

台风影响期间文昌航天发射场浅层风预报技术研究

贵志成¹ 王艳莹² 赵小平¹ 甘思旧¹

(1 63796部队, 西昌 615000; 2 民航海南空管分局, 海口 570100)

摘要: 利用2015—2017年台风影响文昌航天发射场期间26 d浅层风实况和预报数据, 对台风影响过程的24 h浅层风预报能力进行分析, 发现平均预报偏差较大。为提高预报技术, 利用卡方检验的方法分析台风特点与浅层风预报偏差关系, 识别出台风强度、象限和台风与场区的距离三个关键因素; 再利用2015—2016年台风影响场区实况数据, 通过逐步回归方法建立场区24 h浅层风预报回归方程。建立预报回归方程时发现台风强度、台风象限两项显著性不明显, 可将其剔除, 台风与场区距离、数值预报风速两项显著性较好。利用2017年实况数据对回归方程进行预报效果检验, 结果表明通过建立预报回归方程能减小预报偏差的标准差, 从而提高预报准确率。研究结果可以为台风影响场区时的浅层风预报提供定量参考。

关键词: 台风, 浅层风预报, 回归方程

DOI: 10.3969/j.issn.2095-1973.2020.04.019

Study of Forecasting Technique of Shallow Layer Wind over Wenchang Space Launch Site Under Typhoon Impact

Gui Zhicheng¹, Wang Yanying², Zhao Xiaoping¹, Gan Sijiu¹

(1 63796 Troops, Xichang 615000 2 Hainan Air Traffic Management Sub-Bureau, Haikou 570100)

Abstract: By comparing the shallow layer wind data and the forecasts of Wenchang Space Launch Site on the 26 days affected by typhoons in 2015-2017, the 24 h shallow layer wind forecast capacities during typhoon affecting processes were analyzed. Results show that the average forecast deviation was large. In order to improve the forecast technique, the study firstly analyzed the relationship between typhoon characteristics and shallow wind forecast deviations by using chi-square test method, which identified three key factors: typhoon intensity, quadrant and the distance between typhoon and the launch site. The regression equation of shallow wind prediction for 24 h at launch site was then established using real time data from 2015 to 2016. It was found out that the distance between typhoon and the launch site and the numerical prediction data are the main factors of the regression forecast equation, while the contribution of typhoon intensity and quadrant can be ignored. The verified result of the regression equation by using the data of 2017 shows that the forecast accuracy rate could be improved by reducing the standard deviation of the forecast deviation. This study could provide a quantitative reference for the shallow wind forecast when typhoon affects the site.

Keywords: typhoon, shallow layer wind forecast, regression equation

0 引言

浅层风是近地面层中的风, 近地面层是大气边界层最接近地表面的部分, 约为距离地表10~100 m的大气。航天发射任务中, 火箭在垂直转运和加注发射阶段, 遇到强浅层风天气, 可能会影响航天发射任务成败, 因此浅层风预报一直是航天气象保障的重点。热带气旋是最具破坏力的自然灾害之一^[1], 文昌航天发射场(以下简称场区)位于海南岛东北海岸,

年平均受到3.58个热带气旋影响, 大多数台风都是从东南方向移近海南岛, 海南岛东部海岸受台风影响最严重, 受到台风影响时间也较长^[2]。而出现强浅层风天气个例中, 热带气旋造成的浅层风风速大、大风影响时间长、风速变化规律较复杂。夏秋季场区频繁受到台风影响, 受台风影响时浅层风变化特征一直是浅层风预报研究的重点和难点。日常浅层风预报保障中, 台风影响文昌场区时浅层风风速的量化研究较不足, 台风开始影响的具体时间也较难把握, 如2016年台风莎莉嘉登陆万宁时, 24 h风速预报最大偏差为14.5 m/s。场区受台风影响时, 浅层风强风是造成设备设施的损坏、危害人员安全的主要因素, 提高

