第 $n-1$ 层神经元 j 和第 n 层神经元 i 之间的权重; θ_i^n 为第 n 层神经元 i 的偏差。

1.1.5 极限学习机

极限学习机是一类包含单个隐藏层的前馈神经网络^[26],该算法的最大特点是保证学习精度的前提下训练速度更快。极限学习机在训练模型前,首先随机初始化输入权重和偏差,同时确定输出矩阵。不同于传统神经网络,极限学习机在迭代时不需要更新全部参数,仅通过输出矩阵求解相应方程组就能快速训练模型。其损失函数描述为^[27]:

$$E = \sum_{j=1}^N \left(\sum_{i=1}^L \beta_i g(w_i x_j + b_i) - t_j \right)^2 \quad (6)$$

式中: N 为输入神经元数量; L 为隐藏层神经元数量; β_i 为隐藏层和输出层之间权重; $g(\cdot)$ 为激活函数; w_i 为输入层和隐藏层之间权重; x_j 为输入数据; b_i 为偏差; t_j 为输出数据。

1.1.6 卷积神经网络

卷积神经网络能够提取数据特征,在处理时间序列数据方面有良好表现。相较于支持向量机和

人工神经网络等浅层模型,卷积神经网络模型特有的卷积层可以实现自动特征选取,其全连接层处理有短期依赖的高频子集能力强^[28]。卷积层内,各个神经元间无相互联系,只共享卷积核权重。相较于人工神经网络,该模型训练更有效。每个卷积层的数学描述为^[29]:

$$h_{ij}^k = f((w^k x)_{ij} + b_k) \quad (7)$$

式中: f 为激活函数; w^k 为卷积核连接到第 k 个特征图的权重。

1.1.7 循环神经网络

循环神经网络是一种处理时间序列数据的方法,但传统循环神经网络训练过程中常出现梯度消失或梯度爆炸问题。为解决此问题,提出了长短时记忆网络^[30],该方法能有效处理长短期依赖问题。长短时记忆网络的核心组件为输入门、遗忘门和输出门。输入门决定细胞状态中存储的信息,遗忘门决定被遗忘的信息,输出门决定最终输出。单个神经元计算过程为:①输入数据,输入门激活,计算输入细胞信息;②如果遗忘门激活,遗忘部分细胞中保留的历史状态;③输出门激活,计算输出信息。神经元内计算过程可以描述为^[28]:

$$i_t = \sigma(w_{ix}x_t + w_{im}m_{t-1} + w_{ic}c_{t-1} + b_i) \quad (8)$$

$$f_t = \sigma(w_{fx}x_t + w_{fm}m_{t-1} + w_{fc}c_{t-1} + b_f) \quad (9)$$

$$c_t = f \circ c_{t-1} + i_t \circ \tanh(w_{cx}x_t + w_{cm}m_{t-1} + b_c) \quad (10)$$

$$o_t = \sigma(w_{ox}x_t + w_{om}m_{t-1} + w_{oc}c_{t-1} + b_o) \quad (11)$$

$$m_t = o_t \circ \tanh(c_t) \quad (12)$$

$$y_t = w_{ym}m_t + b_y \quad (13)$$

式中: $x=(x_1, x_2, x_3, \dots, x_r)$ 为展开层数据; i_t 为输入门; f_t 为遗忘门; c_t 为记忆细胞的激活矢量; o_t 为输出门; m_t 为每个记忆块的激活矢量; w 为权重矩阵; b 为偏差矢量; \circ 为点积; σ 和 \tanh 为激活函数,激活函数数学描述为^[28]:

$$\sigma(x) = 1/(1 + e^{-x}) \quad (14)$$

$$\tanh(x) = (e^x - e^{-x})/(e^x + e^{-x}) \quad (15)$$

在长短时记忆网络发展中出现了许多变种模型,如双向长短时记忆网络、门控循环单元等。其中,双向长短时记忆网络利用历史信息和未来信息进行预报,门控循环单元减少了门的数量,模型原理与长短时记忆网络相似。此外,还有堆叠式长短时记忆网络、卷积长短时记忆网络和卷积门控循环单元等模型。

1.2 小波包分解

小波包分解源于小波分解,它在分解信号低频子集同时,对高频子集继续分解。因此,该方法适用于分解含有大量高频信息的风速信号,以获取更加准确的信息。小波包分解使用时仅需要设置小波函数和分解层数。根据分解的信号类别,小波包分解可分为离散小波转换和连续小波转换。由于风速的连续性,本研究使用连续小波转换,数学描述为^[31]:

$$CWT(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t) \Psi^*\left(\frac{t-b}{a}\right) dt \quad (16)$$

式中: $x(t)$ 为原始信号; Ψ 为母小波函数; a 为比例因子; b 为转换参数; $*$ 为共轭复数; t 为原始信号对应的时间点。

1.3 信号分解与基础预报模型耦合设计

根据已有文献,常用混合模型大多为信号分解方法和基础预报模型组成。由于信号分解子集各有特点,一些混合模型仅使用一种基础预报模型预报所有子集^[30],另一些模型针对不同频率子集使用不同模型进行预报^[28]。本实验方案为:①选择应用广泛的机器学习和最新的深度学习模型共12个作为基础预报模型(简称单一XXX模型,表1),研究各种单一模型在风速预报中的表现;②小波包分解与12个基础预报模型耦合(简称WPD-XXX混合模

表1 基础预报模型

Table 1 The name of the basic forecast model

模型	英文名称(缩写)
支持向量回归	Support Vector Regression(SVR)
梯度提升回归	Gradient Boosting Regression(GBR)
随机森林	Random Forest(RF)
人工神经网络	Artificial Neural Network(ANN)
极限学习机	Extreme Learning Machine(ELM)
卷积神经网络	Convolutional Neural Network(CNN)
长短时记忆网络	Long Short-term Memory(LSTM)
堆叠式长短时记忆网络	Stacked Long Short-term Memory(SLSTM)
双向长短时记忆网络	Bidirectional Long Short-term Memory (BLSTM)
门控循环单元	Gated Recurrent Unit(GRU)
卷积长短时记忆网络	Convolutional Long Short-term Memory (CLSTM)
卷积门控循环单元	Convolutional Gated Recurrent Unit(CGRU)

型),探索小波包分解在提高模型预报精度中的作用;③小波包分解与 12 个基础预报模型和卷积神经网络(CNN)耦合(简称 WPD-XXX-CNN 混合模型),这里 CNN 用于预报小波包分解子集中的高频子集,而 12 个单一 XXX 模型用于预报低频子集。单一 XXX 模型作为对照组,WPD-XXX 混合模型和 WPD-XXX-CNN 混合模型为实验组。单一 XXX 模型对照组中,使用 12 个基础预报模型直接预报风速;WPD-XXX 实验组中,先将风速分解为低频和高频子集,再由 XXX 模型分别预报所有子集,最后将子集预报结果相加得出最终预报结果;WPD-XXX-CNN 实验组中,利用卷积神经网络模型预报高频子集,XXX 模型预报低频子集,最后将子集预报结果相加,得到最终结果。算法流程见图 1。

