

董林,吕心艳,高拴柱,等,2024.一种基于双重检验逐步回归的台风生成预报方法[J].气象,50(7):835-846. Dong L, Lyu X Y, Gao S Z, et al, 2024. Tropical cyclone genesis forecasting method based on double test stepwise regression[J]. Meteor Mon, 50(7):835-846(in Chinese).

一种基于双重检验逐步回归的台风生成预报方法^{*}

董 林 吕心艳 高拴柱 王 晴

国家气象中心,北京 100081

提 要: 基于 CMA 台风最佳路径和 ECMWF 细网格确定性数值模式资料,计算西北太平洋和南海台风生成期间 16 种环境因子的值,选出 7 个显著性因子,采用双重检验逐步回归方法建立台风生成预报方程。独立样本检验结果表明,利用回归方法预报的风速值作为台风生成判据,较直接使用模式预报的风速和气压作为台风生成判据的命中率更高。双重检验逐步回归的台风生成预报方法在 2022 年的台风预报业务中进行了试用,对 2203 号台风暹芭的生成预报检验结果显示,回归方法表现较好,可为预报员判别台风是否生成及生成时间提供参考。

关键词: 台风生成,逐步回归,双重检验

中图分类号: P456

文献标志码: A

DOI: 10.7519/j.issn.1000-0526.2023.082101

Tropical Cyclone Genesis Forecasting Method Based on Double Test Stepwise Regression

DONG Lin LYU Xinyan GAO Shuanzhu WANG Qian

National Meteorological Centre, Beijing 100081

Abstract: Based on CMA best-track and ECMWF fine grid data, 16 environmental variables during TC genesis are calculated and 7 prediction factors with significance are selected. The stepwise regression algorithm based on double test was used in establishing the TC genesis forecasting equation. The independent sample test results show that the hit rate of the wind speed predicted by the regression method is higher than that of the wind speed and pressure predicted by the numerical model. In 2022, the stepwise regression algorithm based on double test was trialed in the TC genesis forecasting operation. The test results of generating forecast for Typhoon Chaba showed that the regression method has a good performance, and it can provide a reference for forecasters to judge whether and when the typhoon is generated.

Key words: tropical cyclone genesis, stepwise regression, double test

引 言

台风(本文指热带气旋,以下简称 TC)是影响我国最严重的自然灾害之一。为了减小损失,业务

和科研人员对台风路径、强度、结构、风雨影响等方面开展了大量研究,取得了丰富成果(Chen, 2012; Wang, 2012; Emanuel, 2018)。这些研究工作一方面促进了人们对台风的理解和认识,另一方面也提高了台风路径预报的准确率(Elsberry, 2014)。但是

^{*} 国家重点研发计划(2023YFC3107902)、中国气象局创新发展专项(CXFZ2023J015)和国家自然科学基金项目(42175016)共同资助

2023 年 1 月 17 日收稿; 2024 年 3 月 18 日收修定稿

第一作者:董林,主要从事台风预报和相关研究. E-mail: donglin@cma.gov.cn

由于科技水平的制约和洋面上观测资料的缺乏,限制了人们对台风发生发展机理和内部精细结构的深入认识,特别是在台风生成和快速加强两个方面表现尤为突出(Emanuel, 2005)。在台风生成机理研究方面,曾有学者先后提出正压不稳定、第二类条件不稳定(CISK)、风驱动的海气热量交换(WISHE)、多尺度相互作用等多种机制(吕美仲和彭永清, 1990; 张文龙等, 2009; Ooyama, 1969; Charney and Eliassen, 1964; Emanuel, 1986; 1989; Rotunno and Emanuel, 1987),以及涡度下传(top-down)、涡度上传(bottom-up)、“育儿袋”(pouch)等多种假说(Bister and Emanuel, 1997; Zhang and Bao, 1996; Wang et al, 2010; Dunkerton et al, 2009),但是每种机制或假说均不能单独对台风生成过程做出完美的解释。因此,迄今为止仍没有一个被一致认同和广泛接受的台风生成理论。在台风生成预报方法方面,我国学者的研究大多集中在气候特征或大尺度系统,如厄尔尼诺、南极涛动、季风、越赤道气流、热带对流层上部槽(TUTT)、热带大气季节内振荡(MJO)对台风生成时空分布和活动特征的影响等方面(黄勇等, 2008; 王慧等, 2006; 王会军和范可, 2006; 李崇银, 1986; 李崇银等, 2012);基于气候模式研发的预测方法或指数只能提供月、季或年际的台风生成频数和时空分布预报(赵军平等, 2012; 卢秋珍等, 2007; 范可, 2007; 王会军等, 2006; Fan, 2007)。台风群发、台风生成个例分析、扰动加强期间的物理量和卫星云图特征分析等方面的少量研究(吕心艳和端义宏, 2011; 张文龙等, 2008; 叶日新和吴立广, 2016),可以在业务中为实况分析提供线索,但无法作为有效的客观方法为台风生成预报业务提供足够的科技支撑。受到理论发展和方法研究的制约,台风生成预报水平的提高非常缓慢。

尽管如此,随着经济的发展,决策部门和公众对于台风生成预报产品的需求却在不断增加。这样的服务需求促使全球各海域预报中心陆续开展台风生成预报业务,发布台风生成预报产品(如热带天气展望),预报热带扰动在未来不同时效的发展概率(Chand and Walsh, 2011)。国家气象中心负责西北太平洋和南海的台风预报,并对该海域内达到热带风暴(TS)级别或在 48 h 警戒区内的热带低压(TD)进行监测和预报,预报产品通过全球通信系统(GTS)向全球发布(钱传海等, 2012)。美国联合台风警报中心(JTWC)和日本气象厅(JMA)同时也对这个海域内的 TC 发布预报预警产品。JTWC 利用

台风生成警报发布热带扰动未来 24 h 的发展潜势; JMA 在不同的预报产品中,以不同形式表现扰动的发展程度和未来趋势。目前,各预报中心发布的台风生成产品均是预报员通过分析各种资料,对热带扰动的发展趋势做出的综合评估和判断,预报产品的准确率远不及 TC 的路径和强度预报(Kucas and Darlow, 2012)。

综上所述,在台风生成预报方面,机理和方法的研究仍不能满足业务发展和服务方面的需要。因此,自 2015 年起国家气象中心台风预报团队着手研发台风生成主、客观预报方法,并进行台风生成预报业务试验。本文主要介绍一种基于双重检验逐步回归台风生成预报方法的原理和检验结果,供预报和科研人员参考。

1 资料和方法

1.1 资料

本文资料包括 2015—2021 年中国气象局台风最佳路径数据集(CMA-BestTrack; Ming et al, 2014)、2022 年台风业务定位资料和 2015—2022 年欧洲中期天气预报中心(ECMWF)细网格确定性数值模式资料,高空要素的水平分辨率为 $0.25^{\circ} \times 0.25^{\circ}$,地面要素的水平分辨率为 $0.125^{\circ} \times 0.125^{\circ}$,垂直分辨率为 19 层,资料范围为 $0^{\circ} \sim 60^{\circ}\text{N}$ 、 $50^{\circ} \sim 150^{\circ}\text{E}$,其中 2015—2019 年资料为训练数据,2020 年和 2021 年为检验数据,2022 年为业务应用数据。

