

渠鸿宇,董林,马新野,等,2024. 基于 XGBoost 的西北太平洋台风快速增强预报模型[J]. 气象,50(12):1531-1541. Qu H Y, Dong L, Ma X Y, et al, 2024. Forecast model of Northwest Pacific typhoon rapid intensification based on XGBoost[J]. Meteor Mon, 50(12):1531-1541(in Chinese).

基于 XGBoost 的西北太平洋台风快速增强预报模型*

渠鸿宇¹ 董 林¹ 马新野² 向纯怡¹ 黄奕武¹

¹ 国家气象中心,北京 100081

² 中科星图维天信科技股份有限公司,北京 101399

提 要: 台风强度预报,特别是台风快速增强(rapid intensification, RI)预报仍是目前台风预报中非常棘手的问题。基于 XGBoost 模型,利用 2015—2020 年 NCEP GFS 分析和预报数据以及 IBTrACS 数据分别构建了西北太平洋台风未来 24 h 的 RI 预报模型(FM)和预报订正模型(FCM)。通过对 FM 进行预报因子贡献分析发现,对模型预报影响最大的 5 个因子依次为台风丰满度、200 hPa 平均温度、过去 6 h 的强度变化、潜势强度和 200 hPa 平均散度。利用 2021—2022 年数据对模型进行独立检验,结果表明:FM 在利用分析数据测试时,具有较高的准确度,漏报率(FNR)、空报率(FPR)和 TS 分别为 0.25、0.24 和 0.32。但由于预报因子预报误差的影响,使得 FM 在实时预报中的性能下降(FNR、FPR 和 TS 分别为 0.32、0.26 和 0.27),而使用预报数据构建的 FCM 则可以通过学习预报误差进行有效订正,从而有效减小预报误差的影响。FCM 在实时预报检验中的 FNR、FPR 和 TS 分别为 0.28、0.25 和 0.30,相较 FM, FNR 和 FPR 分别降低了 0.04 和 0.01, TS 提升了 0.03。FCM 方便易用,可为台风强度和台风 RI 的实时预报提供参考。

关键词: 台风,快速增强(RI),XGBoost,台风强度

中图分类号: P456, P457

文献标志码: A

DOI: 10.7519/j.issn.1000-0526.2024.041101

Forecast Model of Northwest Pacific Typhoon Rapid Intensification Based on XGBoost

QU Hongyu¹ DONG Lin¹ MA Xinye² XIANG Chunyi¹ HUANG Yiwu¹

¹ National Meteorological Centre, Beijing 100081

² Zhongke Xingtuwei Tianxin Technology Co. Ltd., Beijing 101399

Abstract: The forecast of typhoon intensity, especially the rapid intensification (RI) forecast, is still a very challenging difficulty in current typhoon forecasting. Based on the XGBoost model, this article uses the NCEP GFS analysis and forecast data in 2015—2020, and IBTrACS data to construct RI forecast model (FM) and forecast correction model (FCM) for typhoons in the Northwest Pacific 24 h in advance. Through predictor contribution analysis of the FM, we have found that the five factors that have the greatest impact on model forecasts are typhoon abundance, average temperature at 200 hPa, intensity changes over the past 6 h, potential intensity, and average divergence at 200 hPa. The model is independently tested with the data in 2021—2022, and the results show that the FM has higher accuracy when tested by analytical data, with false negative rate (FNR), false positive rate (FPR) and threat score (TS) being 0.25, 0.24 and 0.32, respectively. However, due to the influence of forecast errors caused by forecast factors, the performance of FM in real-time forecasting decreases (FNR, FPR and TS are 0.32, 0.26 and 0.27,

* 国家重点研发计划(2023YFC3107902)、中国气象局创新发展专项研发项目(CXFZ2024J006)共同资助

2024 年 2 月 5 日收稿; 2024 年 8 月 14 日收修定稿

第一作者:渠鸿宇,主要从事台风与海洋气象预报及研究. E-mail:quhy@cma.gov.cn

通讯作者:董林,主要从事台风预报及研究. E-mail:donglin@cma.gov.cn

respectively). The FCM constructed based on forecast data can effectively correct the forecast errors by learning them, thereby reducing the impact of forecast errors. The FNR, FPR and TS of the FCM in real-time forecasting tests are 0.28, 0.25 and 0.30, respectively; compared with the FM, the FNR and FPR are reduced by 0.04 and 0.01, but the TS rises by 0.03. Thus, the FCM is convenient and easy to use, and can provide reference for real-time forecasting of typhoon intensity and typhoon RI.

Key words: typhoon, rapid intensification (RI), XGBoost, typhoon intensity

引 言

我国位于亚洲东南部,紧邻西北太平洋,拥有狭长的海岸线,是全球受台风(全文台风泛指所有海域的热带气旋)影响最严重的国家之一,平均每年有 7 个台风登陆我国(端义宏等,2012)。这些台风带来的狂风、暴雨、风暴潮以及其他次生灾害,严重威胁着我国沿海省份的生命和财产安全。因此,台风的防灾减灾工作尤为重要,而准确的台风路径和强度预报是关键。

近年来,我国台风路径预报取得长足进步,已达到世界前列水平(许映龙等,2010;端义宏等,2020)。2021 年中央气象台对西北太平洋及南海台风 24 h 预报时效的主观路径预报误差仅为 76 km。相对于台风路径预报,台风强度预报技巧的提升较为缓慢,但也取得一定进步。2021 年中央气象台 24 h 时效主观强度预报平均误差为 $3.6 \text{ m} \cdot \text{s}^{-1}$,较 2016—2020 年提高了 10.7%,但台风的快速增强(rapid intensification, RI)预报仍然是台风预报中的一大难题(杨梦琪等,2023;黄燕燕等,2024)。特别是在近海出现 RI 的台风更容易造成巨大损失,这让准确的台风 RI 预报显得更加重要(向纯怡等,2022)。

台风的强度变化受多种因素影响,包括大尺度环境场、台风内核结构、海气相互作用等(Kaplan and DeMaria, 2003; Kaplan et al, 2015)。传统的数值模式受低分辨率、初始误差和不完善的物理参数化方案限制,在描述台风发展的动力和热力过程方面存在一定的局限性,导致数值模式预报台风强度快速变化的能力不够理想(Mercer and Grimes, 2017; Rozoff et al, 2015)。因此,构建基于统计动力方法的台风 RI 预报模型成为一种有效的解决方案。Kaplan and DeMaria(2003)提出了首个 RI 预报概率模型,该模型利用 5 个 RI 台风与非 RI 台风有显著差异的预报因子,通过简单的阈值方法,预报北大西洋台风未来 24 h 发生 RI 的概率。随后,该模型在 2010 年进行了升级,除使用更多的预报因子

