

文章编号: 1004-4965 (2005) 06-0651-07

热带气旋相似和最大概率集成预报方法 及其预报业务试用

林良勋, 程正泉, 黄 忠, 何夏江

(广州中心气象台, 广东 广州 510080)

摘 要: 利用建立在计算机系统的热带气旋历史资料, 经反复试验提出了一套集热带气旋发生的日期相似、地理位置相似、路径相似和强度相似等相结合的挑选“最相似”热带气旋的方案; 并利用所挑选的路径和强度“最相似”历史样本, 分别对各预报时效的移动方向、移动速度和强度进行分级概率值统计, 根据最大概率值选取相应的移向、移速和强度作为热带气旋路径和强度的预报。本方法已在计算机系统成功开发并投入业务试用, 检验结果表明: 本方法具有一定的预报能力, 其输出的未来 24、48、72 小时的路径和强度预报对预报员有一定参考价值。

关 键 词: 热带气旋; 相似和概率集成; 预报方法应用

中图分类号: P444

文献标识码: B

1 引 言

选用历史相似天气过程, 相似天气形势作预报是综合预报中的常用方法之一。在一个热带气旋整个生命史的活动过程中, 特别是以表征热带气旋运动和结构特征的路径和强度及其变化, 是多因素综合作用的结果^[1]。相似的热带气旋样本, 在一定程度上反映了造成一个热带气旋活动的等效作用。从而有理由用历史上相似的热带气旋及其路径和强度特点推断热带气旋未来可能出现的路径和强度的预报。根据文献[2, 3, 4]的研究, 有关热带气旋预报的气候持续方法主要有三类: 相似统计和回归统计(PC)及相似外推, 这些方法在改善热带气旋的业务预报方面取得了一定的效益。

但对同一目标热带气旋, 采用不同的热带气旋相似方案将有可能产生不同的历史相似样本。因此, 挑选相似热带气旋历史样本是 Cliper 方法的关键。若干“最相似”热带气旋历史样本挑选出来后, 如何利用这些样本进一步输出热带气旋路径和强度的具体预报, 采用不同的数学统计方案极有可能产生差异较大的预报结果。因此, 技术上采用何种更合理、更有效的集成方案, 仍需要作进一步探讨。

本文结合热带气旋业务预报的经验, 经反复试验和分析, 提出了一套挑选“最相似”历史热带气旋的方案。作为一种集成技术的尝试, 利用“最相似”历史样本的最大概率分布, 集成产生路径和强

收稿日期: 2004-08-24; 修订日期: 2004-11-28

基金项目: 广州中心气象台科研基金资助

作者简介: 林良勋(1959-), 男, 广东阳江人, 高级工程师, 从事天气预报及技术研究工作。E-mail: liangxun_lin@grmc.gov.cn

度的预报。

本方法已在计算机系统成功进行了可视化开发, 经过 2002 和 2003 两年的业务试用和检验, 发现本方法具有一定的预报能力和业务参考价值。

2 资料说明

挑选“最相似”热带气旋时, 使用 1949~2001 年已建立在计算机的数字化西太平洋热带气旋年鉴资料, 每隔 6 小时为单位取值, 随年鉴资料的延长, 可同步增加样本资料; 2002~2003 年方法试用和预报性能检验时, 采用目标热带气旋实时定位及强度资料。

3 历史“最相似”热带气旋的挑选

在“最相似”热带气旋的挑选方案中, 相似的判断标准不同, 相似判断流程中的先后顺序的不同, 以及流程的不同阶段所挑选的样本数的不同等, 均会造成最终挑选的“最相似”热带气旋的样本的不同, 从而造成赖于这些样本集成产生的热带气旋路径和强度预报的偏差。根据参考文献[5, 6, 7]和预报员在制作热带气旋预报时的简单人工集成思路, 经反复的分析和试验, 提出集日期相似、地理相似、强度相似、移向移速相似的“最相似”热带气旋的挑选方案(图 1)。