收稿日期: 2019年8月14日; 修回日期: 2020年5月18日

第一作者: 贵志成(1989—), Email: guizhicheng333@sina.com

通信作者: 王艳莹(1990—), Email: 158293284@qq.com

浅层风预报准确率可以为航天发射任务气象保障提供支撑。

目前对台风机理研究和台风引发区域大风研究的主要方法是基于动力学的数值模拟分析^[3-5]。孙学金等^[6]利用WRF-LES模式对大气边界层内大尺度湍流进行解析,提高了干旱湖区近地面风场数值模拟的真实性和准确性。冯慧敏等^[7]利用数值预报资料进行郑州市地面风的多模式预报试验,发现在数值预报的基础上,通过统计降尺度预报能进一步提高地面风预报准确率。

对基层气象预报员来说,统计学方法在计算效率和应用层面上较数值模式更具优势,在数据趋势描述方面,线性回归具有直观性和简洁性。黄奕武等^[8]利用T639数据线性回归方法,建立了台风中心预报误差与环境风整层垂直切变的24~120 h线性预估模型,且预估模型有较好的预估效果。农吉夫^[9]也通过条件数分析方法,建立了西北太平洋台风强度多元回归模型。科研单位主要研究重点为台风发生发展机理和台风结构分析,地方业务单位主要进行地面风预报,对浅层风的预报技术研究较少。风力发电站有浅层风精确预报的需求,但是大多是向气象公司寻求商业服务,利用统计学方法进行台风影响时浅层风特征及其预报技术的研究相对较少,本文针对业务工作中的实际需求,以提高预报准确率为目的,通过统计分析的方法研究台风影响场区时浅层风预报偏差的影响因子,进行因子设计试验,建立受台风影响时场区浅层风24 h预报回归方程。

1 数据来源及统计范围

浅层风实况数据是架设在文昌场区内的三个浅层风测风塔探测2015—2017年50 m高度10 min平均风向风速。资料统计范围为2015年6月—2017年10月影响场区的9个台风个例。

场区浅层风受台风影响日的定义:

1) 台风开始影响日:浅层风风速比上一日平均风速增大4 m/s或以上。

2) 台风影响结束日:当日浅层风平均风速比前一日浅层风平均风速减小4 m/s且最大风速小于10 m/s,则前一日为台风影响结束日。

2 浅层风预报偏差分析

2.1 预报概况

场区24 h浅层风预报共分为四段,分别为18:00—24:00、00:00—06:00、06:00—12:00、12:00—18:00。预报每个时段50 m高度最大风速和最大风对应风向。根据台风影响场区的定义,共筛选

出受台风影响日数26 d,每天数据分为4段,共整理了台风影响场区共26 d(104段)实况和预报数据。26 d浅层风24 h预报准确率为67%,离浅层风精准精细预报的要求还有一定差距。

2.2 预报偏差的过程能力分析

台风影响期间,由于浅层风风速较大,风向变化相对连续^[10],浅层风风向预报准确率高,按照风向的评分标准基本不会出现风向预报错误的情况,因此本文主要对风速预报偏差进行分析。

通过正态检验发现,浅层风预报偏差服从正态分布,且各数据间相互独立,可以进行过程能力分析。过程能力指数反映过程在受控状态下时,客户要求与过程表现(产品质量或服务质量的变动程度)的比值,如果过程表现越能满足顾客要求,则过程能力指数越大^[11]。

根据浅层风预报评定标准,设定目标为浅层风预报准确率大于80%,假设24 h 4段浅层风预报和实况风速偏差均小于等于3 m/s时,预报准确率为84%(在风向全部正确的情况下),满足目标要求。因此将上下限规定为 ± 3 m/s,即USL=3 m/s,LSL=-3 m/s。

通过计算得到各段浅层风预报偏差的能力直方图(图1),其中, C_p 值反映过程的潜在能力, C_p 值越大,则过程受控程度越高, C_{pk} 反映过程的实际表现,能反映出工序偏离规格中心的问题, C_{pk} 值越大,则过程偏移量越小。

浅层风预报平均绝对偏差为3.46 m/s,预报偏差组内标准差为4.72,平均偏差 C_p 值为0.21, C_{pk} 值为0.13,4段浅层风预报偏差都存在正偏倚情况,即实况风速大于预报风速出现次数较多。