外,还采用了线性判别分析方法,并实现了对东北太平洋台风 RI 的预报(Kaplan et al, 2010)。然而,由于台风 RI 是一个复杂的非线性过程,通过线性模型预报台风 RI 准确度有限(Mercer and Grimes, 2017)。为了解决这一问题, Rozoff et al(2015)使用贝叶斯推断和逻辑回归方法构建了北大西洋和东太平洋的 RI 预报模型,有效提高了预报准确度。之后, Mercer and Grimes(2017)利用机器学习优势,使用支持向量机、全连接神经网络和随机森林构建了北大西洋的台风 RI 预报模型,进一步提升了预报准确度。Zhou et al(2022)基于历史卫星数据,利用 ResNet 模型和 LSTM 模型,构建了西太平洋台风 RI 预报模型,该模型利用历史 24 h 的卫星实况数据对未来 12 h 是否发生 RI(台风在未来 12 h 内是否增强 $7 \text{ m} \cdot \text{s}^{-1}$ 以上)进行外推预报,具有不错的准确度。然而,这种数据驱动的模式对数据量的要求很高;在仅依赖历史信息进行外推预报的情况下,随着预报时效的延长,预报误差往往会显著增大;此外,深度学习模型在解释性方面存在不足,应用时只能得到预测结果,而无法深入理解其背后的原因和机制。

本文针对西北太平洋的台风 RI 预报模型进一步深入研究。通过利用台风最佳路径数据和数值模式预报数据,不仅考虑了台风的历史信息,还纳入了大气和海洋在未来 24 h 内的状态变化。在此基础上,结合最新的台风强度机理研究成果,运用动力统计方法,显式地生成多个与台风强度变化或台风 RI 密切相关的预报因子。通过机器学习方法,构建西北太平洋台风 RI 预报模型,从而为西北太平洋台风 RI 的预报提供参考。

1 数据和方法

1.1 RI 事件的定义

本文沿用 Kaplan 方法(Kaplan and DeMaria, 2003; 谢礼江等, 2013),将 24 h 强度变化的第 95%

分位数($15 \text{ m} \cdot \text{s}^{-1}$)作为判断台风 RI 的阈值。若自某一个时刻起的未来 24 h 台风近中心最大平均风速增强了 $15 \text{ m} \cdot \text{s}^{-1}$ 以上,则定义该时刻为一个 RI 事件(RIE),否则为非 RI 事件(NRIE)。台风在增强过程中可能会发生多次 RIE,比如次日 00 时、06 时、12 时、18 时(世界时,下同)较今天的 00 时、06 时、12 时、18 时均发生了 $15 \text{ m} \cdot \text{s}^{-1}$ 以上的增强,则今天连续发生了 4 次 RIE,只要发生 1 次 RIE,就称该台风发生了 RI。

1.2 数据

本文使用的数据包括 2015—2022 年 WMO 的台风最佳路径数据集 IBTrACS (Knapp et al, 2010),以及 NCEP GFS 的分析和预报数据(National Centers for Environmental Prediction et al, 2015),上述数据共同生成台风 RI 预报模型的预报因子和因变量(或预报量,此处指未来 24 h 是否发生 RIE)。

1.3 预报因子

GFS 的空间分辨率为 $0.25^\circ \times 0.25^\circ$,本文使用 GFS 00 时、06 时、12 时、20 时起报的 00 时分析数据以及 06 时、12 时、18 时、24 时预报数据生成基于模式数据的预报因子。由于预报因子不仅表征当前状态,还表征了未来的变化,将其称之为动态预报因子(表 1)。这些因子均为影响台风 RI 变化的重要因素,其中大多数已应用在上述东北太平洋和北大西洋台风 RI 预报模型中。在计算这些预报因子之前首先需要确定 GFS 不同预报时次的台风中心位置,这里采用最大环流法(Nguyen et al, 2014)。根据斯托克斯定理,单位面积的环流最大值即单位面积上相对涡度最大值,因此取距台风中心 100 km 范围内的平均相对涡度最大值作为台风中心。确定了台风中心就可以得出表 1 中的预报因子,下文对这些预报因子做详细介绍。

台风的结构是影响台风强度变化的重要因素。Guo and Tan (2017)提出了台风丰满度(TCF)概念,并指出台风丰满结构($\text{TCF} > 0.8$)是强台风的必然条件,且 TCF 的增加与台风增强关系密切。TCF 计算如式(1)所示:

$$\text{TCF} = 1 - \frac{R_{\text{mw}}}{R_{17}} \quad (1)$$

式中: R_{mw} 表示最大风圈半径, R_{17} 表示 $17 \text{ m} \cdot \text{s}^{-1}$ 风圈半径。通过 GFS 10 m 风场的分析和预报数据计算得到的 TCF 虽然与真值有偏差,但仍能得到与

Guo and Tan (2017)相同的台风强度越强 TCF 越高的结论(图略),因此可将 TCF 作为台风 RI 预报模型的预报因子之一。

海面温度(SST)是海洋热状况的表征,较高的 SST 可通过释放更多的感热和潜热通量促进台风增强(Emanuel, 1986; Holland, 1997)。最大潜势强度(MPI)用于表征台风强度发展的上限,这里使用基于 SST 的经验公式[式(2)]计算获得(Levitus, 1983):

$$\text{MPI} = (A + B e^{C(T - T_0)}) \times 0.5144 \quad (2)$$

式中: $A = 38.21 \text{ kt}$, $B = 170.72 \text{ kt}$, $C = 0.1909 \text{ } ^\circ\text{C}^{-1}$, $T_0 = 30.0^\circ\text{C}$,最后一项乘以 0.5144 将速度单位 kt 转换为 $\text{m} \cdot \text{s}^{-1}$ 。潜势强度(POT)为 MPI 减去台风当前强度,代表当前台风强度可发展的空间。

水平风垂直切变(VWS)在台风生成和加强过程中扮演着重要角色,是影响台风发展的重要动力学因子之一。一般认为在较弱环境 VWS 下台风更容易发生 RI (Gray, 1968; Merrill, 1988; 高拴柱和吕心艳, 2024)。然而,谢礼江等(2013)的研究表明,即使在较大的环境 VWS 情况下,仍有 9.7% 的台风能够发生 RI。为了探究这一现象的原因,Shi and Chen (2021)利用合成分析方法对中等及强环境 VWS 下 RI 发生的机制进行了深入研究。研究结果显示,中等和强环境 VWS 下发生 RI 台风的高层会出现明显的逆风切出流,这种高空出流阻塞了原本的高层环流,从而为台风增强创造了有利的局地弱 VWS 环境。因此,本文将局地 VWS (LVWS, 距台风中心 200 km 内的平均 VWS) 和环境风 VWS (EVWS, 距台风中心 400~800 km 的平均 VWS) 均作为预报因子加入模型。