1.2 TC 生成定义

目前,针对 TC 生成的定义没有统一的规范和标准。一般认为 TC 生成是系统从一个以地表通量为驱动能量的扰动,发展到一个轴对称的、具有暖心的并在地面有低压中心的过程(Emanuel, 2003)。Zehr(1992)从内外强迫、持续时间、最大风速、中心最低气压等 4 个方面对 TC 的生命史进行了定义和划分,其中将生成期又分为热带气旋性扰动和热带低压 2 个阶段。张文龙和崔晓鹏(2013)将这 2 个阶段,分别称为生成前期和生成后期。Zehr(1992)对于 TC 生成的 2 个阶段划分理论虽然晚于 Frank(1988),但是其给出了明确的指标,在业务上具有很强的可操作性,因此成为目前 TC 生成研究中引用率最高、业务上使用最多的 TC 生成阶段划分理论。

本文采用 Zehr(1992)的划分标准,结合张文龙

和崔晓鹏(2013)的命名方式,将 TC 生成后期定义为 TC 生成阶段,即热带低压发展为热带风暴的阶段。结合我国对 TC 等级的划分标准,也就是 TC 中心风速由 $10.8 \text{ m} \cdot \text{s}^{-1}$ 加强为 $17.2 \text{ m} \cdot \text{s}^{-1}$ 的阶段。由于 CMA-BestTrack 资料中只有整数位的风速值,因此本文中的 TC 生成样本指最佳路径资料中 TC 强度首次达到 $18 \text{ m} \cdot \text{s}^{-1}$ 的时刻和位置,未生成样本指资料中热带低压阶段的样本(Ming et al, 2014)。

1.3 涡旋追踪算法

对于 TC 生成前 CMA-BestTrack 没有数据的样本,需要在模式资料中识别和追踪 TC 中心位置。使用 9 个参数的涡旋追踪算法计算 CMA-BestTrack TC 数据样本生成时刻前 48 h 的 TC 中心位置,包括 6 个主参数和 3 个副参数(Marchok, 2002)。6 个主参数分别为 850 hPa、700 hPa 和地面的相对涡度,850 hPa 和 700 hPa 的位势高度及海平面气压;3 个副参数分别为 850 hPa、700 hPa 和地面风速。计算步骤为:(1)确定初猜位置和搜索范围:将 CMA-BestTrack 中样本的位置作为初猜位置,搜索范围为以初猜位置为圆心、半径 300 km 的圆形区域。(2)确定二猜位置:在搜索范围内计算 6 个主参数极值的位置,若其位置与初猜位置距离小于 275 km,则为有效参数,否则为无效参数;查找所有有效参数极值位置,而后与初猜位置一起计算得到的平均位置即为二猜位置。(3)确定最终位置:以二猜位置为圆心,查找 300 km 半径范围内 3 个副参数极值的位置;若副参数极值的位置与二猜位置距离小于 275 km,则为有效参数,否则为无效参数;计算有效副参数极值位置与二猜位置的平均位置作为涡旋中心的最终位置。

此外,为了更合理地确定参数极值所在位置,在计算过程中进行了 2 次更细致的数据处理。(1)每次确定中心位置后,将参数在搜索半径范围内进行重新插值,格点间距不变,即原始资料的分辨率如果为 10 km,那么以中心位置形成的新格点场的分辨率仍为 10 km。(2)参数的位置选择新格点上的极值点位置。但是,为了避免由于数据之间的差别过于微小而导致无法确定极值点,所以新格点上的参数值没有直接使用数值模式的输出值,而是使用 Barnes 算法(Barnes, 1964)对每个格点周围 150 km 半径范围内的格点值做距离倒数加权平均,将其作为该格点上的参数值。通过以上算法可以实现模式分析场和不同预报时效形势场中涡旋位置的判别和追踪,从而形成涡旋路径。

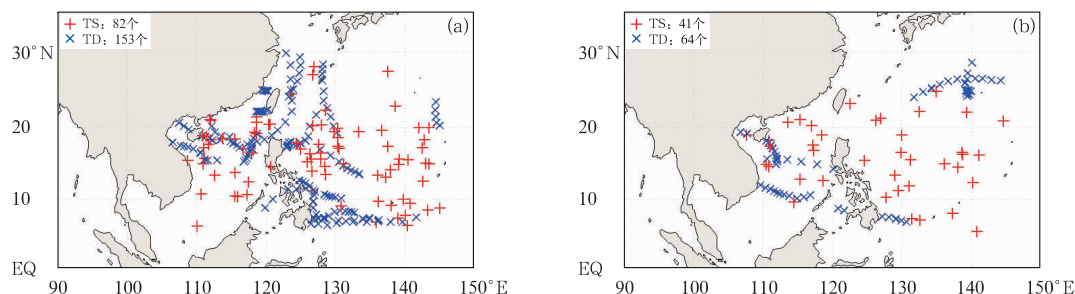
1.4 TC 样本筛选

由于 ECMWF 细网格确定性模式资料的东边界为 150°E ,本文计算 TC 环境场物理量时需要在一定范围内进行平均,因此在筛选样本时设定了 3 个条件,即 TC 中心位置位于 30°N 以南、 145°E 以西且不在陆地上。符合以上条件的样本数共计 340 个(图 1),其中 2015—2019 年的训练样本共计 235 个,包括生成样本 82 个,未生成样本 153 个(图 1a),2020—2021 年的检验样本共计 105 个,包括生成样本 41 个,未生成样本 64 个(图 1b)。

2 台风生成客观预报方法

2.1 TC 生成环境因子筛选

Gray(1968)首先提出了有利于 TC 生成的 6 个环境因子,即暖海温、大气处于条件不稳定、深厚的



注:“+”代表生成样本,“×”代表未生成样本。

图 1 (a)2015—2019 年训练和(b)2020—2021 年检验 TC 样本分布

Fig. 1 TC sample distribution of (a) training samples from 2015 to 2019 and (b) testing samples from 2020 to 2021

中层湿度、存在一定的科里奥利力、较大的低层涡度和弱的环境风垂直切变。陈联寿和丁一汇(1979)、Emanuel(1989)、Gall et al(2010)等先后对这些因子进行了更加细致的分析。McBride and Zehr(1981)、Cheung(2004)、Fu et al(2012)筛选出贡献较大的因子,建立了月、季、年等较长时效的 TC 生成预报方法。本文基于以上研究结果,结合 Demaria and Kaplan(1994; 1999)、Zehr(1992)、高拴柱等(2012; 2021)对 TC 生成或加强期间环境因子的统计分析,分别从海洋和大气两个方面,筛选出表征低层辐合、高层辐散、中层水汽、大气层结条件、高层流

出、低层流入、整层切变以及低层散度和涡度的前后变化等 16 个环境因子。同时,为了便于后面的检验和对比,计算每个样本的 3 个检验因子,即海平面气压最小值、地面风速和 850 hPa 风速最大值。表 1 给出了 19 个因子。其中第 1~3 号为检验因子,为涡旋中心 200 km 半径内的格点值;第 4~19 号为 TC 生成环境因子,其中海温为涡旋中心位置的格点值,其他因子的计算区域半径均为 500 km。第 7 号和第 10 号的 2 个增量为后一个时刻与当前时刻的差值,2 个时刻的时间间隔为 12 h。