3 台风关键因子的识别

从台风特点出发,考虑台风强度、台风相对场区位置和台风与场区距离三个影响因素与预报偏差的关系^[12],通过统计发现,台风与场区距离对预报偏差有正相关关系,即台风距离场区越近,越容易出现预报偏差。接下来将对台风强度、台风相对场区位置两个要素与预报偏差的关系进行分析。

3.1 台风强度与预报偏差的分析

从影响场区浅层风台风强度对预报偏差的卡方检验来看(图2),热带风暴级出现的频数最高,且对预报偏差的贡献率最大,其次为强热带风暴级影响场区时容易出现浅层风预报偏差。热带低压级虽然出现频数第二,但是对预报偏差的贡献率较小,即热带低压影响场区时浅层风预报偏差较小。总的来说不同强度台风对预报偏差有显著差异,但二者并不是呈正/负

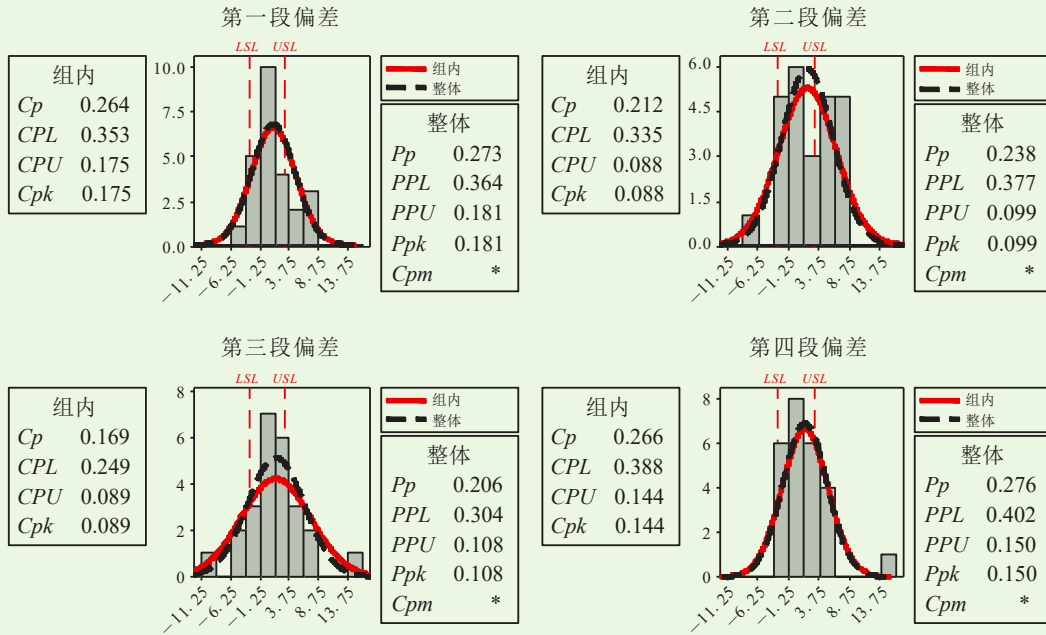


图1 浅层风预报偏差的过程能力分析

Fig. 1 Analysis of process forecast capability based on shallow layer wind forecast deviations

相关关系。

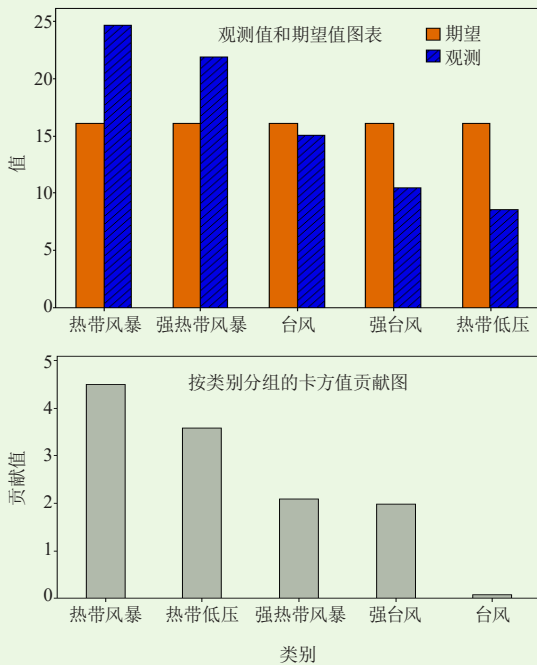


图2 影响场区浅层风预报偏差和台风强度的卡方拟合度检验

Fig. 2 The goodness of Chi-Square fit test between shallow layer wind forecast deviation and typhoon intensity