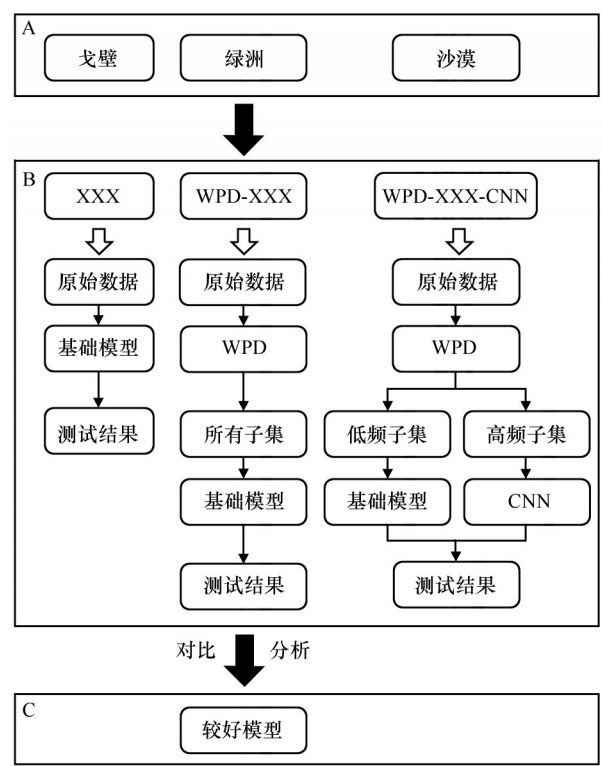


图1 算法流程图(A为3种下垫面,B为3种模型结构,C为较好模型)

Fig.1 Algorithm flowchart (A indicates the dataset, B indicates three kinds of models, C indicates the optimized models)

1.4 数据应用和评价指标

1.4.1 原始数据与重构

数据源于中国西北地区 3 个野外试验观测点,下垫面分别为戈壁、绿洲和沙漠(表 2)。风速时间序列见图 2。

基于现有研究的数据组织方式^[32],本研究风速数据重构过程如图 3。选择 6 个历史数据作为输入数据,数据重构时,每次选取 7 个数据为一组样本,平移 1 个位置选取下一组样本。因此,前后两组样本数据中有 6 个相同数据。最终得到 694 组样本数据,并以 594:100 的比例划分为训练集和测试集。数据重构前进行最大最小值标准化,公式为:

$$x_2 = (x_1 - x_{\min}) / (x_{\max} - x_{\min}) \tag{17}$$

式中: x_1 和 x_2 分别为标准化前和标准化后风速; x_{\max} 和 x_{\min} 分别为标准化前最大和最小风速。

1.4.2 精度指标

为了对比分析模型预报精度,本研究使用 3 种精度指标:平均绝对百分比误差 (Mean Absolute Percentage Error, MAPE)、平均绝对误差 (Mean Absolute Error, MAE) 和 均方根误差 (Root Mean Square Error, RMSE)^[33]。

$$MAPE = \left(\sum_{t=1}^n \left| \frac{x(t) - \hat{x}(t)}{x(t)} \right| \right) / N \tag{18}$$

$$MAE = \left(\sum_{t=1}^n |x(t) - \hat{x}(t)| \right) / N \tag{19}$$

$$RMSE = \sqrt{\left(\sum_{t=1}^n (x(t) - \hat{x}(t))^2 \right) / N} \tag{20}$$

式中: $x(t)$ 为原始风速; $\hat{x}(t)$ 为预报风速; n 为样本数。

2 结果

为了对比分析信号分解方法与不同基础预报模型耦合之间的差异,首先将原始风速数据分解为 1 个低频子集(图 4 A)和 3 个高频子集(图 4 B—D)。在对模型进行对比分析时,加入多模型平均方法,

表 2 数据描述

Table 2 Data description

观测点	时间范围	时间分辨率/min	高度/m
戈壁(40.5593°N,93.1117°E)	2019年4月21日20:40—26日17:10	10	10
绿洲(40.2581°N,94.7551°E)	2018年12月1日21:20—6日17:50	10	10
沙漠(40.0082°N,94.5940°E)	2019年5月21日00:40—25日21:10	10	10