台风的高空由于惯性稳定性远低于中低层,容易与大尺度场发生相互作用。相对角动量涡旋辐合(REFC)常用于表征高空大尺度场和台风的相互作用,REFC 的计算如式(3)所示:

$$\text{REFC} = -\frac{1}{r^2} \frac{\partial}{\partial r} (r^2 \overline{U'V'}) \quad (3)$$

式中: U 是径向风, V 是切向风, r 是半径,上横线表示方位角平均值,撇号表示与方位角平均值的差。Kaplan and DeMaria (2003)指出当 REFC 值较高时,台风不太可能发生 RI。另外,高空出流也是影响台风发展的重要动力条件,本文将距台风中心 800 km 内的 200 hPa 平均散度(D200)作为高空出流指标加入模型。

此外,还加入了距台风中心 200~800 km 的 200 hPa 平均温度(T_{200})和距台风中心 200~800 km 的

850~500 hPa 平均相对湿度(RHLO),有研究指出当前者较低、后者较高时有利于台风增强(Knaff et al,2005)。

IBTrACS 记录了西北太平洋台风逐 6 h 的台风近中心最大平均风速和经纬度等信息,使用该数

据生成表征台风当前状态的预报因子(称为静态预报因子):(1)台风当前最大平均风速(VMAX),(2)台风过去 6 h 的最大平均风速变化(DVMAX),(3)台风过去 6 h 的移动速度(SPD)(表 1)。

表 1 各预报因子的定义
Table 1 Definition of each forecast factor

序号	预报因子	定义	单位
1	TCF	台风丰满度	—
2	POT	潜势强度	$\text{m} \cdot \text{s}^{-1}$
3	EVWS	距台风中心 400~800 km 平均水平风垂直切变	$\text{m} \cdot \text{s}^{-1}$
4	LVWS	距台风中心 200 km 内平均水平风垂直切变	$\text{m} \cdot \text{s}^{-1}$
5	REFC	距台风中心 300~600 km 平均相对角动量涡旋辐合	$\text{m} \cdot \text{s}^{-1} \cdot \text{d}^{-1}$
6	D200	距台风中心 800 km 内 200 hPa 平均散度	10^5 s^{-1}
7	T200	距台风中心 200~800 km 的 200 hPa 平均温度	$^{\circ}\text{C}$
8	RHLO	距台风中心 200~800 km 的 850~500 hPa 平均相对湿度	%
9	VMAX*	台风当前近中心最大平均风速	$\text{m} \cdot \text{s}^{-1}$
10	DVMAX*	台风过去 6 h 的近中心最大平均风速变化	$\text{m} \cdot \text{s}^{-1}$
11	SPD*	台风过去 6 h 的移动速度	$\text{m} \cdot \text{s}^{-1}$

注: * 代表静态预报因子,否则代表动态预报因子;—代表无数据。

1.4 模 型

极端梯度提升(eXtreme Gradient Boosting, XGBoost;Chen et al,2015)是一种高效且可扩展的梯度提升决策树算法,通过迭代训练决策树模型,每棵树都试图纠正前面所有树的残差,从而逐步提升模型的预报性能。XGBoost 的目标函数如式(4)所示:

$$\Phi(t) = \sum_{j=1}^n l[y_j, f_{t-1}(x_j) + f_t(x_j)] + \Omega(f_t) \quad (4)$$

式中: $f_t(x)$ 表示第 t 棵树的预报结果, y_j 表示第 j 个样本的真实输出值, x_j 表示第 j 个样本的特征, $f_{t-1}(x)$ 表示第 $t-1$ 棵树的预报结果。右侧第一项为损失函数,用来衡量模型在第 t 次训练中的预报误差,对于回归问题,通常使用均方误差,对于二分类问题,通常使用交叉熵;第二项表示 t 棵树的复杂度,用来控制模型的复杂度和泛化能力,也被称作正则化项,通过引入正则化项可以有效防止过拟合现象。同时,XGBoost 还采用了特征子采样和数据子采样等技术,可以有效地解决数据倾斜和特征相关性的问题,从而进一步增强模型的鲁棒性和泛化性能。此外,XGBoost 通过并行化处理实现了高效的计算,适用于处理大规模数据集和高维特征空间。相比传统的梯度提升决策树算法,XGBoost 在训练

速度和预报性能上均有显著提升。由于 XGBoost 的众多优势,已广泛应用于分类、回归、排序、推荐系统等领域,包括气象领域(张钧民等,2021;杨绚等,2022)。本文使用 XGBoost 模型对台风未来 24 h 是否发生 RIE 进行预报,即基于 XGBoost 构建一个二分类模型。

1.5 检验指标

RIE 是小概率事件,因此台风 RIE 预报属于样本不平衡的二分类问题。在这种情况下,使用准确率(预报正确的样本占总样本的比例)难以评估模型的性能。因此,本文通过计算气象领域常用的漏报率(FNR)、空报率(FPR)和 TS 来检验模型预报准确度,使用 AUC(Area Under the Curve)评估模型性能以进行参数优化。如表 2 所示,TP 代表观测为 RIE 且预报也为 RIE 的样本数,TN 代表观测为 NRIE 且预报为 NRIE 的样本数,FN 代表观测为 RIE 而预报为 NRIE 的样本数,FP 代表观测为 NRIE 而预报为 RIE 的样本数。

表 2 台风 RIE 预报的混淆矩阵
Table 2 Confusion matrix for typhoon RI event prediction

事件		预报	
		RIE	NRIE
观测	RIE	TP	FN
	NRIE	FP	TN

FNR、FPR 和 TS 的计算如式(5)~式(7)所示。

$$\text{FNR} = \frac{\text{FN}}{\text{TP} + \text{FN}} \quad (5)$$

$$\text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}} \quad (6)$$

$$\text{TS} = \frac{\text{TP}}{\text{FP} + \text{TP} + \text{FN}} \quad (7)$$

式中:FNR 表示实际为 RIE 的样本中,模型错误预报为 NRIE 的样本所占比例,该指标越接近 0,表示模型对 RIE 捕捉能力越高,而越接近 1 则表明模型漏报了大量 RIE 样本。FPR 表示实际为 NRIE 的样本中,模型错误预报为 RIE 的样本所占的比例,该指标越接近 0,表明模型对 NRIE 的预报准确度越高,而越近 1 则表明模型空报了大量 NRIE 样本。TS 表示所有实际为 RIE 样本以及预报为 RIE 样本中,预报正确的 RIE 样本所占的比例,该指标越接近 1,表明模型对 RIE 的预报能力越好。

AUC 是一种常用的用于衡量二分类模型性能的评估指标,它通过绘制接收者操作特征曲线(ROC)来计算。ROC 以不同的分类阈值为基础,绘制模型的查全率(为 $1 - \text{FNR}$,即所有 RIE 样本中被模型正确预报的比例)与 FPR 之间的关系,即利用不同阈值来衡量在尽量提高查全率时会引起多少的空报率。AUC 则是 ROC 下方的面积,取值范围在 0~1,其中 0.5 表示模型的预测性能等同于随机猜测,而 1 表示模型的预测完全准确,因此 AUC 越接近 1,模型的性能越好。AUC 提供了一个单一的数值来概括分类器的性能,使得不同模型之间的比较变得简单;此外,AUC 不受样本不平衡的影响,因此在处理不平衡数据集时具有独特的优势。

