

渠鸿宇,胡海川,黄彬,2024. 基于 SVM 的我国北部沿岸海上风速估算方法研究[J]. 气象,50(2):221-233. Qu H Y, Hu H C, Huang B, 2024. Research on the estimation method of sea surface wind speed along the north coast of China based on SVM[J]. Meteor Mon, 50(2):221-233(in Chinese).

基于 SVM 的我国北部沿岸海上风速估算方法研究^{*}

渠鸿宇 胡海川 黄 彬

国家气象中心,北京 100081

提 要: 沿岸海上观测站点稀少,而沿岸陆地观测站点相对密集,开展海陆分布导致的海陆风速差异特征研究,实现由陆地观测风速估算海上风速,有助于提高海上大风预报服务能力。利用我国北方地区两组浮标及其邻近陆地观测站点的 2016—2020 年逐小时平均风速和阵风风速数据,统计分析海陆风速差异特征及规律,采用支持向量机方法,构建了基于陆地平均风速、陆地阵风风速、海陆站点距离、月份及观测时次的海上风速估算模型。利用另外两组海陆观测站 2021 年观测数据对估算模型进行检验,结果表明:对于 6 级及以上的平均风速和 7 级及以上的阵风风速,模型具有较高的估算准确率,模型估算的两个检验组的海上站点平均风速(阵风风速)RMSE 分别为 $2.40 \text{ m} \cdot \text{s}^{-1}$ ($3.20 \text{ m} \cdot \text{s}^{-1}$) 和 $2.35 \text{ m} \cdot \text{s}^{-1}$ ($2.57 \text{ m} \cdot \text{s}^{-1}$),较 ERA5 分别减少了 24%(14%)和 23%(20%)。在一次温带气旋和冷空气共同影响的大风过程中,模型估算的两个检验组的海上平均风速(阵风风速)平均绝对误差分别为 $1.6 \text{ m} \cdot \text{s}^{-1}$ ($2.3 \text{ m} \cdot \text{s}^{-1}$) 和 $1.1 \text{ m} \cdot \text{s}^{-1}$ ($1.5 \text{ m} \cdot \text{s}^{-1}$),在极值时刻的平均风速(阵风风速)误差分别为 $-1.3 \text{ m} \cdot \text{s}^{-1}$ ($-0.6 \text{ m} \cdot \text{s}^{-1}$) 和 $-1.2 \text{ m} \cdot \text{s}^{-1}$ ($-3.1 \text{ m} \cdot \text{s}^{-1}$),均优于 ERA5 计算结果。基于支持向量机的海上风速估算模型能够利用陆地观测风速估算出较为准确的海上大风,可降低海上观测资料不足的影响,具有一定的应用前景。

关键词: 海陆风速差异,海上平均风,海上阵风,支持向量机

中图分类号: P413,P456

文献标志码: A

DOI: 10.7519/j.issn.1000-0526.2023.010901

Research on the Estimation Method of Sea Surface Wind Speed Along the North Coast of China Based on SVM

QU Hongyu HU Haichuan HUANG Bin

National Meteorological Centre, Beijing 100081

Abstract: Coastal offshore observation stations are rare, but coastal land observation stations are relatively dense. Carrying out research on the characteristics of land-sea wind speed difference caused by land-sea distribution and realizing the estimation of sea surface wind speed from land observation wind speed is helpful to improve the service ability of sea surface wind forecast. In this paper, the hourly mean wind speed and gust speed data of two groups of buoys and their adjacent land observation stations in the north coast of China from 2016 to 2020 are used to statistically analyze the characteristics and regularity of the difference between sea surface wind speed and land wind speed. And the support vector machine method is used to build the sea surface wind speed estimation model based on the land mean wind speed, land gust speed, distance between land-sea stations, month and hour. The estimation model is tested by using the observation data of the other two groups of land-sea observation stations in 2021. The results show that for the

^{*} 国家重点研发计划(2021YFC3000905)资助

2022 年 9 月 20 日收稿; 2023 年 3 月 20 日收修定稿

第一作者:渠鸿宇,主要从事台风海洋预报研究. E-mail:1030184860@qq.com

通讯作者:胡海川,主要从事台风海洋预报研究. E-mail:huhc1988@sina.com

mean wind speed $\geq$ scale 6 and gust wind speed $\geq$ scale 7, the model has a high estimation accuracy. The root mean square errors of the mean wind speed (gust speed) of the two groups estimated by the model are $2.40 \text{ m} \cdot \text{s}^{-1}$ ($3.20 \text{ m} \cdot \text{s}^{-1}$) and $2.35 \text{ m} \cdot \text{s}^{-1}$ ($2.57 \text{ m} \cdot \text{s}^{-1}$), respectively. Compared with ERA5, the errors decreased by 24% (14%) and 23% (20%) respectively. In a strong wind process jointly affected by an extratropical cyclone and cold air, the mean absolute errors of the mean wind speed (gust speed) estimated by the model for the two test groups are $1.6 \text{ m} \cdot \text{s}^{-1}$ ($2.3 \text{ m} \cdot \text{s}^{-1}$) and $1.1 \text{ m} \cdot \text{s}^{-1}$ ($1.5 \text{ m} \cdot \text{s}^{-1}$) respectively, and the mean wind speed (gust speed) errors at the extreme moment are $-1.3 \text{ m} \cdot \text{s}^{-1}$ ($-0.6 \text{ m} \cdot \text{s}^{-1}$) and $-1.2 \text{ m} \cdot \text{s}^{-1}$ ($-3.1 \text{ m} \cdot \text{s}^{-1}$) respectively, which are better than those from ERA5. The sea surface wind speed estimation model based on support vector machine can estimate accurate heavy sea surface wind speed using the land observation wind speed, which can reduce the impact of insufficient sea observation data, and has a certain application prospect.

Key words: land-sea wind speed difference, sea surface mean wind, sea surface gust, support vector machine

引 言

海上大风是主要的海洋气象灾害之一(许小峰等,2009),开展海上大风的研究及预报对保证海上作业安全,提高海洋作业效率及减少开支具有重要意义(尹尽勇等,2012)。随着数值模式的飞速发展,海面风场预报取得长足进步(李敏等,2009),但受动力框架、初始场误差、物理参数化方案等因素影响,模式海面风预报与实况仍有一定差距(Toth et al, 2001; Mu et al, 2002; 季晓阳等, 2005; 胡海川等, 2021)。为了进一步提升海上大风预报能力,林良勋等(2004)、吴曼丽等(2012)、薄文波等(2013)、胡波(2019)、潘静等(2022)诸多学者开展了海上大风统计或动力统计预报方法的研究工作,但由于海洋观测资料稀缺,上述方法中多以沿岸、海岛等陆地下垫面观测站数据近似代表海上资料。为进一步提升海上大风预报准确率,应将由下垫面不同所导致的海陆风速差异考虑其中,基于海陆风速差异特征及规律,利用较密集的陆地站点观测风速估算出更真实的海上风速后,再建立海上大风统计或动力统计预报模型。井传才等(1995)、张新玲和吴增茂(1998)和刘京雄等(2004)分析海陆平均风速差异特征及规律,并基于线性、指数及对数函数拟合方法,使用海岛站、石油平台站、船舶站等陆地平均风速估算海上平均风速。但由于海岛站下垫面依然为陆地,石油平台站测风高度过高以及船舶观测资料连续性较差等问题(陈洪滨等,2019),基于上述站点进行海陆风速差异特征研究和建立由陆地风速估算海上风速模

型会存在一定不足。浮标观测数据具有连续、实时、可靠的特点(王波等,2014),更适合用来研究海陆风速差异特征,构建由陆地风速估算海上风速模型。另外,相比于平均风速,强阵风更容易造成生命财产损失(Nakamura et al, 2007; 胡海川等, 2022),因此同样需要研究海陆阵风风速差异,并构建海上阵风风速的估算模型。

支持向量机(support vector machine, SVM)是建立在统计学习理论和结构风险最小化原理基础上的机器学习方法,相比传统算法,其在解决小样本、非线性、高维问题中具有特有的优势(丁世飞等, 2011),已被广泛应用于降水、雷暴大风、台风等气象要素的预报研究之中(Lin et al, 2009; 施萧等, 2012; 钱燕珍等, 2012; Nayak and Ghosh, 2013; Mercer et al, 2013; 黄威和牛若芸, 2017; 杨璐等, 2018; Pour et al, 2018)。SVM的基本思想是升维和线性化。通过非线性映射函数,把样本空间映射到一个高维的特征空间,使得在高维空间中可以应用线性学习的方法解决原始样本空间中的非线性问题。同时,通过应用核方法(kernel method),无需求解复杂的映射函数即可求得内积,大大减小了计算复杂度。