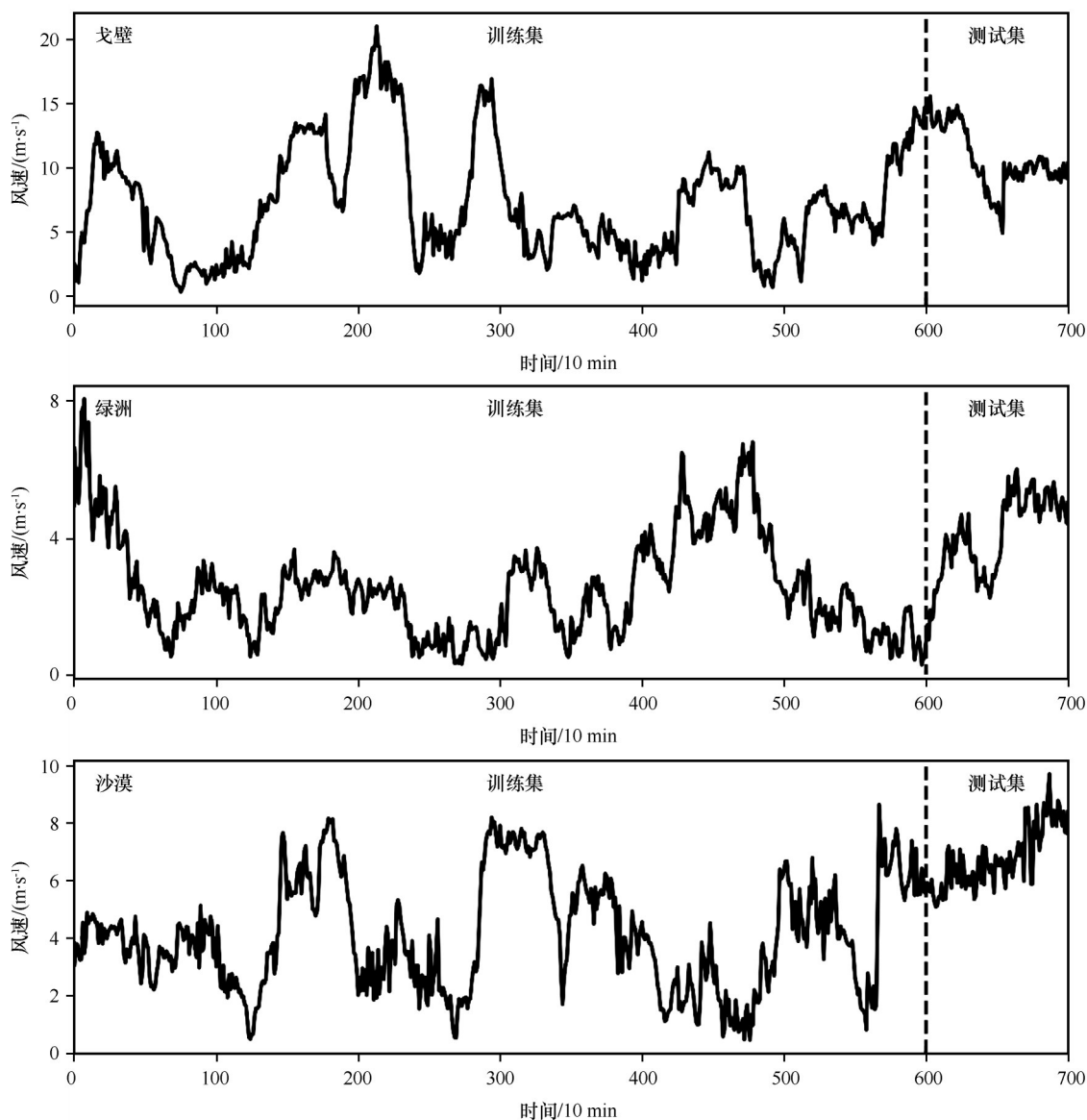


图2 3种下垫面风速时间序列

Fig.2 Wind speed time series

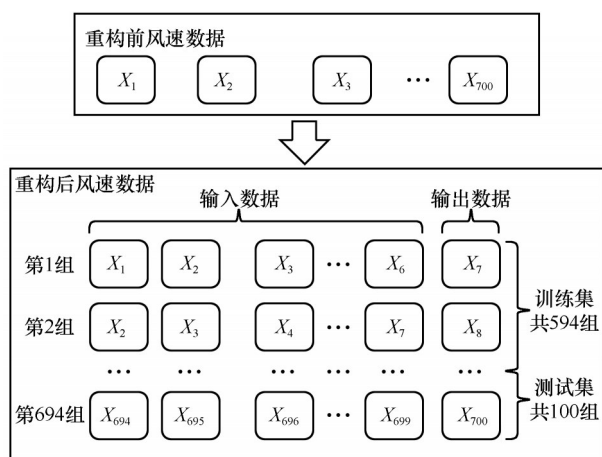


图3 数据重构示意图

Fig.3 Schematic diagram of data reconstruction

探讨其对模型结果的影响,多模型平均方法采用12个模型平均(简称A12)和3个较好模型平均(简称A3)。

2.1 3种模型在不同下垫面的预报精度

模型预报能力在3个下垫面相似(图5—7),下面以戈壁下垫面的预报结果为例说明。

小波包分解显著提高了单一模型精度。以SVR模型为例, WPD-SVR模型误差(MAPE为3.81%, MAE为 $0.40 \text{ m}\cdot\text{s}^{-1}$, RMSE为 $0.52 \text{ m}\cdot\text{s}^{-1}$)远低于SVR模型的误差(MAPE为6.13%, MAE为 $0.64 \text{ m}\cdot\text{s}^{-1}$, RMSE为 $0.93 \text{ m}\cdot\text{s}^{-1}$), WPD-SVR模型相较于SVR模型精度显著提高。其余11个模型规律相同。

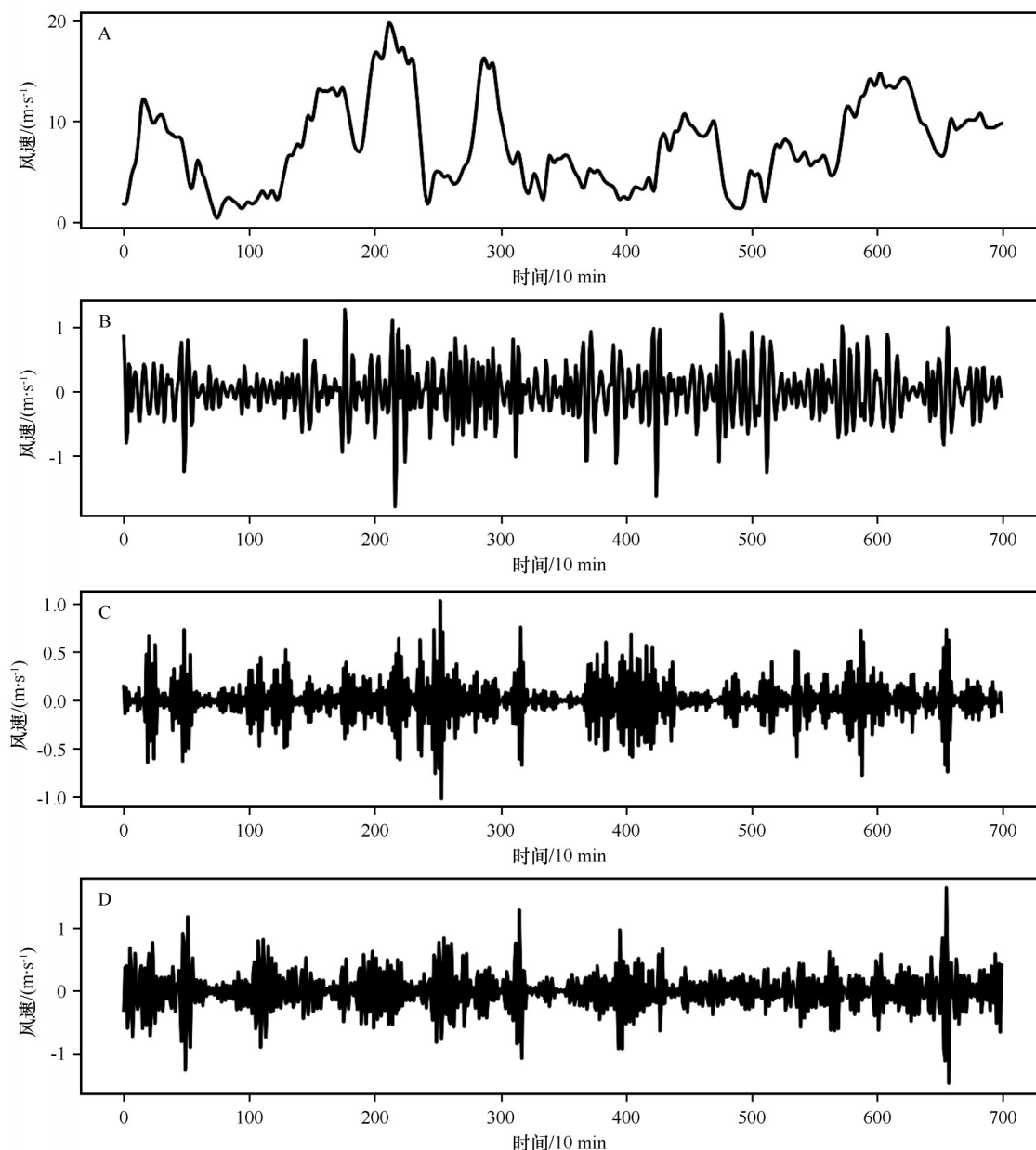


图4 数据集分解子集(以戈壁下垫面为例,A为低频子集,B、C、D为高频子集)

Fig.4 The sub-layers of the gobi dataset (A is the low-frequency sub-layer; B, C, D are the high-frequency sub-layers, respectively)