图 2 为 2015—2019 年 235 个训练样本的物理

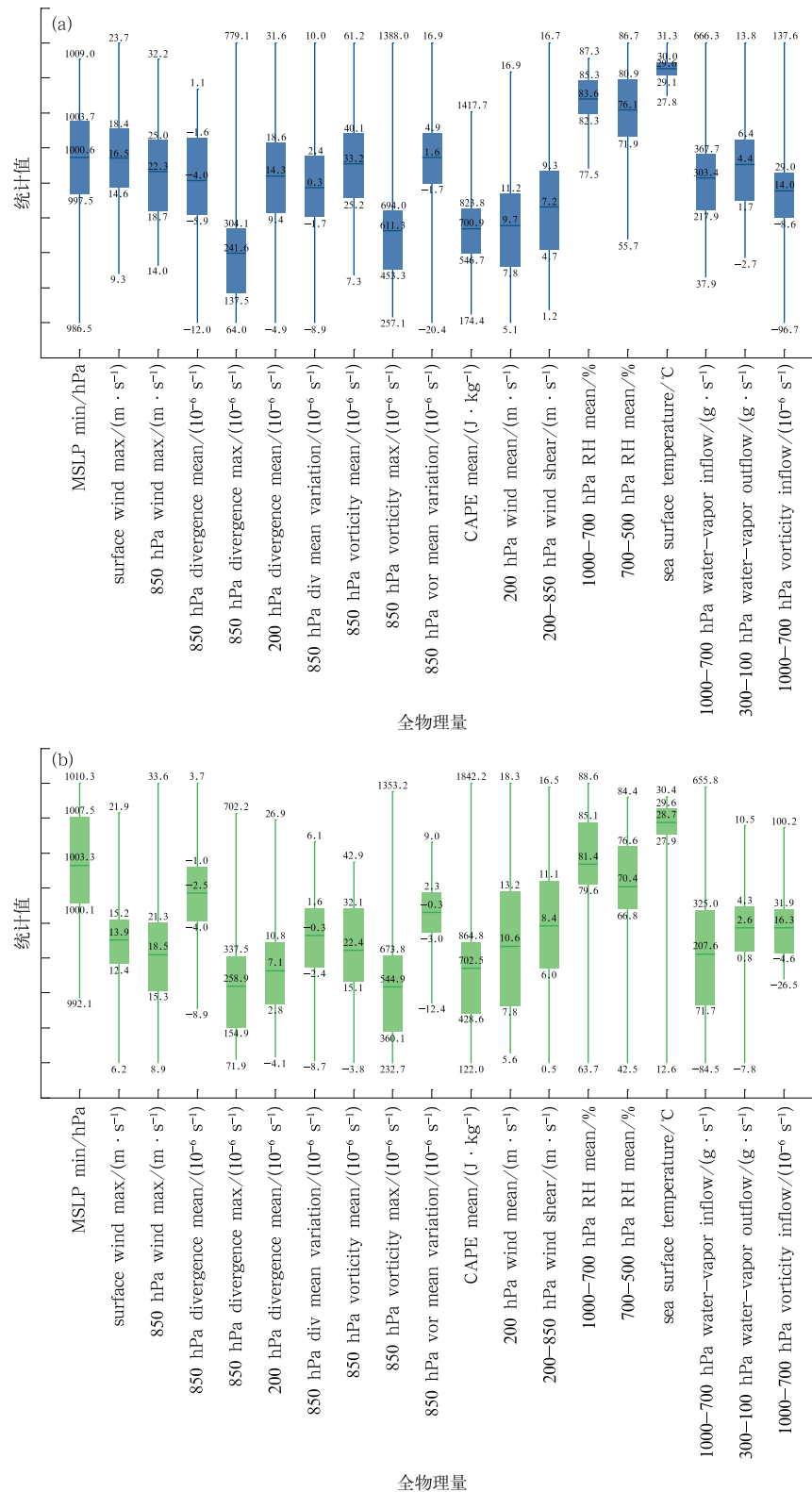
表 1 TC 生成期间的全物理量

Table 1 All physical quantities during TC genesis

序号	中文名称	英文名称	单位
1	海平面气压最小值	MSLP min	hPa
2	地面风速最大值	surface wind max	$\text{m} \cdot \text{s}^{-1}$
3	850 hPa 风速最大值	850 hPa wind max	$\text{m} \cdot \text{s}^{-1}$
4	海面温度	sea surface temperature	$^{\circ}\text{C}$
5	850 hPa 散度最大值	850 hPa divergence max	10^{-6} s^{-1}
6	850 hPa 散度平均值	850 hPa divergence mean	10^{-6} s^{-1}
7	850 hPa 散度平均值增量	850 hPa div mean variation	10^{-6} s^{-1}
8	850 hPa 涡度最大值	850 hPa vorticity max	10^{-6} s^{-1}
9	850 hPa 涡度平均值	850 hPa vorticity mean	10^{-6} s^{-1}
10	850 hPa 涡度平均值增量	850 hPa vor mean variation	10^{-6} s^{-1}
11	200 hPa 平均风速	200 hPa wind mean	$\text{m} \cdot \text{s}^{-1}$
12	200 hPa 散度平均值	200 hPa divergence mean	10^{-6} s^{-1}
13	200~850 hPa 水平风垂直切变	200—850 hPa wind shear	$\text{m} \cdot \text{s}^{-1}$
14	1000~700 hPa 平均相对湿度	1000—700 hPa RH mean	%
15	1000~700 hPa 水汽流入量	1000—700 hPa water-vapor inflow	$\text{g} \cdot \text{s}^{-1}$
16	1000~700 hPa 涡度流入量	1000—700 hPa vorticity inflow	10^{-6} s^{-1}
17	700~500 hPa 平均相对湿度	700—500 hPa RH mean	%
18	300~100 hPa 水汽流出量	300—100 hPa water-vapor outflow	$\text{g} \cdot \text{s}^{-1}$
19	对流有效位能平均值	CAPE mean	$\text{J} \cdot \text{kg}^{-1}$

量分布。从检验因子分布可以看出,生成样本的海平面气压和地面风速的均值分别为 1000.6 hPa 和 $16.5 \text{ m} \cdot \text{s}^{-1}$,相比业务中使用的 TC 生成时刻的阈值(998 hPa 和 $18 \text{ m} \cdot \text{s}^{-1}$)明显偏弱,说明 ECMWF 细网格确定性模式资料在业务认定 TC 生成时刻,其分析场中的海平面气压和地面风速较实况偏弱。这可能会导致模式对 TC 的强度预报偏弱,从而影响 TC 生成预报的准确率或者导致预报的生成时间偏晚。对于 TC 生成时刻的 16 个环境因子,从数值上看,有些因子的差异显著,如 200 hPa 散度、海面温度、850 hPa 平均涡度和整层切变,生成样本的均值分别为 $14.3 \times 10^{-6} \text{ s}^{-1}$ 、 29.6°C 、 $33.2 \times 10^{-6} \text{ s}^{-1}$

和 $7.2 \text{ m} \cdot \text{s}^{-1}$,而未生成样本的均值分别为 $7.1 \times 10^{-6} \text{ s}^{-1}$ 、 28.7°C 、 $22.4 \times 10^{-6} \text{ s}^{-1}$ 和 $8.4 \text{ m} \cdot \text{s}^{-1}$;也在一定程度上表明高层辐散越强、低层环境场的涡度越高、海面越暖、环境风垂直切变越小越有利于 TC 生成。同时,也有一些物理量,生成和未生成样本的数值差异并不明显,但是生成样本的整体分布聚集程度更高,如 1000~700 hPa 的相对湿度和湿度的变化、850 hPa 最大涡度等,表明这些物理量可能不是 TC 生成的关键因素,但是如果湿度大或扰动中有大涡度区等,这些条件可以在生成期间起到一定的积极作用。这些只是定性的分析,如果要定量分析每个物理量的贡献,还需要用逐步回归的方法对



注:箱线图中触须的上、下两端分别为样本中的最大值和最小值,箱体分别为样本的第 25% 和第 75% 分位值,中间为均值;下同。

图 2 2015—2019 年(a)生成 TC 样本,(b)未生成 TC 样本物理量统计分布
Fig. 2 Statistical distribution of physical quantity of (a) genesis samples and (b) non-genesis samples from 2015 to 2019