利用 2016—2020 年逐小时平均风速和阵风风速数据,选取两组位于我国北部沿岸的浮标和与其邻近的陆地观测站点,研究由海陆分布导致的海陆平均风速和阵风风速差异特征及规律,采用 SVM 回归方法构建沿岸海上风速估算模型,以实现利用沿岸陆地风速获取较准确的沿岸海上风速。使用另外两组海陆观测站点 2021 年数据对海上风速估算

模型进行检验,验证其在应用中的效果。

1 资 料

本文所用资料包括 2016—2021 年中国气象局地面实况观测数据中的逐小时风速数据(2 min 平均)和小时内极大风速(3 s 平均)数据(精度为 $0.1 \text{ m} \cdot \text{s}^{-1}$)。选取曹妃甸浮标站和曹妃甸工业区站记为第一训练组,烟台港浮标站和烟台港务局站记为第二训练组,选取曹妃甸浮标站和曹妃甸零公里国家无人观测站记为第一检验组,董家口浮标站和日照水上运动基地站记为第二检验组(表 1),各站点位置如图 1 所示。其中训练组中的 2016—2020 年数据(记为训练集)用于特征分析和模型训练,检验组中的 2021 年数据(记为测试集)用于检验。两组训练组海陆站点直线距离分别为 24.4 km、7.7 km,两组检验组的海陆站点直线距离分别为 37.6 km、35.4 km,尽量保证了海陆站点受相同的天气系统影响(吕美仲等,2004)。同时,沿岸陆地站点的高度均小于 10 m,避免了海拔高度对海陆风速

差异研究的影响(赵鸣和苗曼倩,1992)。

筛选海陆平均风速均大于 3 级($\geq 3.4 \text{ m} \cdot \text{s}^{-1}$)的情况作为研究对象,共获取成对的海陆平均风速和阵风风速样本 14350 对,其中 2016—2020 年的训练组数据共 9809 对,2021 年检验组数据共 4541 对。

为横向比较估算模型准确度,采用双线性插值法,将 2021 年 ERA5(Hersbach et al,2018)逐小时的 10 m 风数据和阵风数据(空间分辨率为 $0.25^\circ \times 0.25^\circ$)插值到检验组的海上站点处,进行 ERA5 风速与海上站点实测风速和模型估测风速的对比分析。

2 差异分析

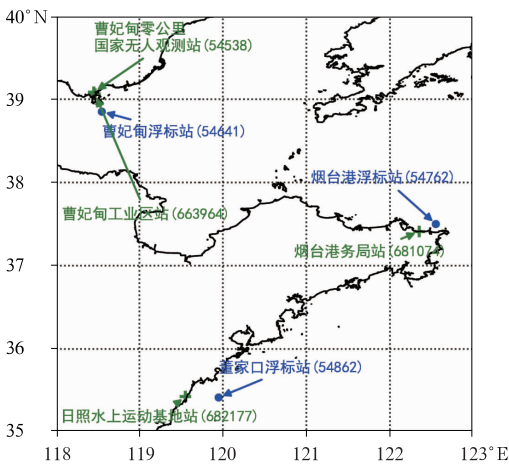
2.1 整体差异

图 2 展示了 2016—2020 年两组海陆站点平均风速和阵风风速的散点分布,图中红色虚线为线性拟合线, k 为线性拟合斜率。如图 2 所示,两组海陆

表 1 训练组站点和检验组站点的名称及站号

Table 1 Name and station number of training group and test group

组名	沿岸海上站点名称	沿岸陆地站点名称
第一训练组	曹妃甸浮标站	曹妃甸工业区站
第二训练组	烟台港浮标站	烟台港务局站
第一检验组	曹妃甸浮标站	曹妃甸零公里国家无人观测站
第二检验组	董家口浮标站	日照水上运动基地站



注:蓝、绿色分别为沿岸海上、陆地站点。

图 1 海陆站点位置分布

Fig. 1 Distribution of land-sea stations

站点中的平均风速与阵风风速的散点均主要集中在对角线左侧,且线性拟合斜率均大于 1,表明海上的平均风速和阵风风速均大于陆地。对于同一训练组海陆站点中的平均风速及阵风风速分布而言,平均风速与对角线的偏差更大。两组海陆站点中的平均风速线性拟合斜率分别为 1.88 和 1.48,阵风风速的线性拟合斜率分别为 1.32 和 1.10,表明海陆下垫面的差异对平均风速影响更为显著,而阵风风速的海陆差异相对较小。无论是平均风速还是阵风风速,第二训练组海陆站点的线性拟合斜率均小于第一训练组,平均风速拟合斜率偏小 0.4,阵风风速拟合斜率偏小 0.21。这可能与海陆站点间的距离有关,距离越近,海陆风速差异有减小的趋势,因此在构建海上风速估算模型中需要考虑到海陆站点距离。

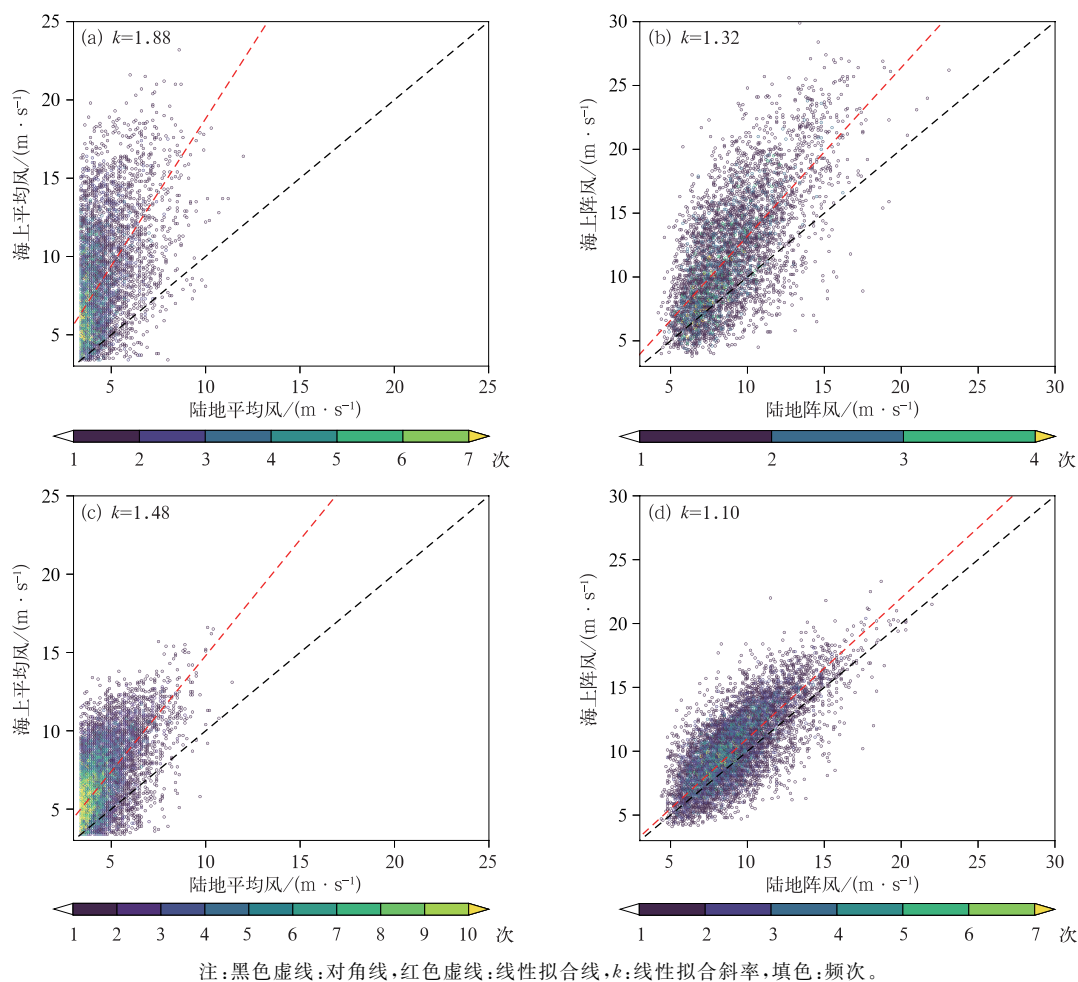


图 2 2016—2020 年海陆(a,c)平均风速和(b,d)阵风风速出现频次的散点分布图
(a,b)第一训练组,(c,d)第二训练组

Fig. 2 Scatter map of frequencies of (a, c) mean wind speed and (b, d) gust speed for the two groups of land-sea stations from 2016 to 2020
(a, b) the first training group, (c, d) the second training group