该结果证明小波包分解能够显著提高单一XXX模型的预报精度。

利用卷积神经网络模型来预报高频子集,对多数WPD-XXX混合模型的预报精度影响不明显。以WPD-SVR模型为例,WPD-SVR-CNN模型误差(MAPE为3.40%,MAE为 $0.36 \text{ m}\cdot\text{s}^{-1}$,RMSE为 $0.47 \text{ m}\cdot\text{s}^{-1}$)低于WPD-SVR模型误差(MAPE为3.81%,MAE为 $0.40 \text{ m}\cdot\text{s}^{-1}$,RMSE为 $0.52 \text{ m}\cdot\text{s}^{-1}$),WPD-SVR-CNN模型精度较好,卷积神经网络小幅度提高了WPD-SVR模型精度。但是对于WPD-SLSTM

模型,WPD-SLSTM-CNN模型误差值(MAPE为3.78%,MAE为 $0.40 \text{ m}\cdot\text{s}^{-1}$,RMSE为 $0.44 \text{ m}\cdot\text{s}^{-1}$)全都大于WPD-SLSTM模型(MAPE为3.59%,MAE为 $0.38 \text{ m}\cdot\text{s}^{-1}$,RMSE为 $0.41 \text{ m}\cdot\text{s}^{-1}$),WPD-SLSTM模型精度好,卷积神经网络没有提高WPD-SLSTM模型精度。戈壁数据集WPD-ELM-CNN模型精度略微高于WPD-ELM模型。加入卷积神经网络模型后,总计8个模型精度小幅度提高,3个模型精度未提高。该结果说明,卷积神经网络对提高WPD-XXX混合模型的精度有限。