3.2 台风方位与预报偏差的分析

影响海南岛台风生成源地为南海和西太平洋, 台风主要从海上向西或西北方向移动影响场区^[13], 不同方位的台风对场区浅层风的影响有一定差别, 为了将

不同方向移近并影响场区的台风进行区分, 对受台风影响时浅层风风向进行划分, 将水平方向风向划分为八个象限, 角度分别为 ($0^\circ \leq$ 第一象限 $< 45^\circ$ 、 $45^\circ \leq$ 第二象限 $< 90^\circ$ 、 $90^\circ \leq$ 第三象限 $< 135^\circ$ 、 $135^\circ \leq$ 第四象限 $< 180^\circ$ 、 $180^\circ \leq$ 第五象限 $< 225^\circ$ 、 $225^\circ \leq$ 第六象限 $< 270^\circ$ 、 $270^\circ \leq$ 第七象限 $< 315^\circ$ 、 $315^\circ \leq$ 第八象限 $< 360^\circ$), 按照八个象限对浅层风风向进行划分, 不同的象限反映出台风相对于场区的方位。通过统计风向和预报偏差的卡方拟合度检验 (图3), 发现场区受到台风影响的风向主要在第一至第六象限, 未出现第七、第八象限风向。场区浅层风风向为七、八象限时, 台风位于场区西北方向, 很有可能已经登陆大陆地区, 台风登陆后强度减弱明显, 大风区影响范围也显著减小, 因此对场区影响也不明显。从图中分析可以发现, 风向为第一、第三象限时对浅层风偏差贡献率最大, 且出现频率也较高, 第六象限卡方值的贡献较大, 但主要为负贡献。台风影响场区时, 南风 (包括西南风、南风、东南风) 比北风 (包括西北风、北风、东北风) 出现偏差的可能性更高, 浅层风为北风时台风位于场区东侧海上, 浅层风为南风时, 台风位于场区西侧, 很有可能已经登陆海南岛, 因此台风中心距离场区位置相对于北风时更近, 风速变化幅度更大, 所以也更容易出现浅层风预报偏差。

4 台风影响时浅层风预报方程建立与检验

台风的预报主要依靠数值预报模式, 数值预报

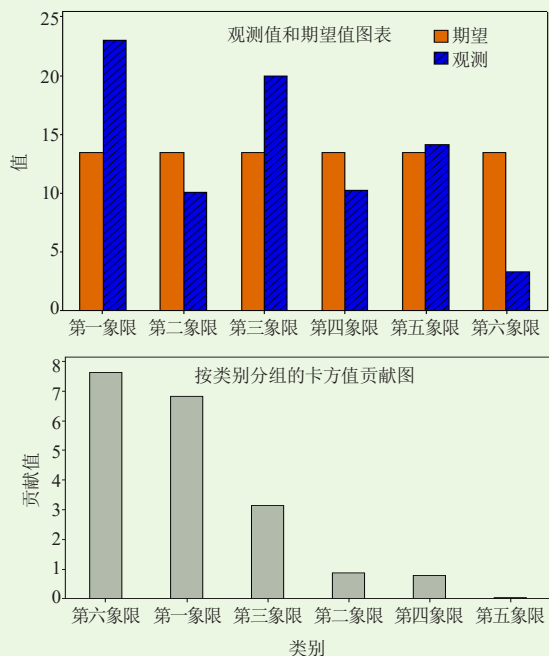


图3 影响场区浅层风预报偏差和风向的卡方拟合度检验
Fig. 3 The goodness of Chi-Square fitting test between shallow wind forecast deviation and typhoon wind direction