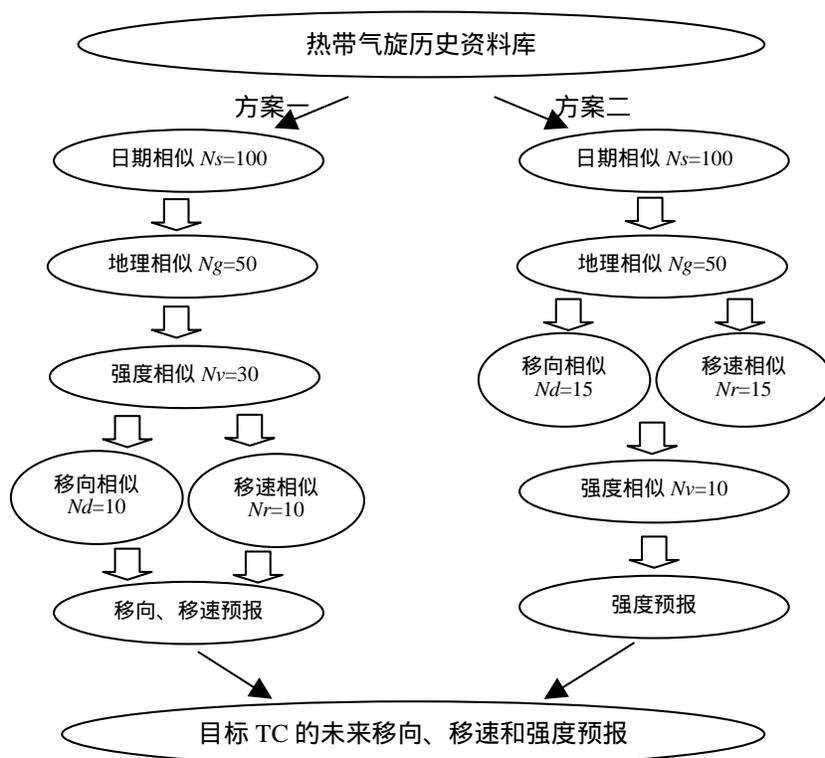


图 1 历史“最相似”热带气旋挑选及预报流程图

方案一、二分别为挑选路径和强度“最相似”气旋的流程; N 为各相似条件下的相似样本数。

3.1 路径“最相似”热带气旋的挑选

具体的相似条件和基本依据以及挑选操作方案如下。

日期相似: 不同的季节具有不同大气环流特征, 不同的大气环流形势也将导致不同的热带气旋活动规律。因此, 日期相似可以在一定程度上减小由于季节差异造成热带气旋预报的误差。经过试验提

出在热带气旋的起报日前后各延伸 10 天的所有历史热带气旋样本,符合日期相似。为下一步的相似判断提供足够的样本数,必要时适当再延伸相似日期以保证 N_S 大于 100。

地理位置相似:热带气旋活动与地理位置密切相关。不同的下垫面对热带气旋的影响明显不同;在不同的经度或纬度上的热带气旋,有其明显的地域性活动规律。期望地理相似能较好地把热带气旋在不同的地理位置的活动规律和特点反映到热带气旋的路径和强度预报上。将热带气旋的起报点定义为地理位置相似的参考点 (R_f),对满足日期相似的所有历史热带气旋分别找出其与 R_f 之间最近的点 R_i ($i=1, 2, \dots, N_S$),并计算 R_f 和 R_i 之间的距离 D_i 。取 D_i 值最小的前 50 个相对应的热带气旋,并定义该 50 个热带气旋符合地理位置相似 ($N_g=50$)。

强度相似:热带气旋的强弱与热带气旋的路径相关;同一热带气旋在不同的强度阶段其路径也会随之发生变化。因此,在进行路径相似挑选时,对强度的相似判断应给予足够的重视。本方案取中心最大风速 V 作为强度参考值,用参考点 R_f 的中心最大风速 V_f 与符合地理相似的 50 个热带气旋最近的点 R_i 的最大风速 V 进行比较,取 $|V_f - V|$ 值最小的前 30 个相对应的热带气旋作为符合强度相似的样本 ($N_v=30$)。

移向相似:移向相似热带气旋从满足强度相似的 30 个样本中产生:分别计算参考点 R_f 和最近点 R_i 的前 24 小时的移动方向 M_f 和 M_i ($i=1, 2, \dots, N_v$),取 $M_f - M_i$ 的绝对值最小的前 10 个相对应的热带气旋作为移动方向最相似的样本 ($N_d=10$)。

移速相似:移速相似热带气旋同样在满足强度相似的 30 个样本中产生:分别计算参考点 R_f 和最近点 R_i 的前 24 小时的移动方向 R_f 和 R_i ($i=1, 2, \dots, N_v$),取 $R_f - R_i$ 的绝对值最小的前 10 个相对应的热带气旋作为移动速度最相似的样本 ($N_r=10$)。