无论是海上还是陆地,阵风因子(在时距 T_0 的时间段内持续时间为 τ 的最大风速与时距为 T_0 的平均风速之比,本文中 T_0 为 2 min, τ 为 3 s) (Monahan and Armendariz, 1971),都会随风速的增大呈减小的趋势,但由于下垫面的差异,也会导致阵风因子的差异。通过对比 2016—2020 年第一训练组海陆站点不同平均风速下的阵风因子可以看出(图 3),在不同平均风速下,陆地阵风因子的最大值、最小值、平均值等均大于海上。在平均风速为 3 级的情况下,陆地阵风因子平均值为 2.00,海上则为 1.52;在平均风速为 6 级情况下,陆地阵风因子平均值为 1.49,明显大于海上(1.30)。对于相同风速,陆地阵风因子的分布更为离散,平均风速在 3~6 级时,陆地阵风因子四分位差分别较海上偏大

0.15、0.18、0.13 和 0.06。第二训练组海陆站点的情况与第一训练组基本相同(图略)。海陆分布的差异会导致陆地上平均风速及阵风风速均小于海上。相比于阵风风速,平均风速受海陆分布影响更为显著,因而导致阵风因子的明显差异。相同平均风速情况下,陆地阵风因子更大且分布更为离散。

2.2 季节影响

图 4 为 2016—2020 年第一训练组海陆站点在不同季节下平均风速和阵风风速均值及两者比值。如图所示,海上与陆地平均风速和阵风风速的比值均在秋、冬季较大,春、夏季较小,平均风速和阵风风速比值的极差分别为 0.61 和 0.43,表明海陆下垫面差异导致的海陆平均风速和阵风风速差异具有显

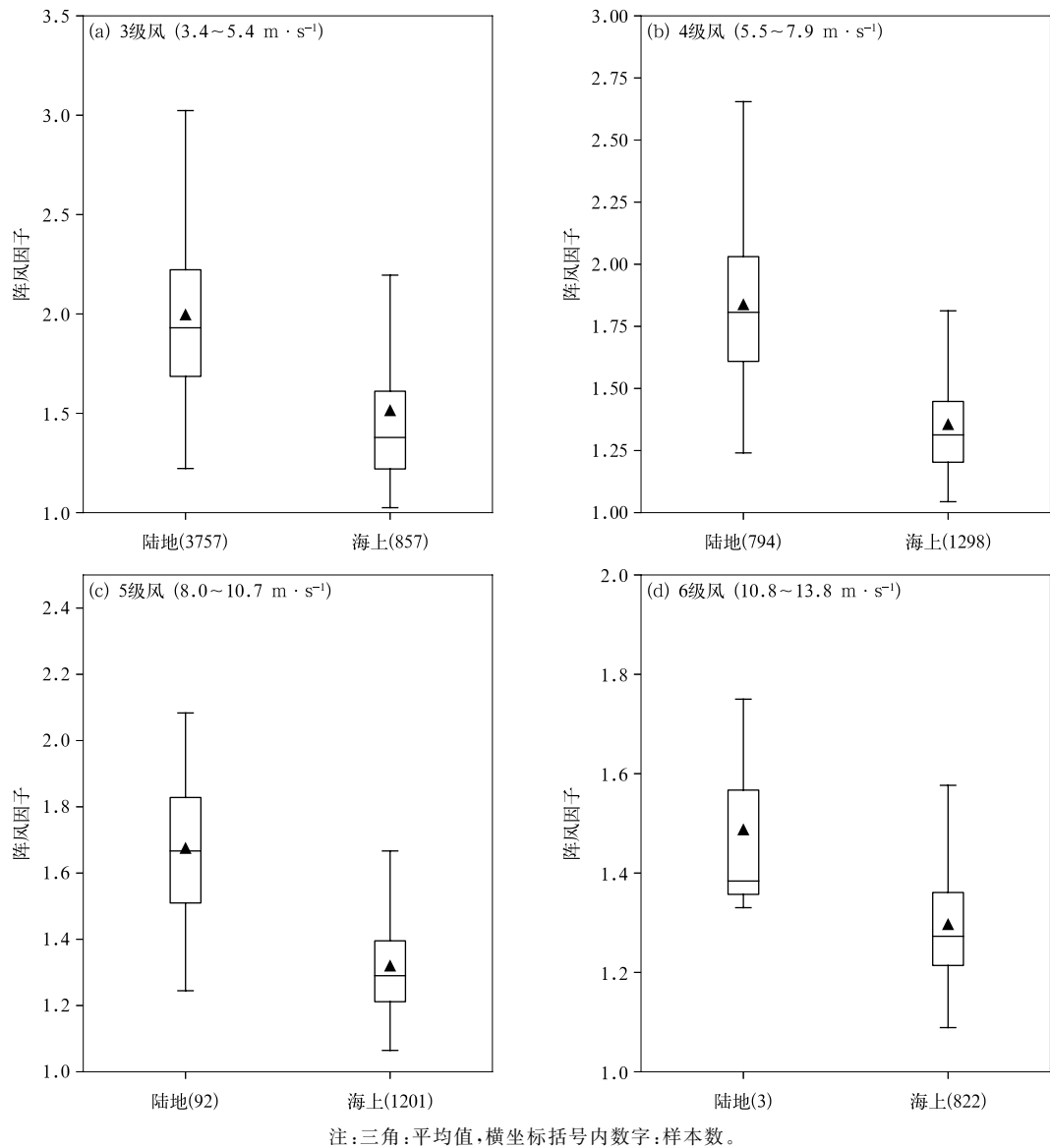


图 3 2016—2020 年第一训练组海陆站点不同平均风风级下海上和陆地阵风因子箱线图

Fig. 3 Box chart of gust factors on sea and land under different mean wind scales for the first training group of sea and land stations from 2016 to 2020

著的季节变化特征。沿岸陆地风速的季节变化并不显著,平均风速极差仅为 $0.21 \text{ m} \cdot \text{s}^{-1}$,阵风风速极差仅为 $0.74 \text{ m} \cdot \text{s}^{-1}$;而沿岸海上风速具有明显的季节变化,平均风速极差为 $2.39 \text{ m} \cdot \text{s}^{-1}$,阵风风速极差为 $3.46 \text{ m} \cdot \text{s}^{-1}$,显著大于陆地风速极差,海上站点平均风速和阵风风速均在秋、冬季较强,春、夏季较弱,使得海陆风速差异也表现出类似的季节变化。第二训练组海陆站点的情况与第一训练组基本相同(图略)。因此在构建海上风速估算模型中需要加入月份作为预报因子,以提高海上风速估算的准

确率。

2.3 日变化

图 5 为 2016—2020 年第一训练组海陆站点的平均风速和阵风风速均值及两者比值的日变化。如图所示,海上与陆地平均风速和阵风风速的比值均在凌晨偏大,午后偏小,平均风速和阵风风速比值的极差分别为 0.73 和 0.44,表明海陆下垫面差异导致的海陆平均风速和阵风风速差异具有显著的日变化特征。沿岸陆地风速的日变化并不显著,平均风

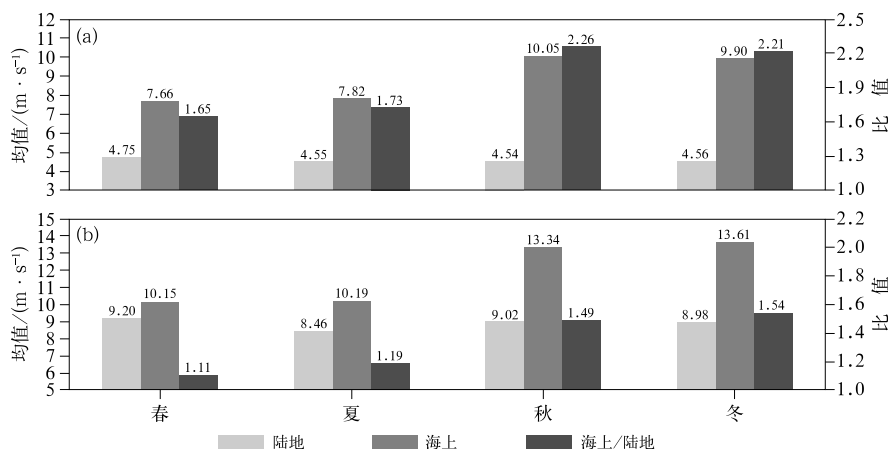


图 4 2016—2020 年第一训练组海陆站点(a)平均风速和

(b)阵风风速的均值及两者比值的季节变化

Fig. 4 Seasonal variation of mean value of (a) mean wind speed, (b) gust speed and the ratio of the first training group of sea and land stations from 2016 to 2020