模式对未来24 h内台风强度的预报准确率可达80%左右^[14], 24 h内台风路径预报的偏差不超过100 km。浅层风预报中主要参考场区附近数值预报低层风速, 在研究中发现, 建立回归方程时选取925 hPa高度风场作为回归方程自变量有较好的显著性。因此在建立回归方程时考虑四个自变量, 分别为925 hPa数值预报风速、台风强度、台风象限、台风相对场区距离。

由于数值预报产品中925 hPa高度风场分辨率为 $0.25^\circ \times 0.25^\circ$, 场区位置没有处于数值预报网格点中, 因此选取场区周围4个网格点进行数据分析(图4), 4个点的经纬度分别为场区左上(19.75°N , 110.75°E)、左下(19.50°N , 110.75°E)、右上(19.75°N , 111.00°E)、右下(19.50°N , 111.00°E), 下文分别以a、b、c、d代表4个网格点, 对4点数据进行单因子试验设计。

4.1 数值预报风速的单因子检验

分别对选取的4点数据进行一天4段的相关性分析(表1), 从表中可以看到, a、b、c、d点和四点平均与实况的相关系数均在0.6以上, 且均通过水平为95%的可信度检验, 纵向来看, 数值预报与第一段实况的相关系数最大为0.81, 与第四段实况的相关系数最小为0.68, 相关系数由第一段实况向第四段实况逐渐递减。横向比较a、b、c、d点和四点平均与实况的相关系数, 发现各点之间相关系数差异较小, a点和四点平均与实况的相关系数略大。

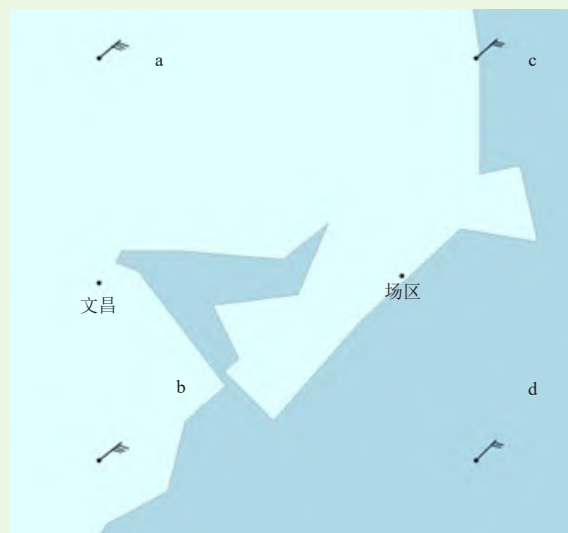


图4 场区位置及网格点的选取
Fig. 4 Location of the site and grid selection

表1 数值预报与实况序列的相关系数
Table 1 Correlation coefficient between numerical prediction and observed data

	a点	b点	c点	d点	四点平均	各段算数平均
第一段实况	0.854	0.777	0.807	0.778	0.816	0.81
第二段实况	0.741	0.752	0.746	0.712	0.742	0.74
第三段实况	0.714	0.747	0.700	0.764	0.737	0.73
第四段实况	0.679	0.657	0.673	0.683	0.687	0.68
各点算数平均	0.747	0.733	0.732	0.734	0.746	

分析各点数据与实况风速偏差的基本统计量(表2), 可以看到, b点的平均偏差值最小, 四点平均偏差数据的标准差、标准误差最小。综合考虑相关系数与基本统计量结果, 选取相关系数最大, 且偏差的标准差、标准误差最小的四点平均数据作为拟合方程中的基本数据。

表2 各点数值预报值与实况偏差的基本统计量
Table 2 Basic statistics of numerical prediction and observed data deviation in different grids

变量	N	平均值	标准差	标准误差	95%置信区间
a	68	0.231	4.018	0.502	(-0.772, 1.235)
b	68	0.137	4.076	0.510	(-0.881, 1.156)
c	68	0.450	4.049	0.506	(-0.561, 1.461)
d	68	0.731	4.038	0.505	(-0.277, 1.740)
四点平均	68	0.388	3.918	0.490	(-0.591, 1.366)