3.2 强度“最相似”热带气旋的挑选

强度“最相似”热带气旋的相似条件和依据相同,但流程上稍有不同,其中移向、移速相似样本在满足地理位置相似的 50 个样本中产生,分别取 15 个;在产生的 30 样本中,取其中 $|V_f - V|$ 值最小的前 10 个相对应的热带气旋作为符合强度相似的样本。

通过上述流程产生的移向移速和强度相似各 10 个样本是最终历史“最相似”热带气旋,也是制作目标热带气旋路径和强度预报的样本。

4 热带气旋路径和强度预报的集成方案

挑选出历史“最相似”热带气旋是相似预报方法的基础,采用何种数学统计方案,充分利用“最相似”热带气旋所包含的信息,特别是如何从“最相似”热带气旋样本中提取对预报有意义的信息,期望使预报的误差达到最小,是相似预报方法的关键^[8]。

无论是热带气旋的移动方向、移动速度,还是中心最大风速,它们在未来时刻的预报值,完全取决于它们各自 10 个“最相似”的历史样本。为了从这 10 个历史样本中找出对预报有意义的信息,从而提出合理、可行的预报集成方案。经过反复的统计分析和试验,如相似样本间差异的最大和最小值、相似程度权重、平均值、概率分布等。通过比较,最终对 10 个历史样本的集成平均给予了重视。显然,集成平均是一个广为使用的统计值,它到底在多大程度上提供对预报制作有意义的信息,仍是一个有待进一步探讨的问题;但普遍认为,当通过均方根差检验时,集成平均有较大的参考价值^[8]。直接用 10 个集成样本的平均值进行试报发现:在路径和强度所发生的变化不大时,预报是有参考价值的;但一旦热带气旋未来的路径和强度与参考点 R_f 相比变化较大时,预报的可信度减小。为了加大后者的预报可信度,提高方法的整体预报能力,提出了以 10 个集成样本平均值为基础,分别对热带气旋的移向、

移速和强度进行分级概率值统计, 根据它们的最大概率值选取相应的移向、移速和强度作为热带气旋路径和强度的预报方案。

4.1 路径预报集成方案

热带气旋的路径预报分解为移向预报和移速预报。分别以移动方向“最相似”10个热带气旋样本最近点 R_i 为参考点, 针对不同的预报时效, 分别计算最近点 R_i 点的后面各时效的10个样本移向和它们的平均移向 $\bar{m}_{0,t}$, 并求出10个“最相似”样本在平均移向 $\bar{m}_{0,t} \pm C_m$ (C_m 为等于 10^0 的常数) 范围内出现的概率 $P_{0,t}$ 。其中 t 是预报时效, 分别为未来24、48、72小时三个时效。

分别给不同预报时效的平均移动方向 $\bar{m}_{0,t}$ 若干个扰动值 $m_{i,t}$, 即

$$\bar{m}_{i,t} = \bar{m}_{0,t} + m_{i,t} \quad (1)$$

式中 $\bar{m}_{i,t}$ 定义为不同预报时效的扰动平均移动方向, i 和 $m_{i,t}$ 的取值见表1。

表1 各预报时效移动方向扰动值 $m_{i,t}$ 取值

i	0	1	2	3	4	5	6	7	8	9	10
$m_{i,t}$	0	-5°	+5°	-10°	+10°	-15°	+15°	-20°	+20°	-25°	+25°

同 $P_{0,t}$ 值的计算方法, 根据 $\bar{m}_{i,t}$ 的值分别求出其相应 $\bar{m}_{i,t} \pm C_m$ 范围内出现的概率 $P_{i,t}$ 。然后从11个分级概率值中找出其最大概率值 $P_{M,t}$ 。即: $P_{M,t} = \text{Max}(P_{i,t})$ 式中 $i=0, 1, 2, \dots, 10$, 此时 $P_{M,t}$ 的值可能是唯一的, 也可能出现若干个概率值同时达到最大。如果 $P_{M,t}$ 的值是唯一时, 取其相应的扰动平均移动方向 $\bar{m}_{i,t}$ 的值作为热带气旋相应预报时效的移动方向预报值; 如果有二个概率值同时达到最大时, 从二个相应的扰动平均移动方向 $\bar{m}_{i,t}$ 中, 取其中最接近平均移动方向 $\bar{m}_{0,t}$ 的值作为热带气旋相应预报时效的移动方向预报值; 否则, 取平均移动方向 $\bar{m}_{0,t}$ 的值作为热带气旋相应预报时效的移动方向预报值。