2 模型构建

2.1 模型的输入和输出

如上文所述,模型的预报因子包括由 GFS 生成

的 8 个动态预报因子,以及由 IBTrACS 生成的 3 个静态预报因子。各动态预报因子为未来 24 h 逐 6 h 间隔的平均值。模型的输出向量为未来 24 h 是否发生 RIE。

2.2 数据集

在构建输入模型的数据集之前,首先对 2015—2022 年的台风数据进行如下过滤:(1)由于只关注台风是否发生 RI,同时为了减小样本不平衡对模型训练的负面影响,因此只保留每个台风首次达到热带风暴级到最后一次出现 24 h 增强的记录;(2)删除由于 GFS 数据缺测无法生成动态预报因子的样本;(3)剔除有效记录不满 24 h 的台风。经过上述筛选后,我们共得到了 173 个台风,2040 个样本,其中发生 RI 的台风数为 90 个,发生 RIE 的样本为 390 个,占总样本的为 19%(表 3)。不同台风等级下发生 RIE 的比例差别较大,其中比例最高的为台风级(TY),占比为 35%,其次为强热带风暴级(STS),而热带风暴级(TS)占比最少。使用 2015—2020 年的样本训练模型,并结合 Optuna 超参数搜索框架(Akiba et al, 2019)和分层交叉验证法(Kohavi, 1995)优化超参数,优化目标为 AUC 最大。使用 2021—2022 年的样本对模型进行独立性检验。训练集中包含 138 个台风,1609 个样本,其中发生 RI 的台风 74 个,发生 RIE 的样本 325 个,占总样本的 20%;测试集中包含 35 个台风,其中 16 个台风发生 RI,431 个总样本中有 65 个发生 RIE,占比为 15%。

2.3 试验设计

为了全面、系统地评估基于上述方法构建的 RI 预报模型的可行性,设计了三组试验:

Exp1:使用由分析数据生成动态预报因子构建的训练集(记为 PP 训练集)训练“完美假设”下的预报模型(记为 FM),并同时使用由分析数据生成的

表 3 数据集样本信息
Table 3 Sample information of dataset

统计项	训练集	测试集	总数据
台风数/个	138	35	173
发生 RI 的台风数/个	74	16	90
热带风暴级(TS)样本数/个,RI 样本数/个,RI 样本占比/%	657,66,10	201,13,6	858,79,9
强热带风暴级(STS)样本数/个,RI 样本数/个,RI 样本占比/%	307,97,32	81,15,19	388,112,29
台风级(TY)样本数/个,RI 样本数/个,RI 样本占比/%	262,88,34	51,23,45	313,111,35
强台风级(STY)样本数/个,RI 样本数/个,RI 样本占比/%	166,43,26	40,10,25	206,53,26
超强台风级(SuperTY)样本数/个,RI 样本数/个,RI 样本占比/%	217,31,14	58,4,7	275,35,13
总样本数/个,RI 样本数/个,RI 样本占比/%	1609,325,20	431,65,15	2040,390,19

测试集(记为 PP 测试集)进行检验。该试验假设预报因子本身没有误差,因此可用于评估和分析模型本身的性能。

Exp2:使用由预报数据生成的测试集(记为 FST 测试集)对 FM 进行预报检验。该试验旨在检验 FM 在实时预报中的准确度。

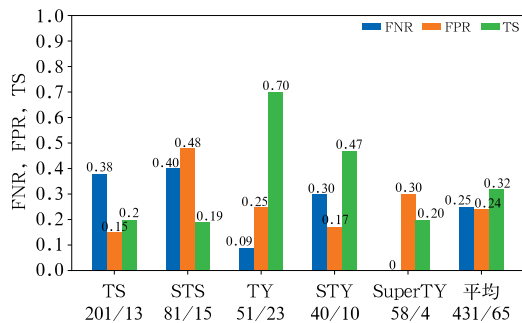
Exp3:考虑到动态预报因子本身预报误差对 FM 预报性能的影响,使用由预报数据生成的 FST 训练集直接训练预报订正模型(记为 FCM),并同样使用 FST 测试集进行检验。该试验旨在检验 FCM 是否能够学习动态预报因子的预报误差,从而减小动态预报因子误差对模型 RI 预报的影响。

为了避免超参数在比较 FM 和 FCM 时的影响,两组试验采用了相同的超参数搜索空间,并且迭代次数均为 1000 次。

3 检验和分析

3.1 “完美假设”下的模型评估和分析

图 1 显示了 FM 在 PP 测试集中的检验结果。由图可知,在动态预报因子无明显误差的假设下,模型具有相当不错的 RI 预报能力,平均的 FNR、FPR、TS 分别为 0.25、0.24、0.32。不同台风等级下的预报准确度有所差异,表现最好的为 TY 级, FNR、FPR 和 TS 分别为 0.09、0.25 和 0.70。表现最差的为 STS 级, FNR、FPR 和 TS 分别为 0.40、0.48 和 0.19。