通过相关性检验可以发现, 数值预报风速与实况风速有很高的相关性, 而台风强度、台风方位、台风距离与实况风速的相关性均不超过 ± 0.3 , 因此在因子设计中主要分析这三个要因与实况风速的检验。

4.2 因子设计分析

为了确定因子对响应变量的影响, 确定关键因素

的适当设置以取得响应变量的变异最小化。对925 hPa数值预报风速、台风强度、台风方位（象限）、台风与场区距离四个因子进行因子设计。考虑到进行因子设计、选定回归模型后要进行模型效果检验，因此在进行因子分析时选取2015—2016年场区受台风影响的17 d×4段共68段数据，在模型效果检验阶段选择2017年5次台风的9 d×4段共36段数据进行效果检验。

表3 试验设计及各试验回归方程
Table 3 Tests designs and the regression equations

试验名称	自变量	不显著项	回归方程
试验1	数值预报风速、距离、强度、象限	强度、象限	$Y=13.9+0.537\times\text{风速}-0.0141\times\text{距离}-0.232\times\text{强度}-0.361\times\text{象限}$
试验2	数值预报风速、距离、象限	象限	$Y=13.3+0.519\times\text{风速}-0.0138\times\text{距离}-0.297\times\text{象限}$
试验3	数值预报风速、距离	无	$Y=12.2+0.523\times\text{风速}-0.0133\times\text{距离}$

标准差， $R-Sq$ 为回归模型误差占总误差的百分比，取值范围为0~100%，数值越大，表明回归模型数据温和越好。 $R-Sq$ （调整）为调整后的 $R-Sq$ ，与 $R-Sq$ 数值越接近，表明回归模型越可靠。各试验间 S 、 $R-Sq$ 和 $R-Sq$ （调整）差距不大，但从因子的显著性来看，试验3中各因子都能通过显著性检验，而且在回归效果差距不大的情况下，选择更少的因子可以节约计算量。

表4 各试验效果对比
Table 4 Comparison of test results

试验序号	自变量	残差标准差S	R-Sq	R-Sq (调整)
1	风速、强度、象限、距离	4.30	62.8%	60.4%
2	风速、距离、象限	4.27	62.7%	61.0%
3	风速、距离	4.26	62.3%	61.2%

4.3 试验3可靠性分析

通过逐步剔除法找到试验3为最理想回归方程，对试验3回归方程的可靠性进行分析，回归方程的显著性检验 P 值为0.0，可以通过可信度为99%的显著性检验。从回归方程的残差图中可以看到（图5），图5a为残差正态性概率图，累计的概率点列沿对角线分布，则说明残差服从正态分布，且概率点越接近对角线分布，说明回归效果越好。图5b为响应变量对应的残差图，残差点在横轴上下随机波动，即不存在某种连续性的波动趋势。图5c为残差的正态分布图，可以看到残差分布集中的区域为-4~4。图5d为按观测值顺序的残差分布，残差分布为随机波动。通过对残差图的分析，认为试验3回归方程（ $Y=12.2+0.523\times\text{风速}-0.0133\times\text{距离}$ ）稳定性较好，且方程结果不发散，是逐步剔除试验中的最优选择。

4.4 回归方程效果检验

将2017年9d 36段风速数据代入试验3回归方程

因子设计阶段共设计三组对照试验（表3），通过因子逐步剔除的方法，将回归方程中不显著项剔除，方程中自变量能通过显著性检验，则方程的稳定性较高，因此试验1至试验3的因子数量逐渐减少。在建立回归方程过程中，首先考虑回归方程的稳定性，其次考虑回归方程拟合度。

对比三组试验回归效果（表4），其中 S 为残差的

$Y=12.2+0.523\times\text{风速}-0.0133\times\text{距离}$ ，算出预报风速后与实况对比，对偏差进行过程能力分析（图6），偏差的组内标准差为3.74，整体标准差为3.68， C_p 值为0.27， C_{pk} 值为0.19。