因子进行筛选。

2.2 建立双重检验的逐步回归方法

基于黄嘉佑(2000)和朱玉祥等(2017)的研究,建立双重检验的逐步回归模型。将表征 TC 在第 i 时刻强度的涡旋中心附近地面最大风速作为预报量 y_i ,使用表 1 中第 4~19 共 16 个量作为预报因子 x_i 。基于 TC 生成时刻 16 个预报因子的数值,采用双重检验的逐步回归方法选出具有显著性的预报因

子。该方法的基本思想是依次引入一个因子,引入因子的条件是该因子的方差贡献是显著的。同时,每引入一个新因子后,要对已被选中的所有因子逐个检验,将方差贡献变为不显著的因子剔除,即引入和剔除因子时均进行显著性水平检验。双重检验的逐步回归方法适合此类自变量比较多的情况,可以将不完全独立的因子剔除。

表 2 给出了使用 235 个台风样本,运用双重检验逐步回归方法得到的回归系数和方差贡献。

表 2 双重检验逐步回归方法的预报因子、方差贡献和 F 统计量
Table 2 Prediction factor, variance contribution and F statistic of stepwise regression method based on double test

变量	被选中预报因子	回归系数	方差贡献	F 统计量数值
x_1	200 hPa 散度平均值	0.7148	0.0335	63.4915
x_2	850 hPa 涡度最大值	0.5567	0.0766	19.8609
x_3	海面温度	0.3369	0.0482	14.1262
x_4	300~100 hPa 水汽流出量	0.1351	0.0352	9.1905
x_5	200~850 hPa 垂直风切变	-0.1045	0.0237	6.5451
x_6	1000~700 hPa 涡度流入量	-0.0095	0.0151	5.9345
x_7	850 hPa 涡度平均值	0.0129	0.0142	5.1956

根据表 2 中的结果,得到的回归方程为

$$\hat{y} = 2.9181 + 0.7148x_1 + 0.5567x_2 + 0.3369x_3 + 0.1351x_4 - 0.1045x_5 - 0.0095x_6 + 0.0129x_7 \quad (1)$$

式中: $\hat{y}$ 的均方差无偏估计量为 $1.8328 \text{ m} \cdot \text{s}^{-1}$,预报中将使用这个值检验 0.05 的显著性水平。

3 预报结果检验

3.1 检验规则

为评估双重检验逐步回归方法的预报效果,本文参考 Cheung and Elsberry(2002)、冯文等(2016)的检验规则,设计了 4 个预报结果:分别为击中、早报、晚报和错报。击中即表明预报的 TC 生成时间在 $\pm 24 \text{ h}$ 以内且位置在 5 个经(纬)度以内。但是早报和晚报较冯文等(2016)的规则更加严苛,将时间上限由 $\pm 72 \text{ h}$ 缩短为 $\pm 48 \text{ h}$ 。表 3 对检验规则给出了具体说明。

3.2 检验结果

利用 2020—2021 年的 TC 样本对回归方法的预报效果进行独立样本检验(样本分布见图 1b)。首先计算每个 TC 生成时刻前后 48 h(未生成样本指该时刻热带低压前后 48 h)的 3 个检验因子的预报值及回归方法的预报风速值。由于 ECMWF 模式起报间隔为 12 h,因此每个样本有 9 组不同起报时间的预报结果,预报时效间隔为 3 h,预报总样本数为 12 262 个。为进行预报效果对比,将训练样本中生成样本计算得到的 3 个检验因子,即海平面气压最小值、地面风速最大值和 850 hPa 风速最大值,这 3 个因子的均值、第 25%分位值、第 75%分位值共 9 个统计量作为对比检验的 TC 生成阈值。

由表 4 的 10 组指标检验结果可见,使用回归风速值进行 TC 生成预报判断的击中率为 78.1%,在所有检验指标中击中率最高;其次是 850 hPa 风速最大值的第 25%分位值($18.7 \text{ m} \cdot \text{s}^{-1}$)、地面风速最大值的第 25%分位值($14.6 \text{ m} \cdot \text{s}^{-1}$)、850 hPa 风

表 3 TC 生成预报的检验规则
Table 3 Test rule of TC genesis forecasting

预报结果	检验规则
击中	满足以下 2 个条件之一即为击中:(1)达到 TC 生成阈值且模式预报的 TC 生成位置与最佳路径给出的 TC 生成位置距离之差小于等于 5 个经(纬)度,且生成时间之差在 24 h 内(包括 24 h);(2)最佳路径指出 TC 未生成,预报指标给出的预报结果也是未生成
早报	模式预报的 TC 生成位置与最佳路径给出的 TC 生成位置距离之差小于等于 5 个经(纬)度,但是模式预报的生成时间较最佳路径生成时间早 24 h 以上,且在 48 h 以下(包括 48 h)
晚报	模式预报的 TC 生成位置与最佳路径给出的 TC 生成位置距离之差小于等于 5 个经(纬)度,但是模式预报的生成时间较最佳路径生成时间晚 24 h 以上,且在 48 h 以下(包括 48 h)
错报	满足以下 2 个条件之一即为错报:(1)模式预报的 TC 生成位置与最佳路径生成位置距离之差大于 5 个经(纬)度,或者生成时间早于 48 h(或者晚于 48 h)以上;(2)在最佳路径生成位置 5 个经(纬)度范围内,或在最佳路径生成时间前后 48 h 内,模式没有预报 TC 生成

表 4 2020—2021 年台风生成的预报效果检验(单位:%)
Table 4 Verification of TC genesis forecasting from 2020 to 2021 (unit: %)

检验要素(阈值)	击中	早报	晚报	错报
回归风速值($18\text{ m}\cdot\text{s}^{-1}$)	78.1	1.4	6.1	14.4
海平面气压最小值的第 25%分位值(997.5 hPa)	68.2	0.2	5.3	26.3
海平面气压最小值的平均值(1000.6 hPa)	64.7	0.3	6.5	28.5
海平面气压最小值的第 75%分位值(1003.7 hPa)	47.3	1.1	6.1	44.5
地面风速最大值的第 25%分位值($14.6\text{ m}\cdot\text{s}^{-1}$)	70.7	0	5.5	23.8
地面风速最大值的平均值($16.5\text{ m}\cdot\text{s}^{-1}$)	66.6	0.2	6.3	26.9
地面风速最大值的第 75%分位值($18.4\text{ m}\cdot\text{s}^{-1}$)	55.4	0.5	6.8	37.3
850 hPa 风速最大值的第 25%分位值($18.7\text{ m}\cdot\text{s}^{-1}$)	71.5	0	4.7	23.8
850 hPa 风速最大值的平均值($22.3\text{ m}\cdot\text{s}^{-1}$)	70.7	0	5.8	23.4
850 hPa 风速最大值的第 75%分位值($25.0\text{ m}\cdot\text{s}^{-1}$)	61.6	0	6.6	31.5

速最大值的平均值($22.3\text{ m}\cdot\text{s}^{-1}$)和海平面气压最小值的第 25%分位值(997.5 hPa),击中率分别为 71.5%、70.7%、70.7%和 68.2%,略逊于回归方法的预报结果。因此,在业务中可以把回归风速值达到 $18\text{ m}\cdot\text{s}^{-1}$ 作为判断 TC 生成的主要参考指标,把另外几个参数作为辅助参考指标。同时,对于没有击中的预报,晚报的比例明显大于早报。可能的原因有两个,一是模式对 TC 生成期间的风速、气压等要素的预报偏弱,即使利用统计回归方法对模式预报的 TC 强度进行了订正,依然无法完全解决这个问题;二是一部分 TC 的可预报性比较差,模式预报时效的提前量小,导致 TC 生成前 24~48 h 的各种物理量较弱。