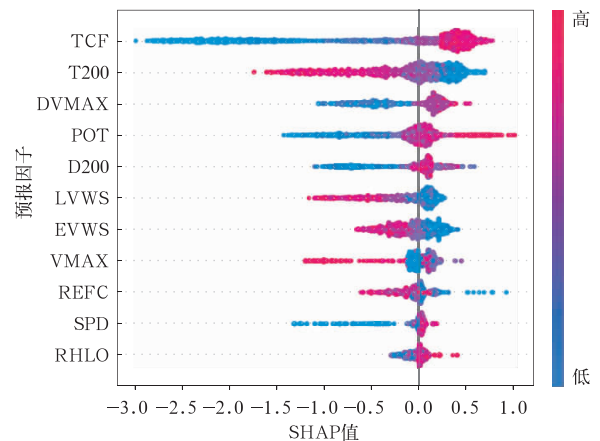


注:横坐标数字代表不同等级总样本数/
RIE 样本数(单位:个)。

图 1 FM 在 PP 测试集中的检验结果

Fig. 1 Verification results of the FM
in the PP test set

在“完美模型”假设下,可以粗略地认为预报因子不存在明显误差,这样便于使用 SHAP(Shapley Additive exPlanations)框架对 XGBoost 分类模型的预报进行分析。SHAP 利用合作博弈论方法为每个预报因子分配一个 SHAP 值,该值反映了其对预报结果的边际贡献,正值(负值)表明该预报因子增大(减小)了 RIE 发生的概率[更多详细原理参考 Lundberg et al(2020)]。图 2 展示了 PP 训练集各预报因子在 FM 中的 SHAP 值分布,其中填色代表各预报因子在标准化后的相对大小,从上到下的排列顺序则代表了各预报因子对预测结果的贡献程度。由图可见,TCF 对预报结果的影响最为显著,其次为 T200。图中,SPD 和 RHLO 的贡献虽然最小,但是通过敏感性试验证明舍弃其中任何一个都会引起误差有所增加。当 TCF 较大、T200 较小、POT 较大、D200 较大、LVWS 和 EVWS 较小、REFC 较小、RHLO 较大时,有利于 RIE 发生,这均与对台风强度发展的认知相一致。另外,DVMAX 较大时,其对应较大的正 SHAP 值,表明当 RIE 发生时,台风已经处于较强的发展阶段;VMAX 较大时,对应较大负 SHAP 值,表明若台风已经很强,则再次发生 RIE 的概率较低;SPD 越小时,其对应较大的负 SHAP 值,这或许与海温的负反馈有关,当移速较慢时更容易受海温负反馈影响,不利于台风增强。需要注意的是,每个因子对预报结果的贡献并不是简单的线性相关。比如 TCF,在其他因子取



注:填色代表各预报因子标准化后的相对大小。

图 2 PP 训练集各预报因子在 FM
中的 SHAP 值分布

Fig. 2 SHAP value distribution of each predictor
of PP training set in the FM

平均的情况下,TCF 与 FM 预报概率的曲线总体呈 S 型(图 3a),即 TCF 较低时对应较低的概率,随着 TCF 增大,RIE 概率逐渐增大并趋于稳定;VMAX 则大概呈 Ω 型(图 3b),这与不同台风等级下发生 RIE 的可能性大小相对应,STS 和 TY 等级下发生 RIE 的比例较高,其他等级则相对较小。

在上述预报因子贡献分析的基础上,可以对 STS 和 TY 等级下 FM 模型预报准确度差异明显的原因进行进一步分析。图 4 所示为 STS 和 TY 等级下 RIE 样本和 NRIE 样本 TCF、T200、DVMAX、POT 和 D200(重要程度最高的 5 个预报因子)的分布差异。由图可见,相较 NRIE 样本,TY 等级下的 RIE 样本的 T200 整体偏低、DVMAX 整体偏高,这些预报因子的显著差异有助于模型区分 RIE 和 NRIE;而 STS 等级下,T200 和 DVMAX 的差异则相对较小,且出现了 RIE 样本的 D200 整体低于 NRIE 样本的情况(与上文预报因子贡献分析不符)的情况,这些都使得模型准确预报 RI 的难度增大。这也在一定程度上说明,仅利用这 11 个预报因子构建模型进行 RI 预报仍存在一定的局限性。

3.2 实时预报检验

在实时业务预报中,由于无法获取未来时刻的分析数据,只能使用预报数据代替,FST 测试集中 FM 模型的检验结果如图 5a 所示,由图可见,受动态预报因子预报误差的影响,FM 在实时预报中的准确度有所降低。整体 FNR、FPR 和 TS 分别为 0.32、0.26 和 0.27,较在 PP 测试集中的结果,FNR、FPR 分别增加了 0.07、0.02,TS 降低 0.05。各等级下的预报准确度也都较在 PP 测试集中的结

果有所下降。

为降低预报因子误差对模型性能的影响,直接使用由预报场生成的 FST 训练集训练和优化台风 RI 预报模型,以求模型在训练的时候能够学习预报误差,从而使其具备一定的预报订正能力。图 5b 为 FCM 在 FST 测试集中的检验结果,整体 FNR、FPR 和 TS 分别为 0.28、0.25 和 0.30,相较 FM,FNR 和 FPR 分别减小了 0.04 和 0.01,TS 增加了 0.03。同时在大部分等级下的预报准确度均优于 FM。此外,模型对 RIE 的预报准确度远优于 NCEP GFS 和 ECMWF,同样本检验下 NCEP GFS 的 FNR、FPR 和 TS 分别为 0.90、0.01 和 0.09,ECMWF 则分别为 0.95、0.01、0.05。

选取 2116 号台风蒲公英和 2216 号台风奥鹿(图 6),对 FM 和 FCM 的实时预报能力进行个例检验,两个台风在发展过程中都经历了多次 RIE。

“蒲公英”于 2021 年 9 月 23 日在西北太平洋洋面生成,此后逐渐朝西北方向行进并持续发展增强,于 9 月 26 日发展为超强台风并达到强度峰值,在此期间,连续经历了 7 次 RIE(图 7a)。FM 和 FCM 都成功预报出了每一次 RIE,但在第一次 RIE 之前和最后一次 RIE 之后均出现了空报的现象,且 FM 较 FCM 多空报了一次。

“奥鹿”于 2022 年 9 月 23 日在菲律宾以东洋面生成,而后逐渐西行增强,并于 24 日开始出现爆发性增强,至 25 日发展为超强台风,在登陆并穿越菲律宾期间有所减弱,进入南海后其强度再度发展,于 27 日再次发展为超强台风。在“奥鹿”两次增强期间,发生了多次 RIE(图 7b)。与“蒲公英”相似,FCM 较成功地预报出两次发展期间发生的 RIE,仅