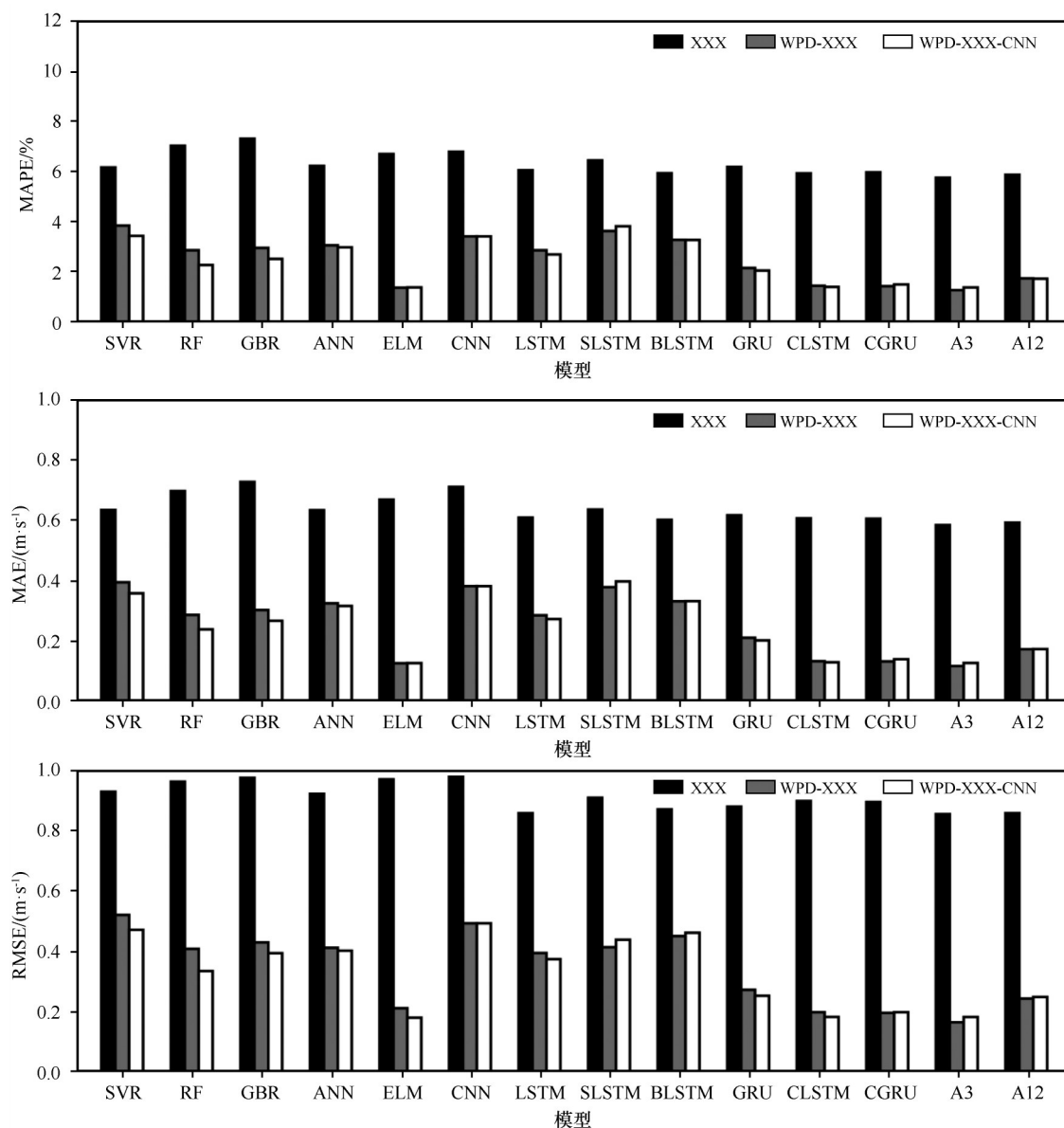


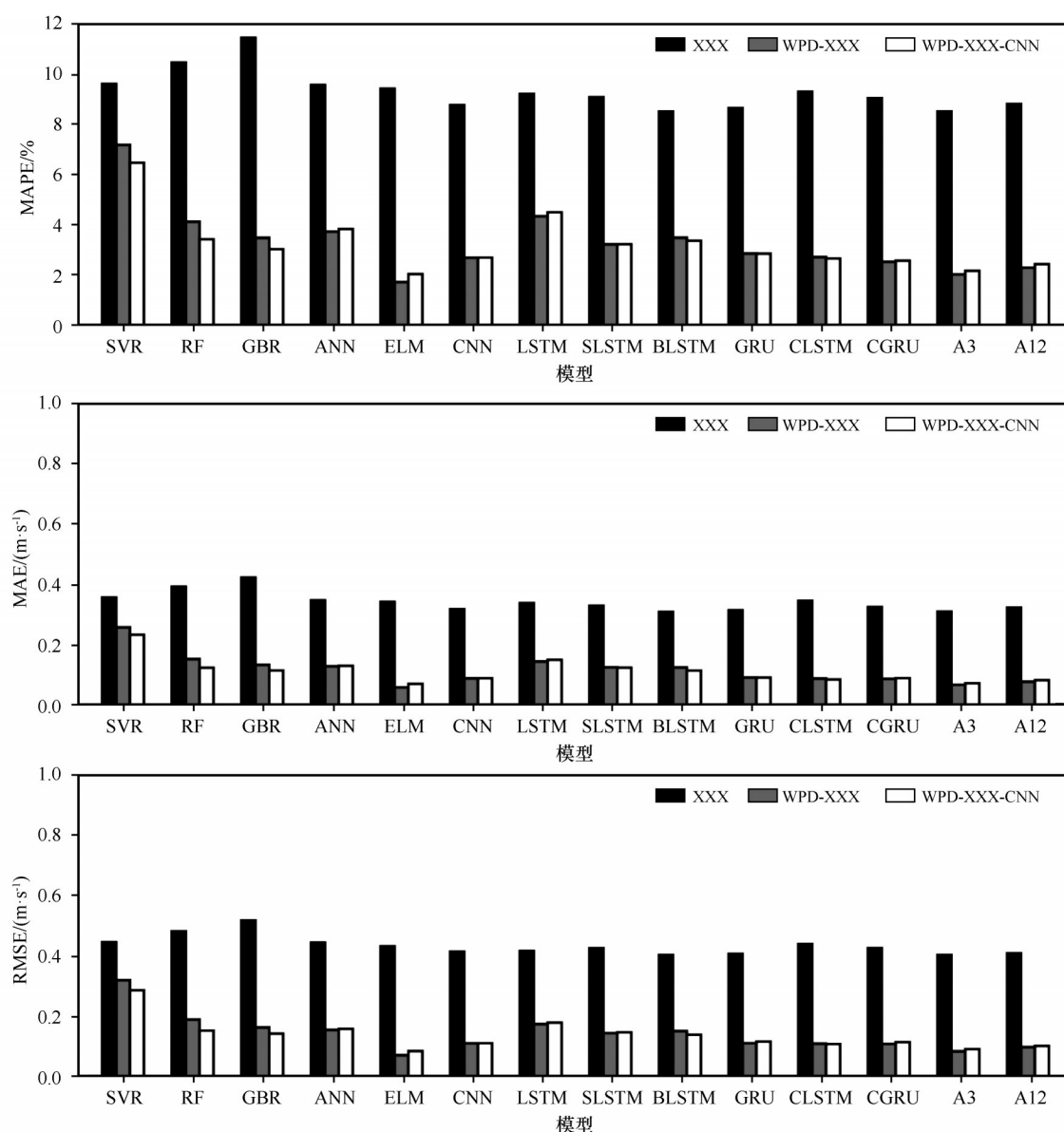
图5 戈壁下垫面3种模型预报精度

Fig.5 The forecast accuracy of three kinds of models in gobi (A3 is the average of three better models, and A12 is the average of all the 12 models)

这些模型中,小波包分解和极限学习机的耦合(WPD-ELM)、小波包分解和集特征提取和记忆功能的卷积长短时记忆网络(CLSTM)和卷积门控循环单元(CGRU)模型耦合具有较好的预报能力。其中,WPD-ELM模型误差值(MAPE为1.33%,MAE为0.13 m·s⁻¹,RMSE为0.21 m·s⁻¹)、WPD-CLSTM模型误差值(MAPE为1.40%,MAE为0.13 m·s⁻¹,RMSE为0.20 m·s⁻¹)和WPD-CGRU模型误差值(MAPE为1.39%,MAE为0.13 m·s⁻¹,RMSE为0.19 m·s⁻¹)相对较低,在3组模型中精度较好。

WPD-XXX混合模型中3个较好模型平均值精度较好。在WPD-XXX混合模型实验组中,3个较好模型误差值(MAPE为1.23%,MAE为0.12 m·s⁻¹,RMSE为0.16 m·s⁻¹)小于WPD-ELM模型误差值(MAPE为1.33%,MAE为0.13 m·s⁻¹,RMSE为0.21 m·s⁻¹),3个较好模型平均值精度优于WPD-ELM模型精度。在其他WPD-XXX模型中,WPD-ELM模型精度相对较好,故而,在WPD-XXX混合模型中,3个较好模型精度较好,能够提高模型精度。

绿洲和沙漠下垫面也有相似结果,见图6



A3为3个较好模型平均值,A12为12个模型平均值

图6 绿洲下垫面3种模型预报精度

Fig.6 The forecast accuracy of three kinds of models in oasis (A3 is the average of three better models, and A12 is the average of all the 12 models)

和图7。

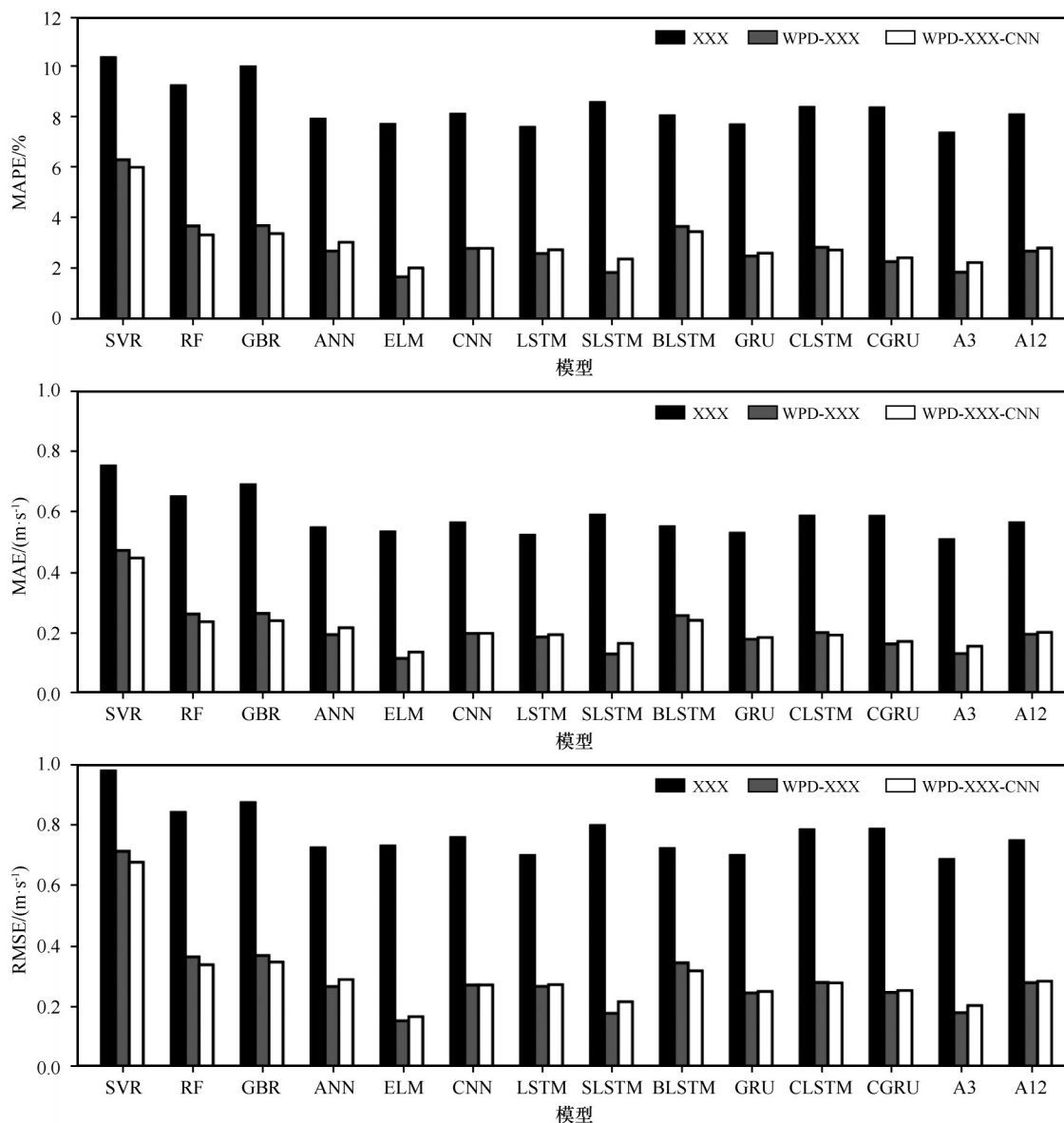
2.2 3种模型预报结果拟合程度

对于3种下垫面,3种模型预报结果的拟合程度相似(表3),在此仅以戈壁下垫面进行说明。

小波包分解显著提高单一模型预报结果的拟合程度。以SVR模型为例,单一XXX模型中SVR模型 R^2 值为0.8468,WPD-XXX混合模型中WPD-SVR模型 R^2 值为0.9520,WPD-SVR模型 R^2 值更大,模型预报结果拟合程度较好,小波包分解能够提高SVR模型的预报结果拟合程度。其余模型有相同