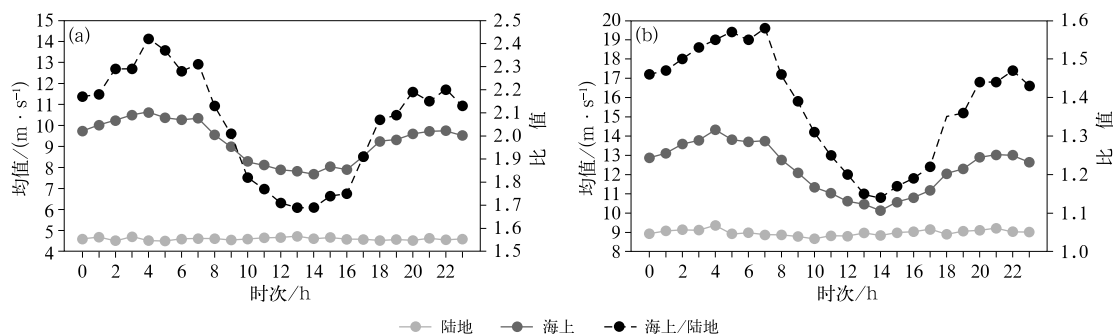


图 5 2016—2020 年第一训练组海陆站点(a)平均风速和(b)阵风风速均值及两者比值的日变化

Fig. 5 Daily variation of mean value of (a) mean wind speed, (b) gust speed and the ratio of the first training group of sea and land stations from 2016 to 2020

速极差仅为 $0.22 \text{ m} \cdot \text{s}^{-1}$, 阵风则仅为 $0.69 \text{ m} \cdot \text{s}^{-1}$; 而沿岸海上风速具有明显的日变化, 平均风速极差为 $2.93 \text{ m} \cdot \text{s}^{-1}$, 阵风则为 $4.19 \text{ m} \cdot \text{s}^{-1}$, 显著大于陆地风速极差, 海上站点风速较大值出现在凌晨, 较小值出现在午后, 使得海陆风速比值也表现出类似的日变化。第二训练组海陆站点的情况与第一训练组基本相同(图略)。因此在构建海上风速估算模型时需要加入观测时次作为预报因子, 以提高海上风速估算准确率。

3 估算模型

利用 2016—2020 年数据建立基于 SVM 回归方法的海上平均风速和阵风风速估算模型。数据预处理以及模型训练、优化、估算和检验流程如图 6 所

示, 其中数据预处理中采用 Z-Score 标准化方法对连续型估算因子(如平均风速、阵风风速和海上陆站点距离)进行标准化处理, 使用训练集估算因子的均值和标准差对测试集估算因子进行标准化; 离散型估算因子(如月份、时次)均为有序型离散变量, 直接使用相应的月份或时次数值表示, 无需进行额外的编码以及标准化处理。

SVM 回归模型与传统回归模型以模型输出值与真值偏差来计算损失不同, SVM 回归模型设置了损失距离 ϵ , 仅当偏差大于 ϵ 时才计算损失。 ϵ 是一个超参数, 设置得太小无法保证所有样本点都在损失距离内, 设置太大又容易被个别离散点影响导致过拟合, 因此, SVM 回归模型通过添加松弛变量允许样本点误差大于 ϵ [如式(3)所示, 其中 ξ_i 和 $\hat{\xi}_i$ 为松弛变量], 从而增加模型的鲁棒性。SVM 回归模

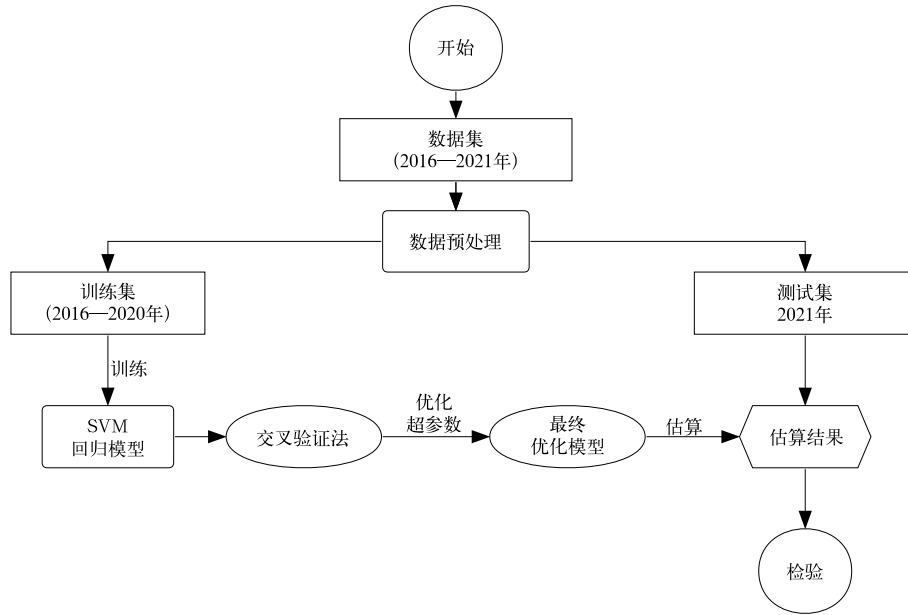


图6 数据预处理以及估算模型的训练、优化、估算和检验流程图

Fig. 6 Flow chart of data preprocessing and training, optimization, estimation and test of estimation model

型可用式(1)表示。

$$\mathbf{w} \cdot \varphi(x_i) + b = 0 \quad (1)$$

$$f(\mathbf{w}, b, \xi_i, \dot{\xi}_i) = \frac{\|\mathbf{w}\|^2}{2} + C \sum_{i=1}^n (\xi_i + \dot{\xi}_i) \quad (2)$$

$$-\epsilon - \dot{\xi}_i \leq y_i - \mathbf{w} \cdot \varphi(x_i) - b \leq \epsilon + \xi_i \quad (3)$$

$$\dot{\xi}_i \geq 0, \xi_i \geq 0 \quad (4)$$

式中: $\mathbf{w}$ 和 b 分别为权重向量和频数, $\varphi(x)$ 代表由低维到高维的映射函数, 模型求解可表示为在限制条件[式(3)和式(4)]下, 求解损失函数[式(2)], 其中 C 为超参数]最小。式(2)中的第一项为权重系数的 L2 范数, 在确定的 ϵ 下, 其尽可能的小可允许更多的样本误差小于 ϵ ; C 用来调节对偏差大于 ϵ 样本点的惩罚程度, C 越大, 模型对这些离散点越敏感, 会提高训练集准确度, 但容易过拟合, 泛化能力差。

SVM 回归模型的训练和预测需要计算升维后向量的内积, 使用核方法可实现隐式求解高维向量内积, 大大减小了计算的复杂度。SVM 回归模型常用的核函数有线性核函数、多项式核函数和径向基核函数(又名高斯核函数)。线性核函数具有参数少、速度快的优势, 但无法实现升维, 主要用于线性可分的问题。多项式核函数和径向基核函数都能实现升维从而解决非线性问题, 但前者由于参数多(三个), 学习复杂度高。相比而言, 径向基核函数参数较少(仅一个超参数), 可隐式计算被映射到无限维

空间的两个向量的内积, 是应用最广泛的核函数。由于使用陆地风速估算海上风速并非简单的线性过程(刘京雄等, 2004), 因此 SVM 回归模型采用径向基核函数, 如式(5)所示。

$$K(\mathbf{x}_1, \mathbf{x}_2) = e^{-\gamma \|\mathbf{x}_1 - \mathbf{x}_2\|^2} \quad \gamma > 0 \quad (5)$$

式中: $\mathbf{x}_1, \mathbf{x}_2$ 表示原始样本空间中的任意两个样本的输入向量, γ 为超参数, 其中 γ 越大, 高斯核函数变化越剧烈, 模型对噪声样本越敏感, 越容易过拟合。

为找出最优的超参数(核函数超参数 γ 、惩罚项 C 和损失距离 ϵ), 结合贝叶斯优化的最佳参数智能搜索方法, 使用交叉验证法(Kohavi, 1995)(取 2016—2020 年训练组数据的其中一年作为验证集, 其余 4 年为训练集, 重复 5 次), 以所有验证集的均方根误差(root mean square error, RMSE)作为评估指标, 对超参数进行优化, 从而得到最终优化后的模型。

为了找出最优的估算因子组合, 结合上述的差异特征分析, 使用表 1 所示的估算因子组合, 分别建模, 并使用交叉验证法寻求最优超参数, 各模型所有验证集的 RMSE 如表 2 所示(按平均风速 RMSE 降序排序)。