4 业务试用

4.1 概 况

双重检验逐步回归的台风生成预报方法(以下将此方法所预报的涡旋强度简称为 TCGI 方法)在 2022 年进行了业务试用,期间 TCGI 方法针对 3 个台风(2203 号台风暹芭、2209 号台风马鞍、2212 号台风梅花)和 1 个南海热带低压启动了作业。从预报结果来看,“暹芭”“马鞍”和南海热带低压预报效果较好,“梅花”预报效果欠佳。下面以“暹芭”为例说明 TCGI 方法的运行和应用情况。

“暹芭”是 2022 年首个在我国登陆的台风。2022 年 6 月 28 日 00 时(世界时,下同)开始针对“暹芭”生成前的热带扰动启动运行 TCGI 方法,29 日 00 时该热带扰动被编为 2022 年第 3 号热带低压,30 日 00 时编号为 2203 号台风,7 月 2 日 07 时“暹芭”在广东省电白市沿海登陆($35 \text{ m} \cdot \text{s}^{-1}$, 965 hPa),4 日 00 时停止编号(图 3b)。

4.2 预报产品

业务试用产品主要分为 5 类,分别为台风路径产品(图 4a,通过 1.3 节的涡旋追踪算法形成的涡旋路径)、TCGI 方法预报结果(图 4b)、3 个检验因子的预报结果(图 4c~4e,模式预报的地面风速、850 hPa 风速和海平面气压)、16 种环境因子的预报结果(图略,表 1 中第 4~19 号环境因子预报和图 2 中生成样本统计量的叠加显示)和全物理量分布(图 4f,各预报时效 19 种物理量预报统计分布)。为了方便预报员对不同起报时次的预报结果进行对比分析,前 4 类产品分别以不同颜色显示该扰动自启动作业以来的每一次预报结果,形成不同起报时间的预报结果集成。最新起报时间以灰色竖线表示,最新的预报结果用粗黑虚线表示,零场值用粗黑实线表示。为了便于分析,图 4a~4e 展示了“暹芭”最后一次运行的预报图,本文在业务试用产品上,后期添加了紫色竖线表示台风生成时间。

4.3 预报效果分析

台风暹芭于 6 月 30 日 00 时编号,在其生成期

间,TCGI 方法从 6 月 28 日 00 时至 7 月 1 日 12 时共有 8 次起报,对前 7 次有检验价值的预报结果进行汇总(表 5)。为了对比预报效果,将表 4 中预报效果最好的前 4 个物理量,即 850 hPa 风速最大值的第 25%分位值、850 hPa 风速最大值的平均值、地面风速最大值的第 25%分位值和海平面气压最小值的第 25%分位值作为对比检验量。

由表 5 结果可见,表现最好的是 TCGI、850 hPa 风速最大值的第 25%分位值和海平面气压最小值的第 25%分位值,均为 7 次起报全部击中,其次是地面风速最大值的第 25%分位值,有 5 次击中和 2 次晚报,最后是 850 hPa 风速最大值的平均值,有 4 次击中、2 次晚报和 1 次错报,而这次错报是由于预报在 7 月 2 日 00 时之后才能达到阈值。在实际业务中,“暹芭”起编之前的 4 次起报是可供预报员参考的有效预报,这 4 次起报表现最好的依然是 TCGI、850 hPa 风速最大值的第 25%分位值和海平面气压最小值的第 25%分位值,4 次起报全部击中。另外,从图 4c 可以看出,如果按照业务中台风生成的阈值(地面最大风速超过 $18 \text{ m} \cdot \text{s}^{-1}$)对模式预报的地面最大风速进行检验,那么 7 次起报中模式只击中了后面 3 次起报;对于台风生成前的 4 次有效预报,模式预报的地面最大风速都没有超过阈值,无法为业务提供参考。因此,与之前业务中直接使用模式预报的地面最大风速判断台风生成相比,利用 TCGI 的预报结果,结合统计出来的 2 个检验阈值,即 850 hPa 风速最大值的第 25%分位值和海平面气压最小值的第 25%分位值,不但可以较好地判断

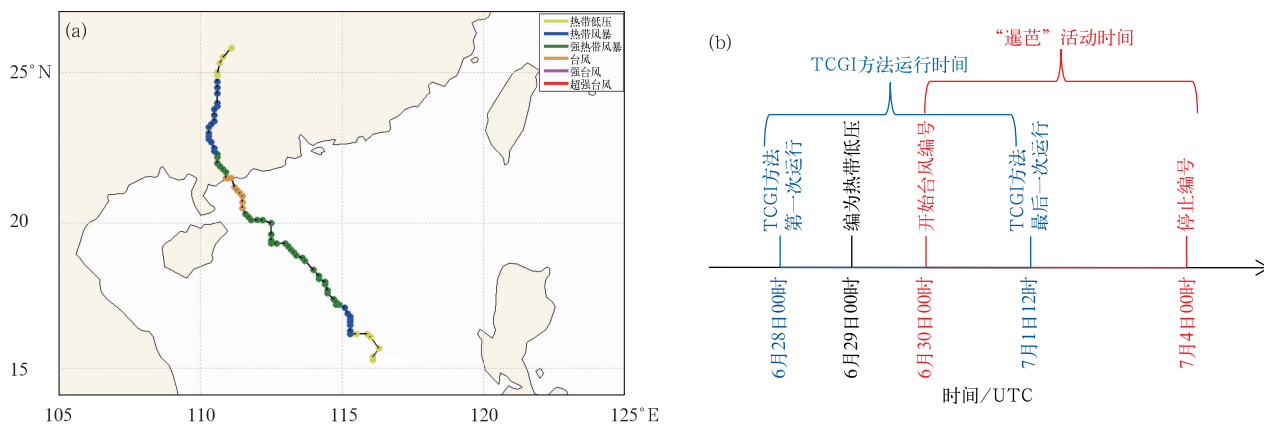
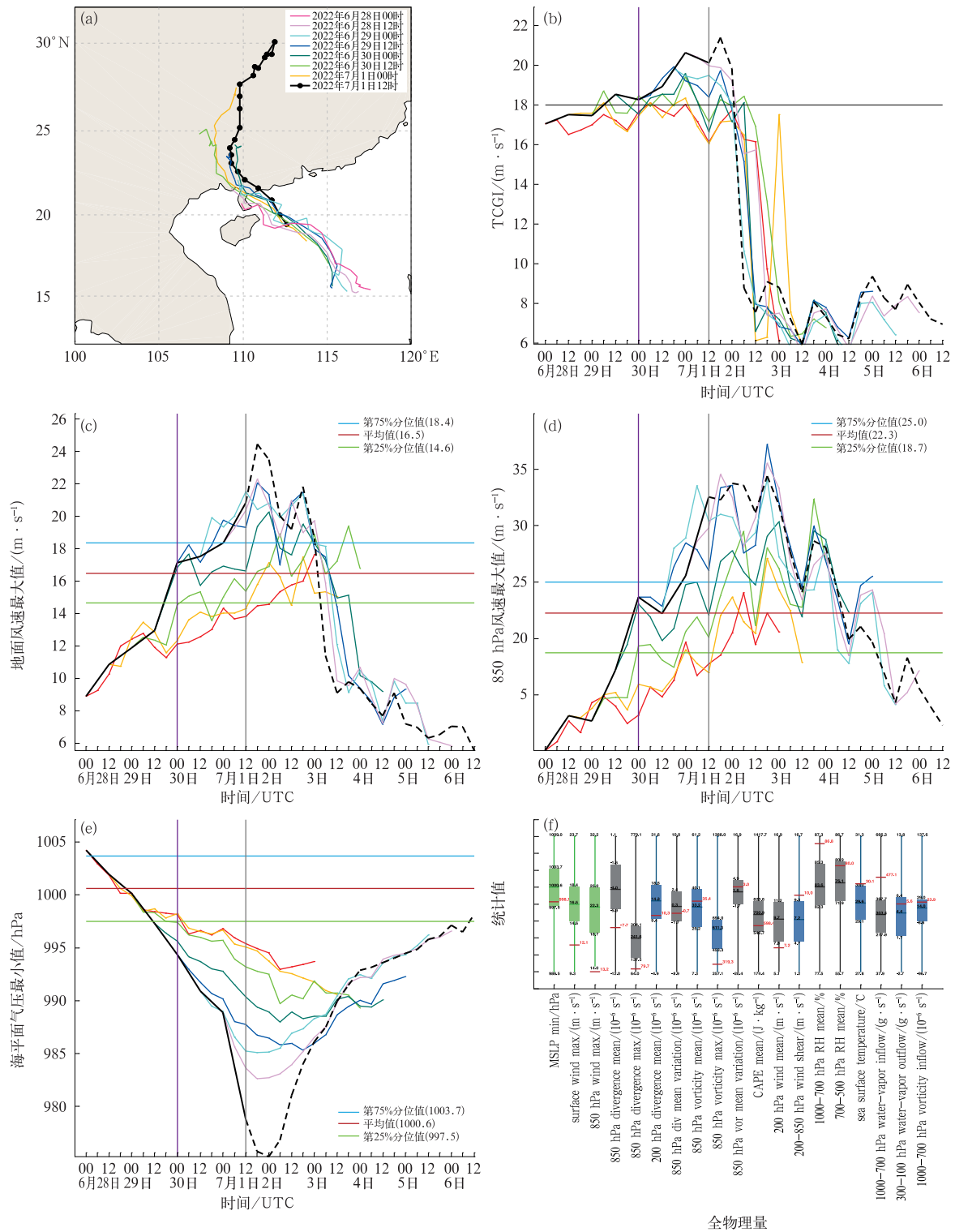