移速预报以移动速度“最相似”的10个热带气旋作为集成样本, 预报的集成处理与上述移动方向的集成思路相同, 有关变量的取值说明如下。

不同预报时效的扰动平均移动速度 $\bar{r}_{i,t}$ 由下式给出

$$\bar{r}_{i,t} = \bar{r}_{0,t} + r_{i,t} \quad (2)$$

式(2)中的 i 和扰动值 $r_{i,t}$ 以及移动速度的分级概率统计范围 $\bar{r}_{i,t} \pm C_r$ 中 C_r 的取值见表2。

表2 各预报时效移动速度扰动值 $r_{i,t}$ 和概率分级范围 C_r 取值 (km/h)

i	0	1	2	3	4	5	6	7	8	9	10
$r_{i,t}$	0	-3	+3	-6	+6	-9	+9	-12	+12	-15	+15
C_r	3	3	3	3	3	4	4	4	4	5	5

移动速度的最大概率值 $P_{r,t}$ 的产生和处理同 $P_{M,t}$ 。

4.2 强度预报集成方案

强度预报以强度“最相似”10个热带气旋作为集成样本, 针对不同的预报时效, 分别以它们的最近点 R_i 为参考点, 读取 R_i 点后各时效的最大风速 V_i 和计算它们的平均最大风速 $\bar{v}_{0,t}$, 并求出10个强度“最相似”热带气旋样本 $\bar{v}_{0,t} \pm C_v$ ($C_v=3$ m/s) 范围内出现的概率 $P_{0,t}$ 。

参照式(1), 不同预报时效的扰动平均最大风速 $\bar{v}_{i,t}$ 由下式产生

$$\bar{v}_{i,t} = \bar{v}_{0,t} + v_{i,t} \quad (3)$$

表 3 各预报时效最大风速扰动值 $v_{i,t}$ 以及概率分级范围 C_v (试验) 取值 (m/s)

i	0	1	2	3	4	5	6	7	8	9	10	11	12
$\bar{v}_{i,t}$	0	-3	+3	-6	+6	-9	+9	-12	+12	-15	+15	-18	+18
C_v	3	3	3	3	3	4	4	4	4	5	5	5	5

式(3)中的 i 和 $\bar{v}_{i,t}$ 以及分级概率的范围 $\bar{v}_{i,t} \pm C_v$ 中的常数 C_v 的取值由表 3 给出。

同理求出 $\bar{v}_{i,t} \pm C_v$ 范围内出现的概率 $P_{i,t}$ 。然后从 13 个分级概率值中找出其最大概率值 $P_{v,t}$ 。最大概率值的处理和最大风速预报值的产生同上述移动方向预报的集成技术。

应该说明的是：无论是移动方向和移动速度，还是中心最大风速，试验表明它们的扰动值绝大多数情况下在 $i \leq 6$ 的范围取值，而在方案中设计 i 的取值范围远大于 6；另外，中心最大风速的扰动值 $\bar{v}_{i,t}$ 的取值与移动方向的扰动值 $m_{i,t}$ 和移动速度的扰动值 $r_{i,t}$ 取值各有不同；概率统计的分级范围 C_m 、 C_r 和 C_v 也不尽相同， C_m 是一个定常量， C_r 和 C_v 是随 i 的增大而出现分级增大的变常量。这种设计是经过对各种类型热带气旋进行反复试验的基础上选定的。尽管所挑选出来的热带气旋已是“最相似”的样本，应该承认这种“最相似”仅仅发生在最近点 R_i 前， R_i 点之后发生明显变化的可能性客观存在。这种取值的设计是期望把离散度较大的概率事件反映出来，特别是对制作时效较长的预报更为必要。

另外，在预报集成时是读取最近点 R_i 之后的相关资料，在制作预报时，特别是制作时效较长的预报时，极可能因为有的历史相似热带气旋样本已减弱消亡，而出现无法全部获取 10 个历史相似样本的相关资料的情况。为此，经试验后提出了如下特殊约定：某一预报时效缺 8 个或以上样本的相关资料，则预报该时效热带气旋减弱消亡；缺 4~7 个样本的相关资料，以集成平均值作为预报值；缺 3 个或以下样本的相关资料，仍按实际获取的样本数进行预报集成。