将试验3的回归验证结果与浅层风预报偏差的过程能力分析进行对比（表5），回归方程得到预报风速有8次偏差大于等于4 m/s。改进前预报偏差有12次风速大于等于4 m/s。建立回归方程后得到的回归风速标准差比改进前减小21%， C_p 值提高28%， C_{pk} 值提高46%，说明运用回归方程的预报准确率有提升、稳定性更好。

5 结论

本文通过分析2015—2017年台风影响场区的26 d浅层风实况、24 h预报数据，分析预报偏差与台风强度、台风与场区方位的关系，找出建立回归方程的4个因子（数值预报925 hPa风速、台风强度、台风方位、台风与场区距离），通过因子试验设计，找到显著性和稳定性最高的浅层风预报回归方程，并对最优回归方程进行数据检验，通过研究发现：

1) 不同台风的强度、台风的方位对浅层风预报偏差影响不同，这种影响为非线性关系。

2) 场区周围4点算数平均的数值预报925 hPa风场数据与实况偏差最小，最适合作为建立回归方程的自变量之一。

3) 利用2015—2016年72段数据，通过因子分析试验发现，台风强度、台风方位两个因子在回归方程中显著性不明显，选取显著性、稳定性最高，且标准差最小的试验3方程作为浅层风预报回归方程。