规律。

卷积神经网络模型对所有WPD-XXX混合模型预报拟合程度提高不显著。以WPD-SVR模型为例,WPD-SVR模型 R^2 值为0.9520,WPD-SVR-CNN模型 R^2 值为0.9607,WPD-SVR-CNN模型 R^2 值更大,模型预报结果拟合程度较好,卷积神经网络对提高WPD-SVR模型预报结果的拟合程度显著。以WPD-SLSTM模型为例,WPD-SLSTM模型 R^2 值为0.9698,WPD-SLSTM-CNN模型 R^2 值为0.9660,WPD-SLSTM模型 R^2 值更大,模型预报结果拟合程度较好,卷积神经网络对提高WPD-SLSTM



A3为3个较好模型平均值,A12为12个模型平均值

图7 沙漠下垫面3种模型预报精度

Fig.7 The forecast accuracy of three kinds of models in desert (A3 is the average of three better models, and A12 is the average of all the 12 models)

模型的预报结果拟合程度不显著。其余模型规律为以上两种情况的一种,卷积神经网络对提高所有模型预报结果拟合程度没有显著的提高。

这3组实验中,小波包分解和卷积长短时记忆网络以及卷积门控循环单元的耦合模型拟合程度相对较好。WPD-ELM模型 R^2 值为0.9922,WPD-CLSTM模型 R^2 值为0.9931,WPD-CGRU模型 R^2 值为0.9933,大于其余模型,预报结果拟合程度较好。从WPD-XXX-CNN模型可以发现同样规律。虽然卷积神经网络的使用使3组较好模型的 R^2 值有上下浮动,但3组较好模型依旧未变。

对于戈壁下垫面,单一XXX模型和WPD-XXX混合模型实验组中,3个较好模型拟合程度优于其余模型拟合程度。单一XXX模型中,3个较好模型 R^2 值为0.8704,值最大,模型预报结果拟合程度最好。WPD-XXX混合模型规律相似。3个较好模型的加入,能够提高模型预报结果的拟合程度。

2.3 多模型预报结果不稳定性

由图8可见:3个较好模型平均的预报曲线与实际观测值接近,多模型平均避免了单个模型在某些点上具有较大误差,如局部极值点处。模型预报范

表 3 模型预报结果拟合程度(R^2)

Table 3 Fitting degree of model forecast results (R^2)

模型	戈壁			绿洲			沙漠		
	单一 XXX 模型	WPD-XXX 混合模型	WPD-XXX-CNN 模型	单一 XXX 模型	WPD-XXX 混合模型	WPD-XXX-CNN 模型	单一 XXX 模型	WPD-XXX 混合模型	WPD-XXX-CNN 模型
SVR	0.8468	0.9520	0.9607	0.8682	0.9322	0.9459	0.1532	0.5506	0.5952
RF	0.8356	0.9706	0.9803	0.8460	0.9763	0.9846	0.3740	0.8827	0.8989
GBR	0.8311	0.9674	0.9726	0.8221	0.9824	0.9866	0.3243	0.8798	0.8935
ANN	0.8491	0.9700	0.9714	0.8692	0.9841	0.9834	0.5360	0.9372	0.9261
ELM	0.8329	0.9922	0.9943	0.8762	0.9967	0.9953	0.5276	0.9792	0.9754
CNN	0.8287	0.9571	0.9571	0.8864	0.9919	0.9919	0.4916	0.9349	0.9349
LSTM	0.8694	0.9725	0.9753	0.8852	0.9799	0.9788	0.5678	0.9369	0.9345
SLSTM	0.8533	0.9698	0.9660	0.8797	0.9862	0.9858	0.4360	0.9719	0.9588
BLSTM	0.8655	0.9641	0.9623	0.8924	0.9849	0.9873	0.5390	0.8950	0.9105
GRU	0.8627	0.9870	0.9888	0.8902	0.9918	0.9911	0.5674	0.9496	0.9449
CLCTM	0.8568	0.9931	0.9942	0.8719	0.9921	0.9923	0.4557	0.9308	0.9315
CGRU	0.8579	0.9933	0.9931	0.8796	0.9922	0.9914	0.4534	0.9461	0.9433
A3	0.8704	0.9953	0.9942	0.8923	0.9953	0.9945	0.5837	0.9714	0.9633
A12	0.8693	0.9896	0.9891	0.8891	0.9937	0.9932	0.5047	0.9311	0.9289

A3 为 3 个较好模型平均值,A12 为 12 个模型平均值。

围(即阴影范围)较大,平均值的使用使预报结果更接近实际观测值。同时,图 8 也体现了小波包分解对模型整体精度的提高(图 8 中阴影范围明显减小)。值得注意的是,沙漠下垫面模型预报结果很不稳定。

3 讨论

本研究中,小波包分解和卷积长短时记忆网络以及卷积门控循环单元的耦合模型在不同数据集上都表现较好。对于戈壁下垫面,WPD-ELM 模型 MAPE 值为 1.33%,WPD-CLSTM 模型 MAPE 值为 1.40%,WPD-CGRU 模型 MAPE 值为 1.39%。相较于文献[2](5.638%)和文献[9](8.92%),本研究中模型 MAPE 值较小,模型预报精度更好。其余下垫面预报结果相似。这表明本研究采用的小波包分解与深度学习模型的耦合模型对风速有较好的预报能力。

主要原因如下:

(1) 信号分解能有效降低数据复杂性,机器学习模型对分解后的风速信号预报能力更强。原始风速具有间歇性和不稳定性,信号复杂,机器学习

模型对其预报能力差。但信号分解方法可将复杂的原始信号转换为更加稳定的分量,降低了原始风速内在特性的影响,机器学习模型对各稳定分量的预报精度更高,因此,单一 XXX 模型与小波包分解组成的混合模型对风速的预报能力大大增强。这与 Zhang 等^[33]的研究结果一致。