5 方法的试用及预报性能检验

本方法已在计算机系统上成功地进行了可视化开发，并投入业务试用。其输出的产品含热带气旋未来 24、48、72 小时的预报位置的经纬度和中心最大风速。经过对 2002~2003 年在广东省气象台预报区 (25°N 以南、130°E 以西洋面) 内的 24 个热带气旋进行的业务试用，并与国内各综合预报、客观预报和数值预报方法的两年平均路径预报误差和强度预报精度进行了对比，发现本方法具有一定的预报能力。

5.1 路径预报结果检验分析

图 2 给出了按不同路径各预报时效路径预报距离平均误差和 2002、2003 年国内综合预报、客观预报和数值预报方法平均距离误差值。与各客观预报方法的平均距离误差相比，本方法的 24 小时的预报总平均误差偏小 10 km，但 48 小时的误差明显偏低；与各数值预报方法的平均距离误差相比，本方法的 24 小时的误差较为接近，而 48 小时的误差偏大 24 km。由于业务发布的各类预报方法尚未对时效为 72 小时的预报进行评定，所以本方法该时效的预报误差暂无法与之相比，但 72 小时的路径平均距离误差为 483 km，表明本方法较长时效的路径预报仍然稳定和具有较大的参考价值。另外，本方法对“西行”和“北到东北行”类路径预报距离误差相对较小，对“抛物线”类路径距离误差最大；对快速移动的热带气旋，其移速预报有偏慢现象；对个别转向热带气旋，对转向现象的预报反应略偏慢。

上海台风研究所. 2002 年热带气旋定位和预报精度评定. 2002. 12

上海台风研究所. 2003 年热带气旋定位和预报精度评定. 2003. 12

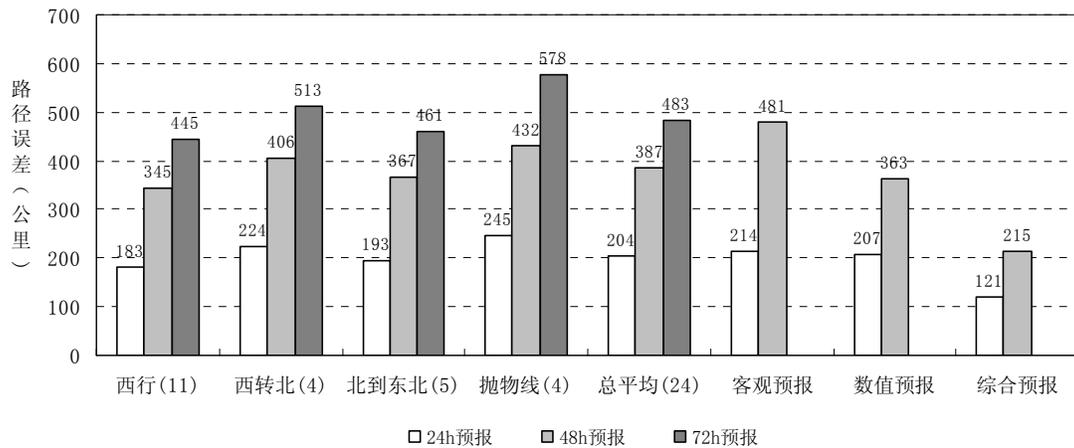


图2 不同路径各预报时效路径预报距离平均误差与国内其它预报方法平均误差对比

为进一步了解方法的路径预报性能,把参与检验的24个气旋按误差低于平均值(<20 km以上)接近平均值(±20 km以内)高于平均值(>20 km)进行了分类分析,同样发现了上述不同路径的误差分布。另外,误差低于和接近平均值的气旋,其移速和移向较稳定;而对误差高于平均值的气旋,方法的较大误差主要出现在:路径异常的目标热带气旋移向预报误差较大;气旋转向前预报反映偏迟;快速东北移时移速预报偏慢。这可能与预报集成方案中未充分考虑非气候因素造成热带气旋路径变化有关。

5.2 强度预报结果检验分析

为了解本方法对热带气旋强度的预报能力,对预报的风速误差、准确率、强度趋势命中率三方面进行了检验和统计(表4)。

从表4可见,与唯一参与评定的上海统计释用方法比,本方法的24小时中心最大风速预报误差与其相当,48小时的误差略有偏小。但中心最大风速预报值略比实况值偏小(表略)。本方法在24小时的“趋势报对率”保持相对较高的水平,并和上海统计释用方法相似,“趋势报对率”均有随预报时效延长而提高的特点。表4中强度预报准确率 P 定义为: $P = F - A / A \cdot 100\%$ 。式中 F 为中心最大风速预报值, A 为相应实况值。