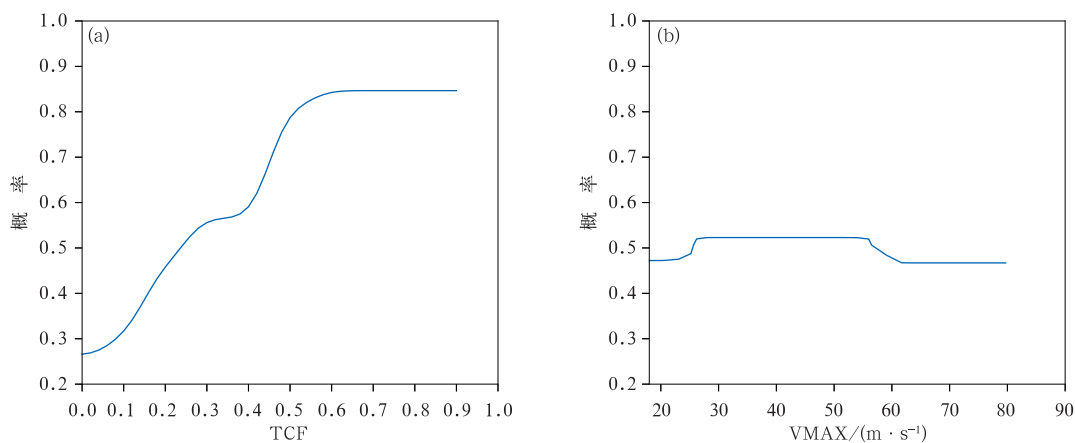
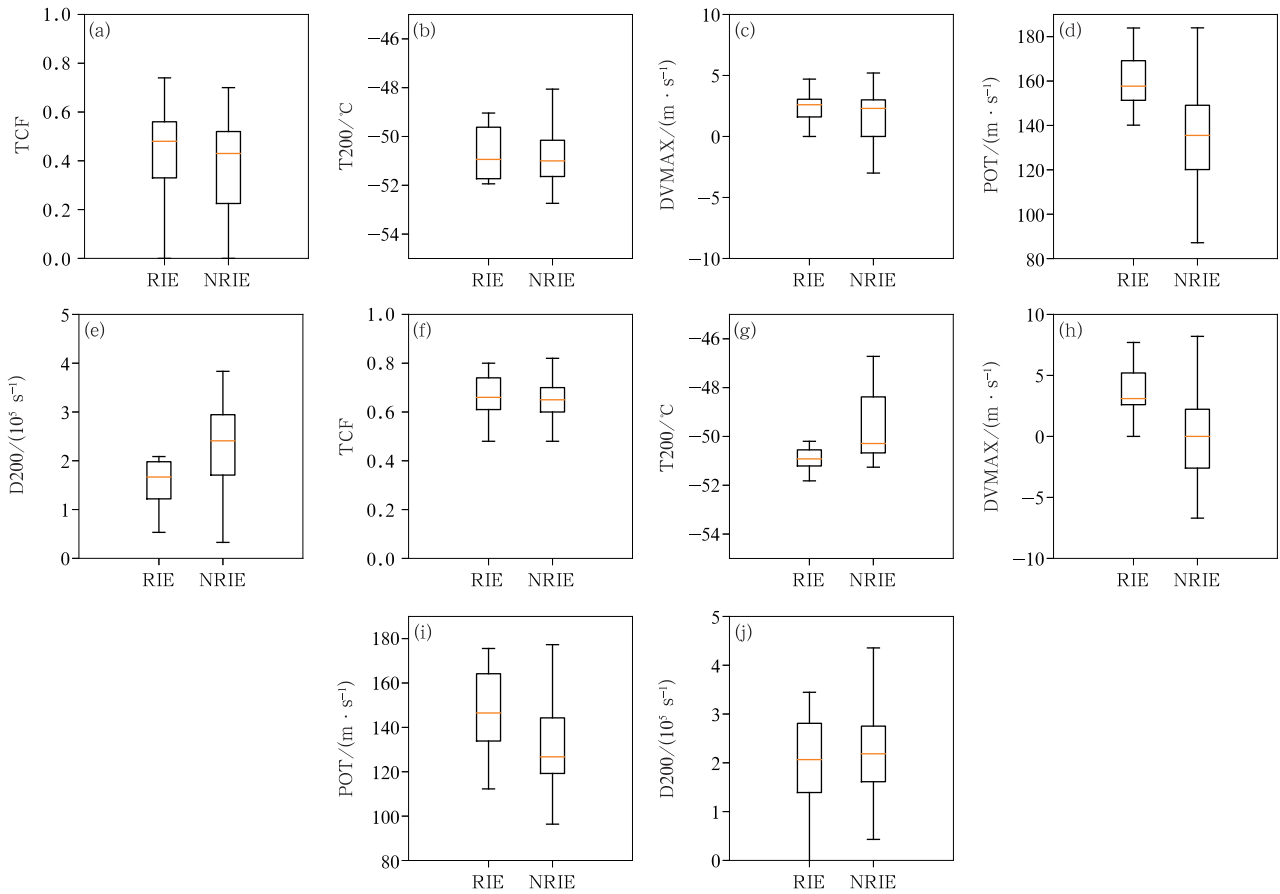


图 3 FM 不同(a)TCF 和(b)VMAX 的平均预报概率

Fig. 3 Average prediction probability of the FM with different (a) TCF and (b) VMAX

有两次漏报,但仍在 NRIE(RIE)和 RIE(NRIE)切换之前(之后)出现了两次空报现象。FM 对“奥鹿”

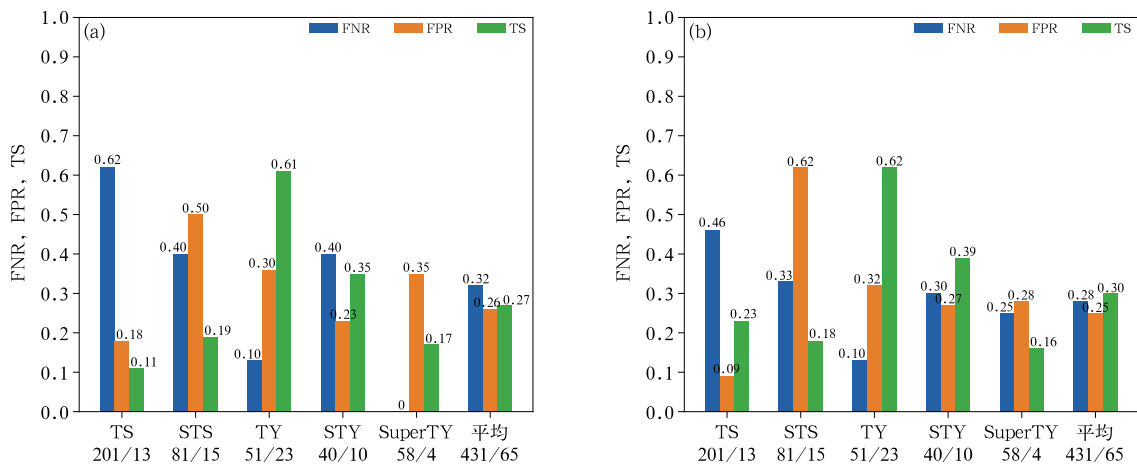
的 RI 预报则相对较差,除出现两次空报外,漏报了第一次快速发展期间的全部 RIE。



注:箱体上、下边界分别代表第 75% 和第 25% 分位数,箱体内横线代表中位数,须线上、下端分别代表最大和最小值(不包括异常值)。

图 4 PP 测试集中(a~e)STS 和(f~j)TY 等级下 RIE 样本和 NRIE 样本预报因子的分布差异

Fig. 4 Distribution differences of predictors of RIE samples and NRIE samples at (a-e) STS and (f-j) TY levels in the PP test set