从表 2 中可以看出, 无论是平均风速还是阵风风速估算模型, 仅使用沿岸陆地平均风速估算海上

风速的 RMSE(平均风速 RMSE 为 $2.577 \text{ m} \cdot \text{s}^{-1}$, 阵风风速 RMSE 为 $3.220 \text{ m} \cdot \text{s}^{-1}$)均高于仅使用沿岸陆地阵风风速估算海上风速的 RMSE(平均风速 RMSE 为 $2.293 \text{ m} \cdot \text{s}^{-1}$, 阵风风速 RMSE 为 $2.684 \text{ m} \cdot \text{s}^{-1}$),这与上文分析得到的海陆下垫面的差异对平均风速影响更为显著,而阵风风速的海陆差异相对较小相对应;当模型同时使用陆地平均风速和阵风风速作为预报因子组合时,平均 RMSE 较模型仅使用一个要素时均有所降低,说明应同时考虑陆地平均风速和阵风风速。在使用以上估算因子组合基础上,逐步增加海陆站点距离、月份以及时次分别建模,三个模型验证集 RMSE 均有进一步降低,平

均风速 RMSE 分别为 2.079 、 1.980 、 $1.738 \text{ m} \cdot \text{s}^{-1}$, 阵风风速 RMSE 分别为 2.497 、 2.204 、 $2.044 \text{ m} \cdot \text{s}^{-1}$,这进一步说明了海陆风速差异与海陆站点距离存在相关性,且具有季节和日变化,在估算因子中增加海陆站点距离、月份以及时次可进一步提高估算模型的准确率。

综上所述,最佳估算因子组合为沿岸陆地平均风速及阵风风速、海陆站点距离、月份、时次,使用该估算因子组合,构建的平均风速估算模型的超参数 C 为 2.030 , γ 为 0.206 , ϵ 为 0.446 , 构建的阵风风速估算模型的超参数 C 为 2.095 , γ 为 0.164 , ϵ 为 0.113 。

表 2 以不同估算因子组合构建模型的平均风速和阵风风速估算的 RMSE(单位: $\text{m} \cdot \text{s}^{-1}$)

Table 2 The root mean square error of the mean wind speed and gust speed of all validation sets of the model constructed with different combinations of estimation factors (unit: $\text{m} \cdot \text{s}^{-1}$)

估算因子组合	平均风速	阵风风速
平均风速	2.577	3.220
阵风风速	2.293	2.684
平均风速+阵风风速	2.289	2.678
平均风速+阵风风速+海陆站点距离	2.079	2.497
平均风速+阵风风速+海陆站点距离+月份	1.980	2.204
平均风速+阵风风速+海陆站点距离+月份+时次	1.738	2.044

4 检验评估

4.1 误差检验

由于海上大风较高的破坏力,是海上风预报服务和研究的重点,因此误差检验仅针对 6 级及以上的平均风以及 7 级及以上的阵风(按海上风分级)。图 7 为两个检验组海上站点 6 级及以上平均风速和 7 级及以上阵风风速的模型估算的与 ERA5 计算的相对误差散点分布。由图可见,平均风速和阵风风速的相对误差散点的横坐标值均整体小于纵坐标值,且较多的样本点位于对角线之上。对于两个检验组的海上站点,ERA5 计算的平均风速的平均相对误差分别为 20.2% 和 18.1% ,而模型估算的分别为 13.9% 和 14.6% ;ERA5 计算的阵风风速平均相对误差分别为 15.1% 和 13.6% ,而模型估算的分别为 13.6% 和 12.4% (表 3);模型估算的平均风速相对误差较 ERA5 小的样本点占比分别为 73.5% 和 63.0% ,模型估算的阵风风速则分别为 57.0% 和 53.0% 。

对于两个检验组海上站点 6 级及以上的平均风

速,ERA5 的 RMSE 分别为 $3.16 \text{ m} \cdot \text{s}^{-1}$ 和 $3.05 \text{ m} \cdot \text{s}^{-1}$,模型估算的 RMSE 分别为 $2.40 \text{ m} \cdot \text{s}^{-1}$ 和 $2.35 \text{ m} \cdot \text{s}^{-1}$,相较于 ERA5 计算结果,模型估算海上平均风速 RMSE 分别减少了 24% 和 23% (表 3);对于两个检验组海上站点 7 级及以上的阵风风速,ERA5 的 RMSE 分别为 $3.70 \text{ m} \cdot \text{s}^{-1}$ 和 $3.21 \text{ m} \cdot \text{s}^{-1}$,模型估算的 RMSE 分别为 $3.20 \text{ m} \cdot \text{s}^{-1}$ 和 $2.57 \text{ m} \cdot \text{s}^{-1}$,相较于 ERA5 计算结果,模型估算的阵风风速 RMSE 分别减少了 14% 和 20% (表 3)。

进一步对 6 级及以上的平均风速以及 7 级及以上阵风风速进行分级检验,检验结果如图 8 所示。由图可见,无论是 ERA5 计算还是模型估算结果在大部分风级下均较实测偏小,但模型估算的平均误差较 ERA5 计算更接近于 $0 \text{ m} \cdot \text{s}^{-1}$ 。在不同风级下,模型估算的两个检验组海上站点的平均风速的平均误差为 $-2.32 \sim 0.38 \text{ m} \cdot \text{s}^{-1}$,阵风风速的平均误差为 $-3.14 \sim 0.92 \text{ m} \cdot \text{s}^{-1}$,而 ERA5 计算的平均风速的平均误差为 $-4.77 \sim -1.80 \text{ m} \cdot \text{s}^{-1}$,阵风风速的平均误差为 $-5.94 \sim -1.39 \text{ m} \cdot \text{s}^{-1}$ 。同时,对于大多数风级,模型估算的海上风速的 RMSE 均较 ERA5 偏小,在不同风级下,前者估算的海上平均风速 RMSE 为 $2.25 \sim 2.79 \text{ m} \cdot \text{s}^{-1}$,阵风风速

RMSE 为 $2.54\sim 3.67\text{ m}\cdot\text{s}^{-1}$ 。

综上所述,对于 6 级及以上的平均风以及 7 级及以上的阵风, SVM 回归模型估算的海上风速具

有较高的准确率,且优于 ERA5 计算结果,可见该模型能较好地估算出海上大风。

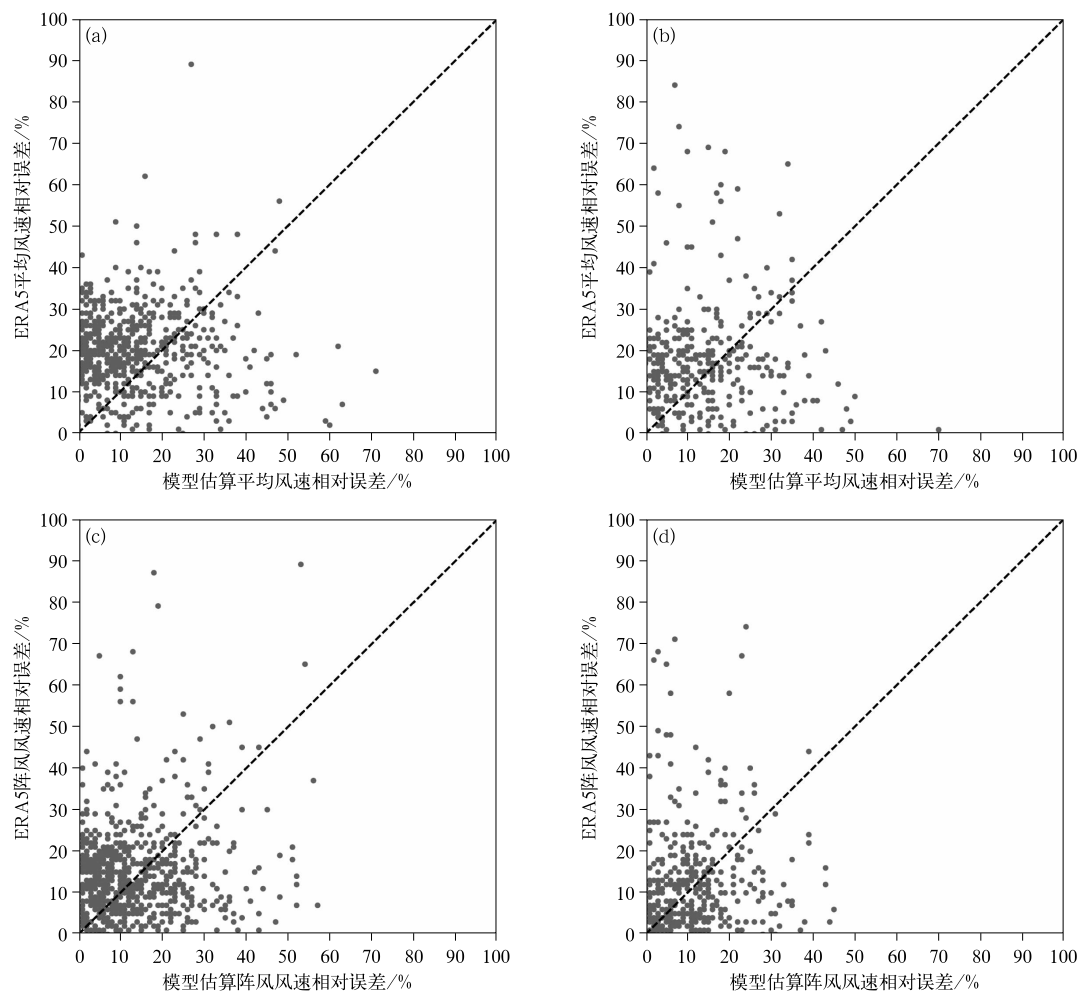


图 7 2021 年模型估算与 ERA5 计算的(a,b)6 级及以上平均风速和(c,d)7 级及以上阵风风速的相对误差散点分布
(a,c)第一检验组,(b,d)第二检验组