表4 各预报时效中心最大风速预报和相关综合预报及客观预报精度

时效/h	风速误差 / (m · s ⁻¹)	准确率 P / %	趋势报对 / %	趋势报错 / %	趋势报平 / %
24	5.2 (4.1, 5.4)	82.8	70.2(76.9, 58.9)	17.0	13.1
48	7.3 (5.6, 8.5)	77.0	71.2(74.0, 71.1)	15.8	13.0
72	10.7 (-, -)	73.0	77.3(-, -)	14.3	8.4

注:表中括号内的数值分别为中央气象台综合预报误差和上海统计释用方法的预报误差。

预报检验还发现,对不同海区活动的热带气旋或热带气旋移到不同的海域,甚至在登陆前的近海,路径和强度预报均表现出较稳定的预报性能。这可能与方法中采用最近距离的地理相似条件和相似样本在 R_i 点以后的有效集成资料有关。

应该指出,至今尚无客观预报方法甚至数值预报方法的路径和强度预报误差低于综合预报误差,本方法也不例外。这点不难解释,因为综合预报是在参考了各种客观预报和数值预报结果后,通过人工进行超级集合产生的预报结果。

6 小结与讨论

经过试验和分析发现,挑选相似的热带气旋样本的流程方案和从相似样本中采用的数学统计方案是相似集成方法的关键。

本文提出的历史最相似热带气旋样本的挑选流程和最大概率集成预报方案仅仅是一种尝试,试用及检验结果初步表明,本方法在路径和强度预报均具有一定的预报能力,作为一种气候方法,在日常预报业务中有一定的参考价值,特别是对未来 48 和 72 小时的长时效预报其参考价值较大。

由于其具有设计思路与预报员的预报思路较为接近的特点,比较容易被预报业务技术人员所接受。

本方法也同样存在气候方法的不足和有待进一步完善之处,如对热带气旋路径和强度的突变,预报反应较慢,异常路径的预报能力偏低等。

就本方法而言,如果有数字化的历史和实况天气形势资料库,有条件把“形势相似”加入相似方案中,相信对提供预报精度会有较大的帮助。

参 考 文 献:

- [1] 陈联寿,丁一汇. 西太平洋台风概论[M]. 北京:科学出版社,1979. 388-391.
- [2] 何夏江,曾琮,许永鏖,等. 热带气旋路径强度预报——SAPC法[J]. 应用气象学报,1996,2(1):45-52.
- [3] 吴中海,徐一鸣,李多武,等. 西太平洋台风路径预报的气候持续性模式[A]. 气象科学技术集刊(5)[C]. 北京:气象出版社,1983. 52-60.
- [4] 霍金JR,纽曼C. 台风路径的一个相似方法——HURRAN[A]. 台风及其预报[C]. 北京:科学出版社,1975. 142-152.
- [5] 何夏江,黄忠,林良勋,等. 华南台风业务系统[J]. 气象,1995(8):13-16.
- [6] 刘春霞. 广东登陆热带气旋活动异常成因分析[J]. 热带气象学报,2004,20(2):161-166.
- [7] 梁建茵. 我国热带气旋登陆时间日变化特征分析[J]. 热带气象学报,2003,19(增):160-165.
- [8] 林良勋. 集合预报及其产品应用[J]. 广东气象,2001(2):44-46.

METHODS OF ANALOGY AND MAXIMUM PROBABILITY ENSEMBLE IN THE FORECAST OF TROPICAL CYCLONES ON AND OPERATIONAL APPLICATION

LIN Liang-xun, CHENG Zheng-quan, HUANG zhong, HE Xia-jiang

(Guangdong Meteorological Observatory, Guangzhou 510080, China)

Abstract: In this paper, a method for tropical cyclone (TC) track and intensity forecast is proposed. Based on the similarity in date, location, track and intensity, the method retrieves the most similar TCs from the historical samples in computer database. The retrieved samples are classified and analyzed statistically in terms of their future direction, speed and intensity. And the TC track and intensity with maximum probability assembly are chosen as forecast results. The method has already been implemented and utilized in operational application. The forecast ability of this method is demonstrated with test results.

Key words: TC forecast; analogy and maximum probability assembly; method and application