图 3 台风暹芭(a)路径和(b)TCGI 方法运行时间示意图

Fig. 3 Diagram of (a) track of Typhoon Chaba and (b) TCGI runtime



注:图 a~e 中,不同颜色曲线表示该扰动自启动作业以来的每一次预报结果,粗黑虚线表示最新预报结果,粗黑实线表示零场值,灰色竖线表示最新起报时间,紫色竖线表示台风生成时间;图 f 中红线为实时预报结果。

图 4 2022 年 7 月 1 日 12 时起报的 120 h 时效台风生成预报产品

(a) 涡旋路径, (b) TCGI, (c) 地面风速最大值, (d) 850 hPa 风速最大值,

(e) 海平面气压最小值, (f) 6 月 28 日 00 时起报的 48 h 时效(30 日 00 时)全物理量预报统计分布

Fig. 4 The 120 h TC genesis forecasting products initiated at 12:00 UTC 1 July

(a) track, (b) TCGI, (c) surface wind max, (d) 850 hPa wind max, (e) MSLP min,

(f) all physical quantities of 48 h TC genesis forecasting initiated at 00:00 UTC 28 June 2022

表 5 台风暹芭生成预报检验
Table 5 Verification of Typhoon Chaba genesis forecasting

对象	起报次数	物理量和阈值				
		TCGI (18 m · s ⁻¹)	850 hPa 风速最大值的第 25%分位值 (18.7 m · s ⁻¹)	850 hPa 风速最大值的平均值 (22.3 m · s ⁻¹)	地面风速最大值的第 25%分位值 (14.6 m · s ⁻¹)	海平面气压最小值的第 25%分位值 (997.5 hPa)
参与检验	7 次	7 次击中	7 次击中	4 次击中 2 次晚报 1 次错报	5 次击中 2 次晚报	7 次击中
有效预报	4 次	4 次击中	4 次击中	1 次击中 2 次晚报 1 次错报	2 次击中 2 次晚报	4 次击中

“暹芭”是否会生成,同时,预报员有 36~48 h 的提前量对台风生成做出及时和有效的判断。

5 结论和讨论

本文基于 2015—2019 年 CMA 台风最佳路径和 ECMWF 细网格确定性模式资料,计算西北太平洋和南海台风生成期间 16 种环境因子的值,采用双重检验逐步回归方法选出具有显著性的 7 个预报因子,建立台风生成预报方程。利用 2020 年和 2021 年的台风生成个例进行独立样本检验,将回归方法预报的风速值作为台风生成判据,计算击中、早报、晚报和错报的分布;并将预报结果与利用海平面气压最小值、地面风速最大值和 850 hPa 风速最大值这 3 个检验因子的不同分位值作为台风生成判据的预报结果进行对比检验,结果表明,回归方法的台风生成预报击中率最高,为 78.1%,其次是 850 hPa 风速最大值的第 25%分位值(18.7 m · s⁻¹)、地面风速最大值的第 25%分位值(14.6 m · s⁻¹)、850 hPa 风速最大值的平均值(22.3 m · s⁻¹)和海平面气压最小值的第 25%分位值(997.5 hPa),击中率分别为 71.5%、70.7%、70.7%和 68.2%。

2022 年,基于双重检验的逐步回归方法在业务上进行了试用,根据方法的预报结果和研发过程中的统计量,业务中可提供涡旋路径、回归方法预报量 TCGI、3 种对比检验物理量的预报值、16 种环境因子预报量以及每个预报时效的 19 种物理量全景图,共 5 类产品。为了方便预报员参考,业务产品设计成不同起报时间预报结果集成的展示方式,并叠加了每种物理量的统计分布。业务试用期间 TCGI 方法针对 2203 号台风暹芭、2209 号台风马鞍、2212 号台风梅花共 3 个台风和 1 个南海热带低压启动了作业。从预报结果来看,“暹芭”“马鞍”和南海热带低压这 3 个从季风槽中发展而来的系统预报效果较

好,但是从大洋中部槽切断的低涡(TUTT-cell)发展而来的“梅花”预报效果欠佳。

通过业务试用可以发现,在参考 TCGI 方法时三点需要关注。第一,TCGI 方法是基于台风生成时刻选取的显著性因子,其预报的风速只适用于台风生成期间,该值不会随着模式预报的台风强度的增加而增大,这说明控制台风生成和加强的环境因子可能是不同的。因此,可以参考 TCGI 和模式预报的风速、气压等来共同判断模式对某一个涡旋在早期发展阶段的强度预报趋势,但是在台风编号之后,除了涡旋路径产品之外,其他产品和各种环境因子的统计量无法继续为台风强度预报提供支撑。第二,TCGI 方法对于季风槽中生成的台风有非常好的预报效果,但是对于从东风波、TUTT-cell 或由冷空气激发等引起的台风生成,预报效果欠佳。这是由于西北太平洋和南海 70%以上的台风生成都是在季风槽中(Zehr,1992),训练样本的统计特征对这种类型的台风生成更具有代表性。第三,TCGI 方法的预报效果依赖于数值模式对台风生成期间各环境因子预报的好坏,如果模式对环境因子的预报效果偏差大,也会影响基于因子得到的预报结果。

未来的改进方向主要有两个,一是增加卫星云图上扰动的云型变化特征作为预报因子,重新建立统计关系,预测 TC 生成;二是改进统计方法,利用预报效果更优的人工智能或者深度学习方法代替双重检验的逐步回归方法(钱奇峰等,2021;吕心艳等,2022;周冠博等,2022)。

参考文献

- 陈联寿,丁一汇,1979. 西太平洋台风概论[M]. 北京:科学出版社:103-145. Chen L S, Ding Y H, 1979. An Introduction to the Western Pacific Typhoon[M]. Beijing: Science Press: 103-145(in Chinese).
- 范可,2007. 西北太平洋台风生成频次的新预测因子和新预测模型[J]. 中国科学 D 辑:地球科学,37(9):1260-1266. Fan K, 2007.