(2) 深度学习模型对复杂时间序列预报能力更强。卷积长短时记忆网络和卷积门控循环单元既具有卷积神经网络的自动特征选择功能,又具有循环神经网络的记忆功能。ELM 为单个隐藏层前馈神经网络,模型结构简单,训练速度快,不需要复杂的参数选择就能快速学习数据的内部规律。Hao 等^[34]使用极限学习机预报风速分量以及风速误差序列,Ghimire 等^[35]在深度太阳辐射预报中使用卷积长短时记忆网络,Liu 等^[36]使用卷积门控循环单元预报风速序列子集,都有着相似发现。

(3) 不同模型具有不同学习能力,本研究中主要体现在极值预报上。在极值点处单一模型预报误差较大,而多模型平均使得预报结果更加平滑,预报值更接近实际观测值。多模型平均能有效利用不同模型特点,在极值预报上实现互补。Feng 等^[25]使用深度特征选择和多模型预报风速,模型预

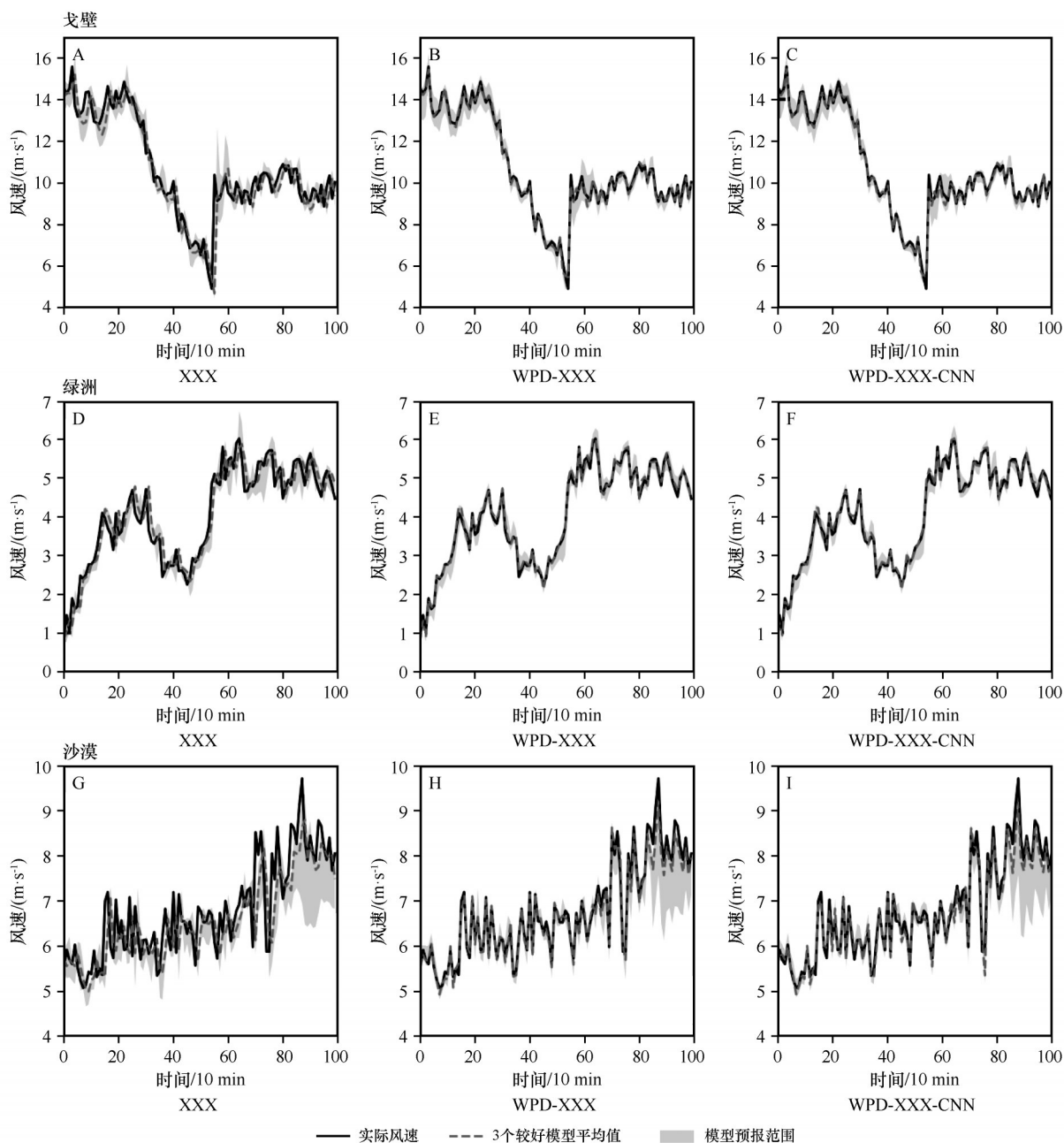


图8 3种下垫面3种模型预报结果

Fig.8 The forecast results of three kinds of models in three datasets

报精度明显优于单一模型。

4 结论

小波包分解可有效降低时间序列数据的复杂性,提高模型预报精度。

小波包分解与具有特征学习和记忆功能的深度学习模型(如卷积长短时记忆网络、卷积门控循环单元)的耦合具有较好的预报能力。

不同模型在时间序列风速不同时间点上预报能力不同,多个模型的集合预报(特别是较好模型

集合如3个较好模型集合)可以取长补短,因此多个模型的集合预报有更好的结果。

参考文献:

- [1] Liu M S, Cao Z M, Zhang J, et al. Short-term wind speed forecasting based on the Jaya-SVM model[J]. International Journal of Electrical Power & Energy Systems, 2020, 121: 106056.
- [2] 王俊, 李霞, 周昔东, 等. 基于VMD和LSTM的超短期风速预测[J]. 电力系统保护与控制, 2020, 48(11): 45-52.
- [3] 张丽英, 叶廷路, 辛耀中, 等. 大规模风电接入电网的相关问题及措施[J]. 中国电机工程学报, 2010, 30(25): 1-9.