4) 利用2017年36段数据对试验3回归方程效果进行检验，发现回归方程能有效减小预报偏差，且减少较大误差出现次数，回归效果较好。

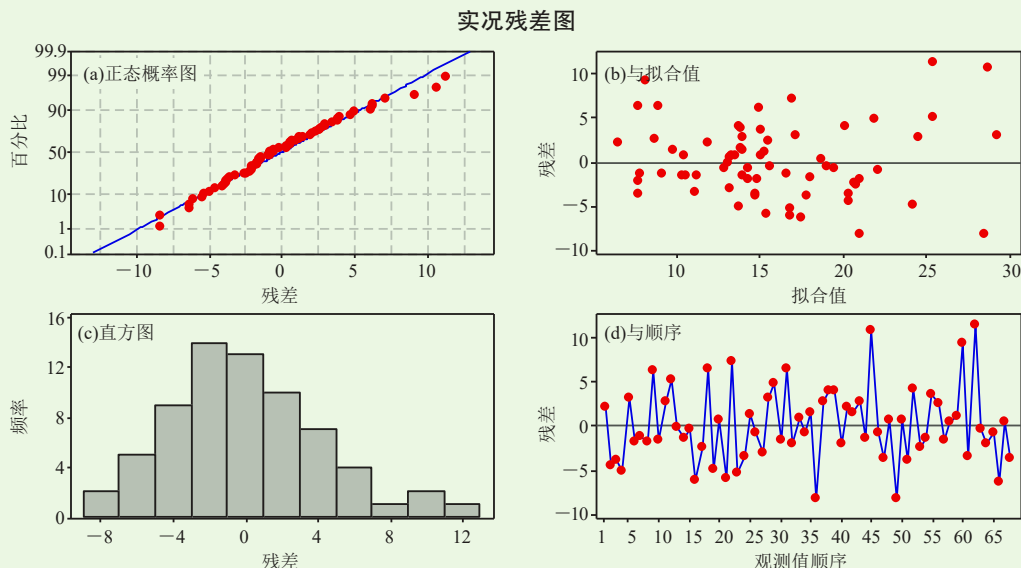


图5 试验3风速的残差四合一图

Fig. 5 Four-in-one diagram of wind speed residuals in test 3

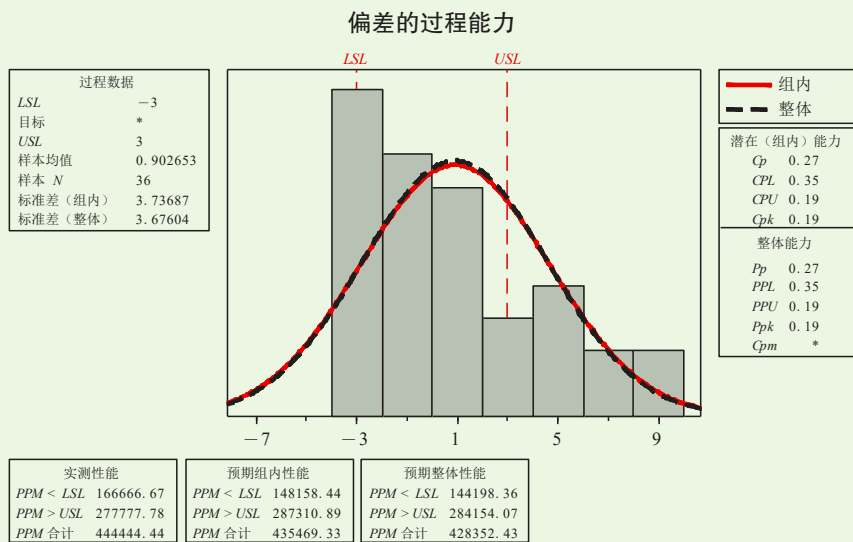


图6 试验3回归方程的效果检验

Fig. 6 The verification of regression equation in test 3

表5 试验效果对比

Table 5 Comparison of verified results

对比	标准差 (组内)	标准差 (整体)	Cp	Cpk
改进前预报偏差	4.72	4.32	0.21	0.13
试验3偏差	3.74	3.68	0.27	0.19

通过建立风速预报回归方程,可以有效提高24 h 预报准确率,但由于场区受台风影响时浅层风数据积累不够多,还需进行资料收集,通过增大样本数据个数改善回归方程。后续研究中还应进行回归方程效果评定,针对台风影响后期风速明显减小时拟合效果不好的情况进行改进提高。

参考文献

[1] Li L Q, Lan H P, Chan Z L, et al. An operational statistical scheme for tropical cyclone induced rainfall forecast. Journal of Tropical Meteorology, 2015, 21(2): 101-110.

[2] 梁海萍, 梁海燕, 车志伟, 等. 近五十年登陆海南省的热带气旋统计特征分析, 海洋预报, 2015, 32(4): 68-74.

[3] 许映龙, 张玲, 高拴柱. 我国台风预报业务的现状及思考. 气象, 2010, 36(7): 43-49.

[4] 端义宏, 陈联寿, 许映龙, 等. 我国台风监测预报预警体系的现状及建议. 中国工程科学, 2012, 14(9): 4-9.

[5] 周荣卫, 何晓凤. 模式分辨率对东北地区近地层风场预报性能的影响. 气象科技进展, 2017, 7(1): 95-100.

[6] 孙学金, 李岩, 张燕鸿, 等. 基于WRF-LES的干旱湖区近地面风场模拟与敏感性研究. 高原气象, 2017, 36(3): 835-844.

[7] 冯慧敏, 智协飞, 李荣. 郑州市地面风场的统计降尺度预报研究. 中国科技论文, 2017, 12(15): 1785-1791.

[8] 黄奕武, 高拴柱, 钱奇峰. T639台风预报误差与环境场变量的相关分析和回归分析. 气象, 2016, 42(12): 1506-1512.

[9] 农吉夫. 预报因子选择的条件数方法及其在台风强度预报中的应用. 数学的实践与认识, 2014, 44(33): 146-152.

[10] 应俊, 陈光华, 黄荣辉, 等. 西北太平洋热带气旋变性阶段强度变化的比较研究, 大气科学, 2013, 37 (4) : 773-785.

[11] 何桢. 六西格玛绿带手册. 中国人民大学出版社, 102-107.

[12] 王延青, 李晴岚, 兰红平, 等. 热带气旋影响下的深圳港湾码头最大阵风预报方法研究. 气象科技进展, 2017, 7 (1) : 128-133.

[13] 戴高菊, 文永仁, 李英. 西北太平洋热带气旋运动及其突变的若干统计特征. 热带气象学报, 2014, 30 (1) : 23-33.

[14] 何晓凤, 周荣卫, 孙逸涵. 3个全球模式对近地层风场预报能力的对比检验. 高原气象, 2014, 33(5): 1315-1322.