Fig.7 Scatter diagram of relative errors estimated by the model and ERA5 for the (a, b) mean wind speed $\geq$ scale 6 and (c, d) gust speed $\geq$ scale 7 in 2021
(a, c) the first test group, (b, d) the second test group

表 3 2021 年模型估算与 ERA5 计算的 6 级及以上平均风速和 7 级及以上阵风风速的 RMSE 和平均相对误差
Table 3 Root mean square errors and mean relative errors estimated by the model and ERA5 for the mean wind speed $\geq$ scale 6 and gust speed $\geq$ scale 7 in 2021

组名	方法	平均风速		阵风风速	
		RMSE/($\text{m}\cdot\text{s}^{-1}$)	平均相对误差/%	RMSE/($\text{m}\cdot\text{s}^{-1}$)	平均相对误差/%
第一检验组	ERA5	3.16	20.2	3.70	15.1
	模型估算	2.40	13.9	3.20	13.6
第二检验组	ERA5	3.05	18.1	3.21	13.6
	模型估算	2.35	14.6	2.57	12.4

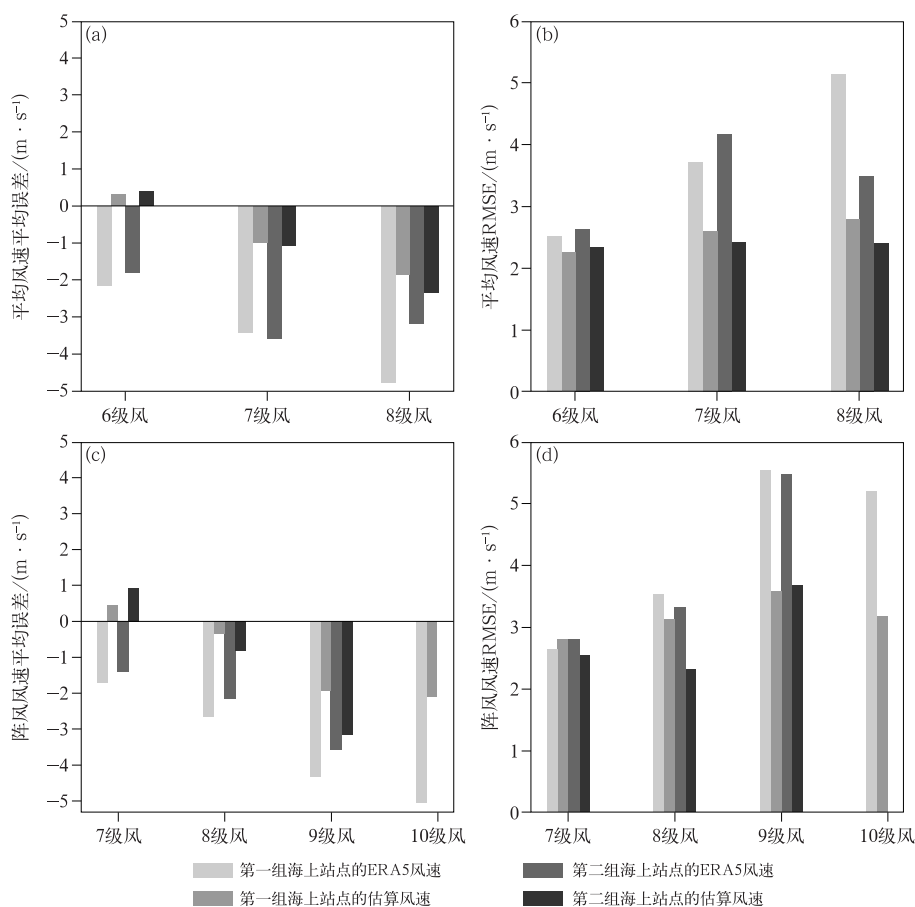


图 8 2021 年 ERA5 计算和模型估算的(a,b)平均风速和(c,d)阵风风速
在各风级下的(a,c)平均误差和(b,d)RMSE

Fig. 8 (a, c) Mean error and (b, d) RMSE of (a, b) mean wind speed
and (c, d) gust speed calculated by ERA5 and estimated
by model under different wind scales in 2021

4.2 个例对比

2021 年 11 月 21—22 日(北京时,下同),受冷空气和温带气旋的共同影响,渤海、渤海海峡、黄海、东海、台湾海峡、台湾以东洋面、巴士海峡、南海北部海域、北部湾出现了 8~9 级、阵风 10 级的大风(聂高臻和黄彬,2022)。

图 9 为两个检验组在 21 日 12 时至 22 日 12 时实测到的风速时间序列,可以看出,海陆风速差异明显,且模型估算的海上风速较 ERA5 计算更接近实测。对于第一检验组,海上实测的平均风速和阵风风速均值分别为 $15.8 \text{ m} \cdot \text{s}^{-1}$ 和 $21.5 \text{ m} \cdot \text{s}^{-1}$,远大于其对应的陆地站点实测风速,模型估算的平均风速和阵风风速平均绝对误差分别为 $1.6 \text{ m} \cdot \text{s}^{-1}$ 和 $2.3 \text{ m} \cdot \text{s}^{-1}$,较 ERA5(分别为 $3.4 \text{ m} \cdot \text{s}^{-1}$ 和 $4.9 \text{ m} \cdot$

s^{-1})偏小。此次过程的最大平均风速和阵风风速出现在 21 日 20 时,分别为 $21.4 \text{ m} \cdot \text{s}^{-1}$ 和 $25.9 \text{ m} \cdot \text{s}^{-1}$,同一时刻模型的估算偏差分别为 $-1.3 \text{ m} \cdot \text{s}^{-1}$ 和 $-0.6 \text{ m} \cdot \text{s}^{-1}$,小于 ERA5 在此时的风速偏差(分别为 $-6.9 \text{ m} \cdot \text{s}^{-1}$ 和 $-5.0 \text{ m} \cdot \text{s}^{-1}$)。对于第二检验组,模型估算的风速同样整体优于 ERA5,前者平均风速和阵风风速平均绝对误差分别为 $1.1 \text{ m} \cdot \text{s}^{-1}$ 和 $1.5 \text{ m} \cdot \text{s}^{-1}$,后者为 $2.1 \text{ m} \cdot \text{s}^{-1}$ 和 $1.8 \text{ m} \cdot \text{s}^{-1}$,在最大风速时刻(21 日 23 时),前者的平均风速和阵风风速误差为 $-1.2 \text{ m} \cdot \text{s}^{-1}$ 和 $-3.1 \text{ m} \cdot \text{s}^{-1}$,后者为 $-3.2 \text{ m} \cdot \text{s}^{-1}$ 和 $-3.6 \text{ m} \cdot \text{s}^{-1}$ 。综上所述,SVM 回归模型估算风速较准确地反映了此次冷空气和温带气旋造成的两个检验组海上站点的大风风速变化,且准确率整体优于 ERA5。