- [4] Liu H, Mi X W, Li Y F. Comparison of two new intelligent wind speed forecasting approaches based on wavelet packet decomposition, complete ensemble empirical mode decomposition with adaptive noise and artificial neural networks[J]. *Energy Conversion and Management*, 2018, 155: 188–200.
- [5] Dong Q L, Sun Y H, Li P Z. A novel forecasting model based on a hybrid processing strategy and an optimized local linear fuzzy neural network to make wind power forecasting: a case study of wind farms in China[J]. *Renewable Energy*, 2017, 102: 241–257.
- [6] Howard T, Clark P. Correction and downscaling of NWP wind speed forecasts[J]. *Meteorological Applications*, 2007, 14(2): 105–116.
- [7] 黄彦辉, 王龙杰, 杨薛明. 基于混沌时间序列的支持向量机短期风速预测模型研究[J]. *电测与仪表*, 2015, 52(17): 32–37.
- [8] 李应求, 安勃, 李恒通. 基于NARX及混沌支持向量机的短期风速预测[J]. *电力系统保护与控制*, 2019, 47(23): 65–73.
- [9] 梁智, 孙国强, 俞娜燕, 等. 基于高斯过程回归和粒子滤波的短期风速预测[J]. *太阳能学报*, 2020, 41(3): 45–51.
- [10] 喻敏, 常毓婵, 袁浩, 等. 基于HP-EMD和ARMA的短期风速预测[J]. *中国科技论文*, 2016, 11(5): 566–570.
- [11] 孙驷洲, 陈亮, 郭兴众, 等. 基于EEMD与极限学习机的短期风速组合预测模型[J]. *安徽工程大学学报*, 2018, 33(4): 56–63.
- [12] 王晨, 寇鹏. 基于卷积神经网络和简单循环单元集成模型的风电场多风机风速预测[J]. *电工技术学报*, 2020: 1–14.
- [13] Yu C J, Li Y L, Xiang H Y, et al. Data mining-assisted short-term wind speed forecasting by wavelet packet decomposition and elman neural network[J]. *Journal of Wind Engineering and Industrial Aerodynamics*, 2018, 175: 136–143.
- [14] Harbola S, Coors V. One dimensional convolutional neural network architectures for wind prediction[J]. *Energy Conversion and Management*, 2019, 195: 70–75.
- [15] Liu H, Tian H Q, Li Y F. Four wind speed multi-step forecasting models using extreme learning machines and signal decomposing algorithms[J]. *Energy Conversion and Management*, 2015, 100: 16–22.
- [16] Hao Y, Tian C S. A novel two-stage forecasting model based on error factor and ensemble method for multi-step wind power forecasting[J]. *Applied Energy*, 2019, 238: 368–383.
- [17] Liu H, Tian H Q, Liang X F, et al. Wind speed forecasting approach using secondary decomposition algorithm and elman neural networks[J]. *Applied Energy*, 2015, 157: 183–194.
- [18] Pei S Q, Qin H, Zhang Z, et al. Wind speed prediction method based on empirical wavelet transform and new cell update long short-term memory network[J]. *Energy Conversion and Management*, 2019, 196: 779–792.
- [19] 叶瑞丽, 郭志忠, 刘瑞叶, 等. 基于小波包分解和改进Elman神经网络的风电场风速和风电功率预测[J]. *电工技术学报*, 2017, 32(21): 103–111.
- [20] 王宁, 罗汝斌, 廖俊, 等. 基于小波包分解的BP神经网络的短期风速预测[J]. *控制与信息技术*, 2019(4): 44–49.
- [21] Mi X W, Liu H, Li Y F. Wind speed forecasting method using wavelet, extreme learning machine and outlier correction algorithm[J]. *Energy Conversion and Management*, 2017, 151: 709–722.
- [22] Liu H, Yu C Q, Wu H P, et al. A new hybrid ensemble deep reinforcement learning model for wind speed short term forecasting[J]. *Energy*, 2020, 202: 1–36.
- [23] Moreno S R, Silva R G D, Mariani V C, et al. Multi-step wind speed forecasting based on hybrid multi-stage decomposition model and long short-term memory neural network[J]. *Energy Conversion and Management*, 2020, 213: 112869.
- [24] Mi X W, Liu H, Li Y F. Wind speed prediction model using singular spectrum analysis, empirical mode decomposition and convolutional support vector machine[J]. *Energy Conversion and Management*, 2019, 180: 196–205.
- [25] Feng C, Cui M J, Hodge B M, et al. A data-driven multi-model methodology with deep feature selection for short-term wind forecasting[J]. *Applied Energy*, 2017, 190: 1245–1257.
- [26] Liu H, Mi X W, Li Y F. An experimental investigation of three new hybrid wind speed forecasting models using multi-decomposing strategy and ELM algorithm[J]. *Renewable Energy*, 2018, 123: 694–705.
- [27] Liu H, Mi X W, Li Y F. Smart multi-step deep learning model for wind speed forecasting based on variational mode decomposition, singular spectrum analysis, LSTM network and ELM[J]. *Energy Conversion and Management*, 2018, 159: 54–64.
- [28] Liu H, Mi X W, Li Y F. Smart deep learning based wind speed prediction model using wavelet packet decomposition, convolutional neural network and convolutional long short term memory network[J]. *Energy Conversion and Management*, 2018, 166: 120–131.
- [29] Ding L Y, Fang W L, Luo H B, et al. A deep hybrid learning model to detect unsafe behavior: integrating convolution neural networks and long short-term memory[J]. *Automation in Construction*, 2018, 86: 118–124.
- [30] Memarzadeh G, Keynia F. A new short-term wind speed forecasting method based on fine-tuned LSTM neural network and optimal input sets[J]. *Energy Conversion and Management*, 2020, 213: 112824.
- [31] Li Y F, Shi H P, Han F Z, et al. Smart wind speed forecasting approach using various boosting algorithms, big multi-step forecasting strategy[J]. *Renewable Energy*, 2019, 135: 540–553.
- [32] Zhang Y G, Pan G F, Chen B, et al. Short-term wind speed prediction model based on GA-ANN improved by VMD[J]. *Renewable Energy*, 2020, 156: 1373–1388.
- [33] Zhang J L, Wei Y M, Tan Z F. An adaptive hybrid model for short term wind speed forecasting[J]. *Energy*, 2019, 190: 115615.
- [34] Hao Y, Tian C. A novel two-stage forecasting model based on error factor and ensemble method for multi-step wind power

- forecasting[J]. *Applied Energy*, 2019, 238: 368–383.
- [35] Ghimire S, Deo R C, Raj N, et al. Deep solar radiation forecasting with convolutional neural network and long short-term memory network algorithms [J]. *Applied Energy*, 2019, 253: 113541.
- [36] Liu H, Mi X W, Li Y F, et al. Smart wind speed deep learning based multi-step forecasting model using singular spectrum analysis, convolutional gated recurrent unit network and support vector regression [J]. *Renewable Energy*, 2019, 143: 842–854.

The hybrid of wavelet packet decomposition and machine learning models in wind speed forecasting

Wang Tongliang^{1a,2}, Ma Shaoxiu^{1a}, Gao Yang^{1a,2}, Gong Yulai^{1a,2}, An Zhishan^{1ab}

(1.a. *Key Laboratory of Desert and Desertification* / b. *Public Technical Service Center, Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences, Lanzhou 730000, China*; 2. *University of Chinese Academy of Sciences, Beijing 100049, China*)

Abstract: Wind speed forecasting is an effective method to improve wind power utilization and power system stability. Previous studies proposed a amount of wind speed forecast models, but there are few comparative studies on wind speed forecast models for different underlying land cover. This study mainly explores the ability of wavelet packet decomposition, 12 machine learning models and their coupled models to forecast the wind speed for three different underlying land cover and landforms (gobi, oasis and desert), and explores the optimized coupled model for wind speed forecast. Three sets of model experiments were set up for comparison in the article: single XXX model, WPD-XXX hybrid model and WPD-XXX-CNN hybrid model. We found that the deep learning models with feature selection, memory functions (such as convolutional long short-term memory networks) and extreme learning machine have better forecast capabilities for wind speed forecast. The wavelet packet decomposition significantly improves model accuracy. The coupled models of wavelet packet decomposition and convolutional long short-term memory network, convolutional gated recurrent unit network has better performance. This shows that the use of signal decomposition and the coupling of deep learning models can effectively improve the forecast accuracy of the models. It could be widely used in industry practices.

Key words: wind speed forecast; wavelet packet decomposition; deep learning model; arid and semi-arid area; hybrid model