- New predictors and a new prediction model for the typhoon frequency over western North Pacific[J]. *Sci China Ser D: Earth Sci*, 37(9):1260-1266(in Chinese).
- 冯文, 黄伟健, 谭绮婷, 等, 2016. 一种基于多物理量阈值定义的热带气旋生成预报方法及在 ECMWF 全球模式中的应用验证[J]. *热带气象学报*, 32(6):908-917. Feng W, Wong W K, Tam Y T, et al, 2016. Tropical cyclone genesis forecasting based on thresholds of multiple physical parameters and verification of performance using ECMWF model[J]. *J Trop Meteor*, 32(6):908-917(in Chinese).
- 高拴柱, 吕心艳, 王海平, 等, 2012. 热带气旋莫兰蒂(1010)强度的观测研究和增强条件的诊断分析[J]. *气象*, 38(7):834-840. Gao S Z, Lü X Y, Wang H P, et al, 2012. An observational and diagnostic analysis on the intensity and intensity changes of Typhoon Meranti (1010)[J]. *Meteor Mon*, 38(7):834-840(in Chinese).
- 高拴柱, 张胜军, 吕心艳, 等, 2021. 南海台风生成前 48 h 环流特征及热力与动力条件[J]. *应用气象学报*, 32(3):272-288. Gao S Z, Zhang S J, Lü X Y, et al, 2021. Circulation characteristics and thermal and dynamic conditions 48 hours before typhoon formation in the South China Sea[J]. *J Appl Meteor Sci*, 32(3):272-288(in Chinese).
- 黄嘉佑, 2000. 气象统计分析与预报方法:第2版[M]. 北京:气象出版社:60-75. Huang J Y, 2000. *Statistic Analysis and Forecast Methods in Meteorology*[M]. 2nd ed. Beijing:China Meteorological Press:60-75(in Chinese).
- 黄勇, 李崇银, 王颖, 等, 2008. 近百年西北太平洋热带气旋频数变化特征与 ENSO 的关系[J]. *海洋预报*, 25(1):80-87. Huang Y, Li C Y, Wang Y, et al, 2008. Study on the variation characteristics of annual frequency for tropical cyclones in western North Pacific during the last hundred years and relationship with ENSO[J]. *Marine Forecasts*, 25(1):80-87(in Chinese).
- 李崇银, 1986. TUTT 对台风形成作用的动力学研究[J]. *气象科学研究院院刊*, 1(2):158-164. Li C Y, 1986. A dynamic study on the role of TUTT in the generation of typhoon[J]. *J Appl Meteor Sci*, 1(2):158-164(in Chinese).
- 李崇银, 潘静, 田华, 等, 2012. 西北太平洋台风活动与大气季节内振荡[J]. *气象*, 38(1):1-16. Li C Y, Pan J, Tian H, et al, 2012. Typhoon activities over the western North Pacific and atmospheric intraseasonal oscillation[J]. *Meteor Mon*, 38(1):1-16(in Chinese).
- 卢秋珍, 胡邦辉, 王学忠, 等, 2007. 西北太平洋台风活动的年代际变化与大尺度环流因子的关系[J]. *热带气象学报*, 23(6):629-635. Lu Q Z, Hu B H, Wang X Z, et al, 2007. Large scale circulation associated with interdecadal variations of typhoon activity over the western North Pacific[J]. *J Trop Meteor*, 23(6):629-635(in Chinese).
- 吕美仲, 彭永清, 1990. 动力气象学教程[M]. 北京:气象出版社:390-391. Lü M Z, Peng Y Q, 1990. *Dynamic Meteorology*[M]. Beijing:China Meteorological Press:390-391(in Chinese).
- 吕心艳, 端义宏, 2011. 季风槽内热带气旋生成的基本特征分析[J]. *气象学报*, 69(6):990-1000. Lü X Y, Duan Y H, 2011. Characteristics of the tropical cyclogenesis in the summer monsoon trough[J]. *Acta Meteor Sin*, 69(6):990-1000(in Chinese).
- 吕心艳, 钱奇峰, 王登科, 等, 2022. 基于深度图像目标检测的智能台风涡旋识别技术[J]. *热带气象学报*, 38(4):492-501. Lyu X Y, Qian Q F, Wang D K, et al, 2022. Intelligent technique of typhoon vortex detection based on object detection with deep learning of satellite image[J]. *J Trop Meteor*, 38(4):492-501(in Chinese).
- 钱传海, 端义宏, 麻素红, 等, 2012. 我国台风业务现状及其关键技术[J]. *气象科技进展*, 2(5):36-43. Qian C H, Duan Y H, Ma S H, et al, 2012. The current status and future development of China operational typhoon forecasting and its key technologies[J]. *Adv Meteor Sci Technol*, 2(5):36-43(in Chinese).
- 钱奇峰, 王川, 徐雅静, 等, 2021. 一种基于深度学习的台风强度估测技术[J]. *气象*, 47(5):601-608. Qian Q F, Wang C, Xu Y J, et al, 2021. A deep learning technique of typhoon intensity estimation[J]. *Meteor Mon*, 47(5):601-608(in Chinese).
- 王慧, 丁一汇, 何金海, 2006. 西北太平洋夏季风的变化对台风生成的影响[J]. *气象学报*, 64(3):345-356. Wang H, Ding Y H, He J H, 2006. Influence of western North Pacific summer monsoon changes on typhoon genesis[J]. *Acta Meteor Sin*, 64(3):345-356(in Chinese).
- 王会军, 范可, 2006. 西北太平洋台风生成频次与南极涛动的关系[J]. *科学通报*, 51(24):2910-2914. Wang H J, Fan K, 2006. Relationship between the Antarctic oscillation in the western North Pacific typhoon frequency[J]. *Chin Sci Bull*, 51(24):2910-2914(in Chinese).
- 王会军, 郎威海, 范可, 等, 2006. 关于 2006 年西太平洋台风活动频次的气候预测试验[J]. *气候与环境研究*, 11(2):133-137. Wang H J, Lang X M, Fan K, et al, 2006. Real-time climate prediction experiment for the typhoon frequency in the western North Pacific for 2006[J]. *Climatic Environ Res*, 11(2):133-137(in Chinese).
- 叶日新, 吴立广, 2016. 1998—2009 年间台风不活跃期间热带云团活动分析[J]. *气象科学*, 36(3):291-300. Ye R X, Wu L G, 2016. Analysis of tropical cloud clusters in typhoon inactive periods during 1998—2009[J]. *J Meteor Sci*, 36(3):291-300(in Chinese).
- 张文龙, 崔晓鹏, 2013. 热带气旋生成问题研究综述[J]. *热带气象学报*, 29(2):337-346. Zhang W L, Cui X P, 2013. Review of the studies on tropical cyclone genesis[J]. *J Trop Meteor*, 29(2):337-346(in Chinese).
- 张文龙, 崔晓鹏, 王昂生, 等, 2008. 2001 年台风“榴莲”生成前期对流“热塔”的数值模拟[J]. *热带气象学报*, 24(6):619-628. Zhang W L, Cui X P, Wang A S, et al, 2008. Numerical simulation of hot towers during pre-genesis stage of Typhoon Durian(2001)[J]. *J Trop Meteor*, 24(6):619-628(in Chinese).
- 张文龙, 张大林, 王昂生, 等, 2009. 台风榴莲(2001)在季风槽中生成的机制探讨[J]. *气象学报*, 67(5):811-827. Zhang W L, Zhang D L, Wang A S, et al, 2009. An investigation of the genesis of Typhoon Durian (2001) from a monsoon trough[J]. *Acta Meteor Sin*, 67(5):811-827(in Chinese).
- 赵军平, 吴立广, 赵海坤, 2012. 西北太平洋热带气旋潜在生成指数的改进[J]. *气象科学*, 32(6):591-599. Zhao J P, Wu L G, Zhao H K, 2012. Improvement of tropical cyclone genesis potential index