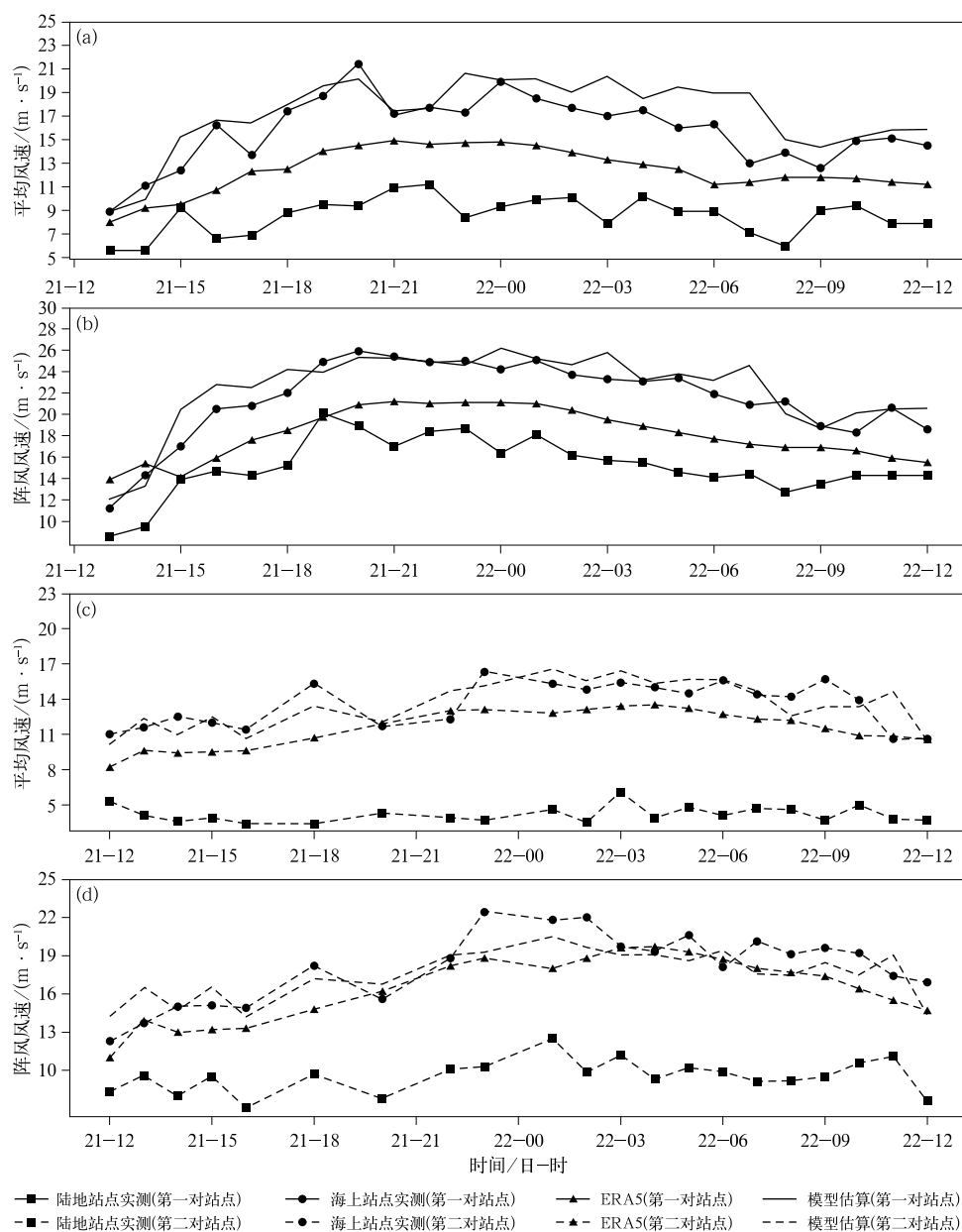


图9 2021年11月21日12时至22日12时(a,b)第一检验组和(c,d)第二检验组站点的陆地实测风速、海上实测风速、ERA5风速以及模型估算风速对比
(a,c)平均风速,(b,d)阵风风速

Fig. 9 Comparison of the measured land wind speed, the measured sea surface wind speed, ERA5 wind speed and the model estimated wind speed of (a, b) the first and (c, d) the second test groups from 12:00 BT 21 to 12:00 BT 22 November 2021
(a, c) mean wind speed, (b, d) gust speed

5 结论与讨论

利用位于我国北部沿岸的两组浮标及其邻近陆地观测站2016—2020年逐小时平均风速和阵风风速数据,统计分析了海陆风速差异特征及规律,基于

SVM回归方法,使用沿岸陆地平均风速及阵风风速、海陆站点距离、月份、时次作为估算因子组合,构建了沿岸海上平均风速和阵风风速估算模型,并使用另外两组海陆站点的2021年数据对估算结果进行检验。主要得到以下结论:

(1)海陆下垫面差异对平均风速分布的影响显

著大于阵风风速,两组站点的海陆平均风速的线性拟合斜率分别为 1.88 和 1.48,海陆阵风风速拟合斜率分别为 1.32 和 1.10;海陆风速差异具有显著的季节变化和日变化,第一训练组站点,不同季节的海陆平均风速和阵风风速比值极差分别为 0.61 和 0.43,不同时次的海陆平均风速和阵风风速比值极差分别为 0.73 和 0.44,第二训练组情况类似;另外,相同平均风速情况下,陆地阵风因子较海上阵风因子更大且分布更离散。

(2)使用 2021 年两个检验组数据对模型估算结果进行检验发现,对于 6 级及以上的平均风速和 7 级及以上的阵风风速,模型具有较高准确率,且优于 ERA5 计算结果。对于两个检验组海上站点 6 级及以上的平均风速,模型估算的 RMSE 分别为 $2.40 \text{ m} \cdot \text{s}^{-1}$ 和 $2.35 \text{ m} \cdot \text{s}^{-1}$,相较 ERA5 分别减少了 24%和 23%;对于两个检验组海上站点 7 级及以上的阵风风速,模型估算的 RMSE 分别为 $3.20 \text{ m} \cdot \text{s}^{-1}$ 和 $2.57 \text{ m} \cdot \text{s}^{-1}$,较 ERA5 分别减少了 14%和 20%。对一次温带气旋和冷空气共同影响的大风过程进行检验发现,模型估算风速较准确地反映出了两个海上站点的大风风速变化。此次过程中,模型估算的两个检验组海上站点的平均风速的平均绝对误差分别为 $1.6 \text{ m} \cdot \text{s}^{-1}$ 和 $1.1 \text{ m} \cdot \text{s}^{-1}$,阵风风速平均绝对误差分别为 $2.3 \text{ m} \cdot \text{s}^{-1}$ 和 $1.5 \text{ m} \cdot \text{s}^{-1}$,均优于 ERA5 计算结果。

通过一系列分析,本文统计了海陆下垫面差异导致的海陆风速差异特征及规律,并使用 SVM 回归模型,以沿岸陆地平均风速及阵风风速、海陆站点距离、月份、时次为估算因子组合,较准确地估算了沿岸海上大风。为使用丰富的沿岸陆地风速观测数据推测沿岸海上大风,从而弥补海上观测资料不足,提供了一种可靠、便于应用的方法。值得注意的是,受制于海上观测资料有限,以及海陆站点距离和海拔高度的要求,本文选取的站点均位于我国北部沿岸,因此研究结论仅适用于我国北部沿岸地区。另外,基于统计结果,将月份、时次作为离散的分类型预报因子加入模型中并不具备物理意义。随着海上观测数据的不断丰富,后期可开展我国近海其他海区的海陆风速差异特征研究,并进一步分析不同天气系统下的海陆风速差异特征,深入分析其背后的物理机制,结合物理机制和统计学规律,构建针对我国近海全海域且更加准确的估算模型。