注:横坐标中的数字代表不同等级的总样本数/RIE 样本数(单位:个)。

图 5 (a)FM 和(b)FCM 在 FST 测试集中的检验结果

Fig. 5 Verification results of the (a) FM and (b) FCM in the FST test set

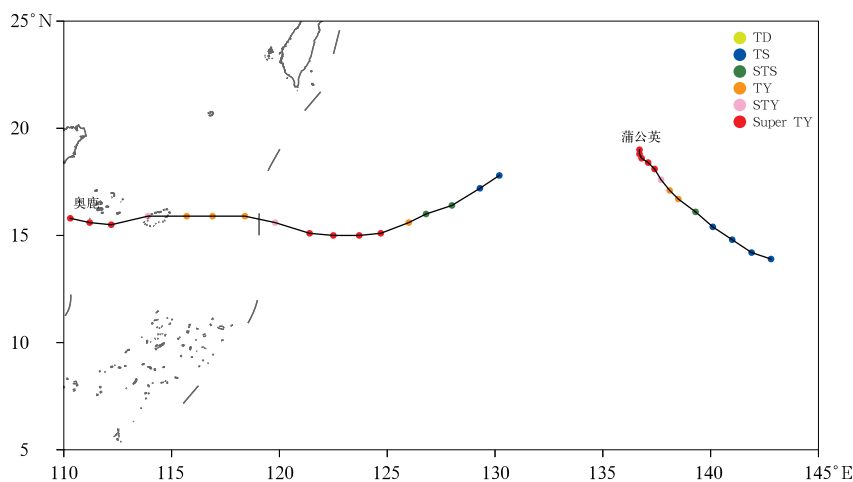


图 6 台风蒲公英和奥鹿的部分路径

Fig. 6 Partial tracks of Typhoon Mindulle and Typhoon Noru

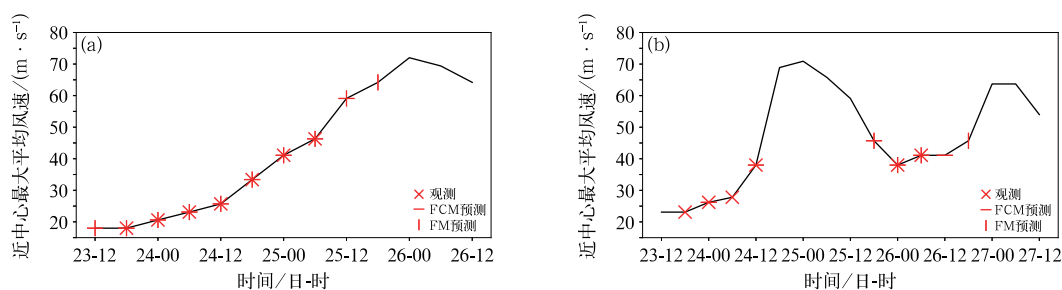


图 7 (a)2021 年 9 月 23 日 12 时至 26 日 12 时台风蒲公英和

(b)2022 年 9 月 23 日 12 时至 27 日 12 时台风奥鹿的强度变化(黑色实线)

Fig. 7 Intensity changes (solid black line) of (a) Typhoon Mindulle

from 12:00 BT 23 to 12:00 BT 26 September 2021,

(b) Typhoon Noru from 12:00 BT 23 to 12:00 BT 27 September 2022

综上所述,使用预报场作为输入会带来一定的误差,但通过训练模型来学习预报误差并进行订正,可以在一定程度上消除预报误差的影响来提高模型的预报准确度。但是,尽管 FCM 相较 FM 在实时预报中的准确度有所提升,但仍无法达到 FM 在 PP 测试集中的准确度。

4 结论与讨论

本文基于 XGBoost 模型,利用 2015—2020 年 NCEP GFS 分析和预报数据以及 IBTrACS 数据分别构建了西北太平洋台风未来 24 h 的 RI 预报模型(FM)和预报订正模型(FCM),并使用 2021—2022 年数据进行独立检验。主要得到以下结论:

(1)在“完美模型”假设下构建的 FM 具有较高的性能,2021—2022 年回算的 FNR 为 0.25, FPR

为 0.24, TS 为 0.32。不同台风等级下的预报准确度有所差异,表现最好的为 TY 等级(FNR、FPR 和 TS 分别为 0.09、0.25 和 0.70),表现最差的为 STS 等级(FNR、FPR 和 TS 分别为 0.40、0.48 和 0.19)。利用 SHAP 进行预报因子贡献分析发现,对 FM 预报影响最大的 5 个因子依次为 TCF、T200、DVMAX、POT 和 D200。

(2)在实时预报中,由于预报因子本身的误差,“完美模型”假设下构建的 FM 的预报性能会受到影响,整体 FNR、FPR 和 TS 分别为 0.32、0.26 和 0.27,较在 PP 测试集下的结果, FNR 和 FPR 分别增加了 0.07、0.02, TS 减小了 0.05。然而,通过训练模型学习这些误差并进行订正,可以在一定程度上消除这些影响,提高模型的实时预报准确度。FCM 的 FNR、FPR 和 TS 分别为 0.28、0.25 和 0.30,相较 FM, FNR 和 FPR 分别减小了 0.04 和

0.01, TS 增加了 0.03, 同时在大部分等级下的预报准确度均优于 FM。通过对比 FCM 和 FM 对 2116 号台风蒲公英和 2216 号台风奥鹿的实时预报准确度, 发现 FCM 相较 FM 具有更高的性能, 几乎捕捉到了全部的 RIE, 但在 NRIE(RIE) 和 RIE(NRIE) 切换之前(之后)出现了多次空报的现象。

综上所述, 本文利用 NCEP GFS 数据构建了可进行实时预报的西北太平洋 RI 预报模型, 并具有一定的性能, 能够为台风强度以及台风 RI 的实时预报提供参考。但是, 模型仍存在一些不足: 一方面, 模型对不同台风等级下预报准确的相差较大, 这其中背后的原因还需要进行进一步深入分析; 另一方面, 模型选取的预报因子不够丰富, 缺乏对海气相互作用的考虑, 且数据来源单一。未来, 会对台风不同阶段强度发展的主要影响因子进行进一步深入分析, 在机理研究的基础上采用更多的预报因子、多源的数据, 进一步改进优化模型。