- in the western North Pacific basin[J]. *J Meteor Sci*, 32(6):591-599(in Chinese).
- 周冠博, 钱奇峰, 吕心艳, 等, 2022. 人工智能在台风监测和预报中的探索与展望[J]. *气象研究与应用*, 43(2):1-8. Zhou G B, Qian Q F, Lv X Y, et al, 2022. Application and expectation of artificial intelligence in typhoon monitoring and forecasting[J]. *J Meteor Res Appl*, 43(2):1-8(in Chinese).
- Daniel S. Wilks, 2017. 大气科学中的统计方法[M]. 朱玉祥, 等, 译. 北京: 气象出版社, 173-188. Daniel S. Wilks, 2017. *Statistical Methods in the Atmospheric Sciences*. Zhu Y X, et al, trans. Beijing: China Meteorological Press, 173-188(in Chinese).
- Barnes S L, 1964. A technique for maximizing details in numerical weather map analysis[J]. *J Appl Meteor Climatol*, 3(4):396-409.
- Bister M, Emanuel K A, 1997. The genesis of Hurricane Guillermo: TEXMEX analyses and a modeling study[J]. *Mon Wea Rev*, 125(10):2662-2682.
- Chand S S, Walsh K J E, 2011. Forecasting tropical cyclone formation in the Fiji Region: a probit regression approach using Bayesian fitting[J]. *Wea Forecasting*, 26(2):150-165.
- Charney J G, Eliassen A, 1964. On the growth of the hurricane depression[J]. *J Atmos Sci*, 21(1):68-75.
- Chen L S, 2012. Research progress on the structure and intensity change for the landfalling tropical cyclones[J]. *J Trop Meteor*, 18(2):113-118.
- Cheung K K W, 2004. Large-scale environmental parameters associated with tropical cyclone formations in the western North Pacific[J]. *J Climate*, 17(3):466-484.
- Cheung K K W, Elsberry R L, 2002. Tropical cyclone formations over the western North Pacific in the navy operational global atmospheric prediction system forecasts[J]. *Wea Forecasting*, 17(4):800-820.
- DeMaria M, Kaplan J, 1994. A statistical hurricane intensity prediction scheme (SHIPS) for the Atlantic Basin[J]. *Wea Forecasting*, 9(2):209-220.
- DeMaria M, Kaplan J, 1999. An updated Statistical Hurricane Intensity Prediction Scheme (SHIPS) for the Atlantic and eastern North Pacific basins[J]. *Wea Forecasting*, 14(3):326-337.
- Dunkerton T J, Montgomery M T, Wang Z, 2009. Tropical cyclogenesis in a tropical wave critical layer: easterly waves[J]. *Atmos Chem Phys*, 9(15):5587-5646.
- Elsberry R L, 2014. Advances in research and forecasting of tropical cyclones from 1963-2013[J]. *Asia-Pac J Atmos Sci*, 50(1):3-16.
- 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.
- Emanuel K A, 1989. The finite-amplitude nature of tropical cyclogenesis[J]. *J Atmos Sci*, 46(22):3431-3456.
- Emanuel K A, 2003. Tropical cyclone[J]. *Annu Rev Earth Planet Sci*, 31:75-104.
- Emanuel K A, 2005. *Divine Wind: The History and Science of Hurricanes*[M]. New York: Oxford University Press:285.
- Emanuel K A, 2018. 100 years of progress in tropical cyclone research[J]. *Meteor Monogr*, 59(1):15. 1-15. 68.
- Fan K, 2007. New predictors and a new prediction model for the typhoon frequency over western North Pacific[J]. *Sci China Ser D: Earth Sci*, 50(9):1417-1423.
- Frank W M, 1988. Tropical cyclone formation[M]//Elsberry R L. *A Global View of Tropical Cyclones*. Arlington: Office of Naval Research:53-90.
- Fu B, Peng M S, Li T, et al, 2012. Developing versus nondeveloping disturbances for tropical cyclone formation. Part II: western North Pacific[J]. *Mon Wea Rev*, 140(4):1067-1080.
- Gall J S, Frank W M, Wheeler M C, 2010. The role of equatorial Rossby waves in tropical cyclogenesis. Part I: idealized numerical simulations in an initially quiescent background environment[J]. *Mon Wea Rev*, 138(4):1368-1382.
- Gray W M, 1968. Global view of the origin of tropical disturbances and storms[J]. *Mon Wea Rev*, 96(10):669-700.
- Kucas M E, Darlow J W E, 2012. Subjective methods for assessing tropical cyclogenesis and intensity change at the Joint Typhoon Warning Center[R/OL]//AMS 30th Conference on Hurricanes and Tropical Meteorology. <https://ams.confex.com/ams/30Hurricane/webprogram/Paper205894.html>.
- Marchok T P, 2002. How the NCEP tropical cyclone tracker works[R/OL]//25th Conference on Hurricanes and Tropical Meteorology. https://ams.confex.com/ams/25HURR/techprogram/paper_37628.htm.
- McBride J L, Zehr R, 1981. Observational analysis of tropical cyclone formation. Part II: comparison of non-developing versus developing systems[J]. *J Atmos Sci*, 38(6):1132-1151.
- Ming Y, Zhang W, Yu H, et al, 2014. An overview of the China meteorological administration tropical cyclone database[J]. *J Atmos Oceanic Technol*, 31(2):287-301.
- Ooyama K, 1969. Numerical simulation of the life cycle of tropical cyclones[J]. *J Atmos Sci*, 26(1):3-40.
- Rotunno R, Emanuel K A, 1987. An air-sea interaction theory for tropical cyclones. Part II: evolutionary study using a nonhydrostatic axisymmetric numerical model[J]. *J Atmos Sci*, 44(3):542-561.
- Wang Y Q, 2012. Recent research progress on tropical cyclone structure and intensity[J]. *Trop Cyclone Res Rev*, 1(2):254-275.
- Wang Z, Montgomery M T, Dunkerton T J, 2010. Genesis of pre-Hurricane Felix (2007). Part I: the role of the easterly wave critical layer[J]. *J Atmos Sci*, 67(6):1711-1729.
- Zehr R M, 1992. Tropical cyclogenesis in the western North Pacific[R]. Washington: NOAA:181.
- Zhang D L, Bao N, 1996. Oceanic cyclogenesis as induced by a mesoscale convective system moving offshore. Part II: genesis and thermodynamic transformation[J]. *Mon Wea Rev*, 124(10):2206-2226.

(本文责编:何晓欢)