参考文献

- 薄文波,高山红,王永明,2013. 基于 WRF 模式的渤海海面风的预报与订正[J]. 海洋湖沼通报, (3): 37-44. Bo W B, Gao S H, Wang Y M, 2013. Forecast and correction of the Bohai Sea surface wind based on WRF model[J]. Trans Oceanol Limnol, (3): 37-44(in Chinese).
- 陈洪滨,李军,马舒庆,等,2019. 海洋气象观测技术研发进展[J]. 科技导报, 37(6): 91-97. Chen H B, Li J, Ma S Q, et al, 2019. Progress of the marine meteorological observation technologies[J]. Sci Technol Rev, 37(6): 91-97(in Chinese).
- 丁世飞,齐丙娟,谭红艳,2011. 支持向量机理论与算法研究综述[J]. 电子科技大学学报, 40(1): 2-10. Ding S F, Qi B J, Tan H Y, 2011. An overview on theory and algorithm of support vector machines[J]. J Univ Electron Sci Technol China, 40(1): 2-10(in Chinese).
- 胡波,2019. 阵风因子与大气边界层要素的关系及预报试验[J]. 气象科技, 47(2): 282-288. Hu B, 2019. Relationship between gust factors and ABL parameters and its prediction test[J]. Meteor Sci Technol, 47(2): 282-288(in Chinese).
- 胡海川,刘珺,林建,2022. 基于预报方程的我国近海阵风预报[J]. 气象, 48(3): 334-344. Hu H C, Liu J, Lin J, 2022. Application of prediction equation to gust forecasting for Chinese offshore areas [J]. Meteor Mon, 48(3): 334-344(in Chinese).
- 胡海川,赵伟,董林,2021. 概率密度匹配方法在我国近海海面 10 m 风速预报中的应用[J]. 热带气象学报, 37(1): 91-101. Hu H C, Zhao W, Dong L, 2021. Application of probability density function matching in the offshore 10 m wind speed forecasting in China[J]. J Trop Meteor, 37(1): 91-101(in Chinese).
- 黄威,牛若芸,2017. 基于集合预报和支持向量机的中期强降水集成预报试验[J]. 气象, 43(9): 1110-1116. Huang W, Niu R Y, 2017. The medium-term multi-model integration forecast experimentation for heavy rain based on support vector machine[J]. Meteor Mon, 43(9): 1110-1116(in Chinese).
- 季晓阳,吴辉庭,杨学联,2005. 海面风场数值预报的历史和现状[J]. 海洋预报, 22(S1): 167-171. Ji X Y, Wu H D, Yang X L, 2005. A history and current situation on numerical prediction of sea surface wind field[J]. Mar. Forec, 22(S1): 167-171(in Chinese).
- 井传才,唐万林,王建平,等,1995. 青岛近海海陆大风的对比分析[J]. 黄渤海海洋, 13(4): 17-22. Jing C C, Tang W L, Wang J P, et al, 1995. A contrastive analysis of gale on the sea and the land of Qingdao Coastal Area[J]. J Oceanogr Huanghai Bohai Seas, 13(4): 17-22(in Chinese).
- 李敏,王辉,金敬华,2009. 中国近海海面风场预报方法综述[J]. 海洋预报, 26(3): 114-120. Li M, Wang H, Jin Q H, 2009. A review on the forecast method of China offshore wind[J]. Mar. Forec, 26(3): 114-120(in Chinese).
- 林良勋,程正泉,张兵,等,2004. 完全预报(PP)方法在广东冬半年海面强风业务预报中的应用[J]. 应用气象学报, 15(4): 485-493. Lin L X, Cheng Z Q, Zhang B, et al, 2004. The application of perfect prognosis method to gale wind along the Guangdong

- Coast in winter[J]. J Appl Meteor Sci, 15(4):485-493(in Chinese).
- 刘京雄,唐文伟,朱持则,等,2004. 浙闽沿海和台湾海峡海域冬季大风风速计算方法探讨[J]. 台湾海峡,23(1):8-13. Liu J X, Tang W W, Zhu C Z, et al, 2004. Discussion on calculation method of strong wind velocity on Zhejiang-Fujian coasts and Taiwan Strait in winter[J]. J Oceanogr Taiwan Strait, 23(1):8-13(in Chinese).
- 吕美仲,侯志明,周毅,2004. 动力气象学[M]. 北京:气象出版社. Lü M Z, Hou Z M, Zhou Y, 2004. Dynamic Meteorology[M]. Beijing: China Meteorological Press(in Chinese).
- 聂高臻,黄彬,2022. 2021年秋季海洋天气评述[J]. 海洋气象学报, 42(1):74-82. Nie G Z, Huang B, 2022. Autumn 2021 marine weather review[J]. J Mar Meteor, 42(1):74-82(in Chinese).
- 潘静,陈峥蓉,黄燕波,等,2022. 基于MOS的广西北部湾沿海秋冬极大风速精细化预报[J]. 气象研究与应用, 43(1):26-30. Pan J, Chen Z R, Huang Y B, et al, 2022. Refined prediction of maximum wind speed in autumn and winter along the coast of Guangxi Beibu Gulf based on MOS[J]. J Meteor Res Appl, 43(1):26-30(in Chinese).
- 钱燕珍,孙军波,余晖,等,2012. 用支持向量机方法做登陆热带气旋站点大风预报[J]. 气象, 38(3):300-306. Qian Y Z, Sun J B, Yu H, et al, 2012. Application of SVM method to the station strong wind forecast in landfalling tropical cyclones[J]. Meteor Mon, 38(3):300-306(in Chinese).
- 施萧,徐幼平,胡邦辉,等,2012. 支持向量机在雷暴预报中的应用[J]. 气象, 38(9):1115-1120. Shi X, Xu Y P, Hu B H, et al, 2012. Application of support vector machine to thunderstorm forecasting[J]. Meteor Mon, 38(9):1115-1120(in Chinese).
- 王波,李民,刘世萱,等,2014. 海洋资料浮标观测技术应用现状及发展趋势[J]. 仪器仪表学报, 35(11):2401-2414. Wang B, Li M, Liu S X, et al, 2014. Current status and trend of ocean data buoy observation technology applications[J]. Chin J Sci Instrum, 35(11):2401-2414(in Chinese).
- 吴曼丽,陈宇,王瀛,等,2012. 黄渤海北部沿海大风时空变化特征[J]. 气象与环境学报, 28(6):65-71. Wu M L, Chen Y, Wang Y, et al, 2012. The temporal and spatial characters of gale in the northern coast of the Yellow Sea and the Bohai Sea, China[J]. J Meteor Environ, 28(6):65-71(in Chinese).
- 许小峰,顾建峰,李永平,2009. 海洋气象灾害[M]. 北京:气象出版社. Xu X F, Gu J F, Li Y P, 2009. Marine Meteorological Disaster[M]. Beijing: China Meteorological Press(in Chinese).
- 杨璐,韩丰,陈明轩,等,2018. 基于支持向量机的雷暴大风识别方法[J]. 应用气象学报, 29(6):680-689. Yang L, Han F, Chen M X, et al, 2018. Thunderstorm gale identification method based on support vector machine[J]. J Appl Meteor Sci, 29(6):680-689(in Chinese).
- 尹尽勇,徐晶,曹越男,等,2012. 我国海洋气象预报业务现状与发展[J]. 气象科技进展, 2(6):17-26. Yin J Y, Xu J, Cao Y N, et al, 2012. The development and current status of marine weather forecasting operation in China[J]. Adv Meteor Sci Technol, 2(6):17-26(in Chinese).
- 张新玲,吴增茂,1998. 渤海海上测风与沿岸实测风的对比分析[J]. 海洋预报, 15(4):25-32. Zhang X L, Wu Z M, 1998. Comparing analyses of the observations of sea and coastal wind of Bohai[J]. Marin Forec, 15(4):25-32(in Chinese).
- 赵鸣,苗曼倩,1992. 大气边界层[M]. 北京:气象出版社. Zhao M, Miao M Q, 1992. Atmospheric Boundary Layer[M]. Beijing: China Meteorological Press(in Chinese).
- Hersbach H, Bell B, Berrisford P, et al, 2018. ERA5 hourly data on single levels from 1959 to present[Z]. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). DOI: 10.24381/cds.adbb2d47.
- Kohavi R, 1995. A study of cross-validation and bootstrap for accuracy estimation and model selection[C]// Proceedings of the 14th International Joint Conference on Artificial Intelligence-Volume 2. Montreal: Morgan Kaufmann Publishers Inc; 1137-1143.
- Lin G F, Chen G R, Wu M C, et al, 2009. Effective forecasting of hourly typhoon rainfall using support vector machines[J]. Water Resour Res, 45(8):W08440.
- Mercer A, Dyer J, Zhang S, 2013. Warm-season thermodynamically-driven rainfall prediction with support vector machines[J]. Proc Comput Sci, 20:128-133.
- Monahan H H, Armendariz M, 1971. Gust factor variations with height and atmospheric stability[J]. J Geophys Res, 76(24):5807-5818.
- Mu M, Duan W S, Wang J C, 2002. The predictability problems in numerical weather and climate prediction[J]. Adv Atmos Sci, 19(2):191-204.
- Nakamura K, Kershaw R, Gait N, 2007. Prediction of near-surface gusts generated by deep convection[J]. Meteor Appl, 3(2):157-167.
- Nayak M A, Ghosh S, 2013. Prediction of extreme rainfall event using weather pattern recognition and support vector machine classifier[J]. Theor Appl Climatol, 114(3-4):583-603.
- Pour S H, Shahid S, Chung E S, et al, 2018. Model output statistics downscaling using support vector machine for the projection of spatial and temporal changes in rainfall of Bangladesh[J]. Atmos Res, 213:149-162.
- Toth Z, Zhu Y J, Marchok T, 2001. The use of ensembles to identify forecasts with small and large uncertainty[J]. Wea Forecasting, 16(4):463-477.

(本文责编:何晓欢)