参考文献

- 端义宏, 陈联寿, 许映龙, 等, 2012. 我国台风监测预报预警体系的现状及建议[J]. 中国工程科学, 14(9): 4-9. Duan Y H, Chen L S, Xu Y L, et al, 2012. The status and suggestions of the improvement in the typhoon observation, forecasting and warning systems in China[J]. Strateg Study CAE, 14(9): 4-9(in Chinese).
- 端义宏, 方娟, 程正泉, 等, 2020. 热带气旋研究和业务预报进展—第九届世界气象组织热带气旋国际研讨会(IWTC-9)综述[J]. 气象学报, 78(3): 537-550. Duan Y H, Fang J, Cheng Z Q, et al, 2020. Advances and trends in tropical cyclone research and forecasting: an overview of the ninth World Meteorological Organization International Workshop on Tropical Cyclones (IWTC-9) [J]. Acta Meteor Sin, 78(3): 537-550(in Chinese).
- 高拴柱, 吕心艳, 2023. 台风快速增强爆发前后的背景环流和热力动力条件的演变[J]. 气象学报, 81(5): 702-716. Gao S X, Lü X Y, 2023. Evolution of environmental circulation and dynamic and thermodynamic conditions before and after the onset of typhoon rapid intensification[J]. Acta Meteor Sin, 81(5): 702-716(in Chinese).
- 黄燕燕, 陈子通, 冯业荣, 等, 2024. 南海台风模式对台风利奇马快速增强预报能力研究[J]. 气象, 50(5): 532-546. Huang Y Y, Chen Z T, Feng Y R, et al, 2024. Research on the forecasting capability of CMA-TRAMS Model for rapid intensification of Typhoon Lekima[J]. Meteor Mon, 50(5): 532-546(in Chinese).
- 向纯怡, 许映龙, 高拴柱, 等, 2022. 2021 年西北太平洋台风活动特征和预报难点分析[J]. 气象, 48(9): 1195-1208. Xiang C Y, Xu Y L, Gao S Z, et al, 2022. Analysis of the characteristics and forecast difficulties of typhoon over the western North Pacific in 2021[J]. Meteor Mon, 48(9): 1195-1208(in Chinese).
- 谢礼江, 邱新法, 王伟, 2013. 西北太平洋热带气旋快速增强与环境垂直风切变统计分析[J]. 热带地理, 33(3): 242-249, 263. Xie L J, Qiu X F, Wang W, 2013. Rapid intensification of tropical cyclones and vertical wind shear over the Northwest-Pacific[J]. Trop Geogr, 33(3): 242-249, 263(in Chinese).
- 许映龙, 张玲, 高拴柱, 2010. 我国台风预报业务的现状及思考[J]. 气象, 36(7): 43-49. Xu Y L, Zhang L, Gao S Z, 2010. The advances and discussions on China operational typhoon forecasting[J]. Meteor Mon, 36(7): 43-49(in Chinese).
- 杨梦琪, 陈国民, 张喜平, 等, 2023. 2021 年西北太平洋和南海台风预报精度评定[J]. 气象, 49(8): 1005-1019. Yang M Q, Chen G M, Zhang X P, et al, 2023. Verification on typhoons forecasts over the Western North Pacific and the South China Sea in 2021[J]. Meteor Mon, 49(8): 1005-1019(in Chinese).
- 杨绚, 张立生, 王铸, 2022. 基于机器学习算法的县域台风灾害经济损失风险评估[J]. 热带气象学报, 38(5): 651-661. Yang X, Zhang L S, Wang Z, 2022. Risk assessment for typhoon economic losses in county-based units using machine learning[J]. J Trop Meteor, 38(5): 651-661(in Chinese).
- 张钧民, 阮惠华, 许剑辉, 等, 2021. 基于 XGBoost 的多源降水数据融合方法研究[J]. 热带地理, 41(4): 845-856. Zhang J M, Ruan H H, Xu J H, et al, 2021. An XGBoost-merging method for high-resolution daily precipitation estimation for a regional rainstorm event[J]. Trop Geogr, 41(4): 845-856(in Chinese).
- Akiba T, Sano S, Yanase T, et al, 2019. Optuna: a next-generation hyperparameter optimization framework[C] // Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. Anchorage: Association for Computing Machinery.
- Chen T, He T, Benesty M, et al, 2015. Xgboost: extreme gradient boosting[EB/OL]. (2023-10-30). R package version 0. 4-2, 1(4): 1-4. <https://cran.r-project.org/web/packages/xgboost/index.html>.
- Emanuel K A, 1986. An air-sea interaction theory for tropical cyclones. Part I: steady-state maintenance[J]. J Atmos Sci, 43(6): 585-605.
- Gray W M, 1968. Global view of the origin of tropical disturbances and storms[J]. Mon Wea Rev, 96(10): 669-700.
- Guo X, Tan Z M, 2017. Tropical cyclone fullness: a new concept for interpreting storm intensity[J]. Geophys Res Lett, 44(9): 4324-4331.
- Holland G J, 1997. The maximum potential intensity of tropical cyclones[J]. J Atmos Sci, 54(21): 2519-2541.
- Kaplan J, DeMaria M, 2003. Large-scale characteristics of rapidly intensifying tropical cyclones in the North Atlantic Basin[J]. Wea Forecasting, 18(6): 1093-1108.
- Kaplan J, DeMaria M, Knaff J A, 2010. A revised tropical cyclone rapid intensification index for the Atlantic and eastern North Pacific Basins[J]. Wea Forecasting, 25(1): 220-241.
- Kaplan J, Rozoff C M, DeMaria M, et al, 2015. Evaluating environmental impacts on tropical cyclone rapid intensification predictability

- utilizing statistical models[J]. *Wea Forecasting*, 30(5):1374-1396.
- Knaff J A, Sampson C R, Demaria M, 2005. An operational statistical typhoon intensity prediction scheme for the western North Pacific[J]. *Wea Forecasting*, 20(4):688-699.
- Knapp K R, Kruk M C, Levinson D H, et al, 2010. The International Best Track Archive for Climate Stewardship (IBTrACS): unifying tropical cyclone data[J]. *Bull Amer Meteor Soc*, 91(3):363-376.
- 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*. San Francisco: Morgan Kaufmann Publishers Inc:1137-1145.
- Levitus S, 1983. Climatological atlas of the world ocean[J]. *Eos*, 64(49):962-963.
- Lundberg S M, Erion G, Chen H, et al, 2020. From local explanations to global understanding with explainable AI for trees[J]. *Nat Mach Intell*, 2(1):56-67.
- Mercer A, Grimes A, 2017. Atlantic tropical cyclone rapid intensification probabilistic forecasts from an ensemble of machine learning methods[J]. *Proc Comput Sci*, 114:333-340.
- Merrill R T, 1988. Environmental influences on hurricane intensification[J]. *J Atmos Sci*, 45(11):1678-1687.
- National Centers for Environmental Prediction, National Weather Service, NOAA, et al, 2015. NCEP GFS 0.25 Degree Global Forecast Grids Historical Archive, Research Data Archive, National Center for Atmospheric Research, Computational and Information Systems Laboratory[EB/OL]. [2023-02-10]. <https://doi.org/10.5065/D65D8PWK>.
- Nguyen L T, Molinari J, Thomas D, 2014. Evaluation of tropical cyclone center identification methods in numerical models[J]. *Mon Wea Rev*, 142(11):4326-4339.
- Rozoff C M, Velden C S, Kaplan J, et al, 2015. Improvements in the probabilistic prediction of tropical cyclone rapid intensification with passive microwave observations[J]. *Wea Forecasting*, 30(4):1016-1038.
- Shi D L, Chen G H, 2021. The implication of outflow structure for the rapid intensification of tropical cyclones under vertical wind shear[J]. *Mon Wea Rev*, 149(12):4107-4127.
- Zhou G B, Xu J, Qian Q F, et al, 2022. Discriminating technique of typhoon rapid intensification trend based on artificial intelligence[J]. *Atmosphere*, 13(3):448.

(本文责编:何晓欢)