

热带气旋路径客观综合集成预报方法的研制¹⁾

谢玲娟²⁾

(广州热带海洋气象研究所, 510080)

提 要

在正确评估我国现有发报的7种客观预报方法性能的基础上^[1], 分别对西北太平洋、东海、南海3个责任海区, 采用典型相关、条件概率组合、新回归系数估计法等多种统计模型, 建立包括数值预报产品在内的热带气旋路径的综合集成预报方法。

关键词: 热带气旋路径 集成预报 数值预报产品

引 言

热带气旋路径业务预报方法的精确度直接影响热带气旋预报服务的精确性。多年来虽然建立了一系列热带气旋路径客观预报方法, 但类别较多, 其预报结果不尽相同, 尚没有一个占绝对优势的客观预报模式, 缺乏一个具有较高水平的客观预报集成模式投入业务应用。为此, 在八五科研计划期间, 试图挑选几种应用方便、效果较好的客观预报方法, 分别对西北太平洋、东海、南海3个责任海区, 采用典型相关、条件概率组合、新的回归系数估计法等多种统计模型加以集成, 建立包括数值预报产品在内的热带气旋路径的综合集成预报方法, 以弥补目前台风客观预报工作的某些不足。从其拟合结果及试报情况来看, 已取得令人鼓舞的结果。

1 防区范围及资料选取

15°N 以北、115—140°E 属西北太平洋责任区; 15—30°N、120—132°E 属东海责任区; 10—25°N、110—130°E 属南海责任区。适用期为每年6月15日—10月15日, 预报时效达60小时。

考虑到各子方法在相同时期内均能作出预报(日本台风数值预报模式是1988年开始), 因此资料选用1988—1992年共5年的热带气旋资料, 按3个海区范围分别选取361、202、151个样本。

2 集成模式建立的关键技术

以往, 一般的集成预报大多采用平均法或加权平均法, 而没有涉及到集成的最大难点, 即如何处理权重系数的变化和分配这一实质性的问题。本文针对这一问题提出一些新的设想和技术方法, 且在此基础上作初步探讨, 试验证明具有一定的效果。

2.1 典型相关法³⁾

(1) 先求出组成子方法的协方差矩阵。

$$S_{xx} = \begin{bmatrix} s_{11} & s_{12} & s_{13} & s_{14} \\ s_{21} & s_{22} & s_{23} & s_{24} \\ s_{31} & s_{32} & s_{33} & s_{34} \\ s_{41} & s_{42} & s_{43} & s_{44} \end{bmatrix}$$

下标1—4为子方法序号。

(2) 求 $S_{xx}^{-1} \cdot S_{xy} \cdot S_{yy}^{-1} \cdot S_{yx}$ 即4个矩阵乘积。

S_{xx}^{-1}, S_{yy}^{-1} 为逆矩阵

1) 本文为85-906-05-04-02子专题成果;

2) 为本文执笔。参加此项工作的还有吴天泉、丁长根、刘近、薛宗元、陈雪峰、万齐林等人;

3) 上海台风所集成课题组, 用典型相关试作集成预报。

$$S_{xy} = \begin{bmatrix} S_{1y} \\ S_{2y} \\ S_{3y} \\ S_{4y} \end{bmatrix}, S_{yx} = [S_{y1} \quad S_{y2} \quad S_{y3} \quad S_{y4}]$$

(3)求 F 最大特征值与特征向量。

$$F = \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \end{bmatrix} \cdot S_{yx} \quad (F \text{ 为 } 4 \text{ 行 } 4 \text{ 列矩阵})$$

其中,

$$E = \begin{bmatrix} D_1 \\ D_2 \\ D_3 \\ D_4 \end{bmatrix} \cdot 1/S_{yy}, \quad D = \begin{bmatrix} D_1 \\ D_2 \\ D_3 \\ D_4 \end{bmatrix}$$

(4)以 4 个子方法预报值与最大特征值的特征向量相乘作因子,再与实况求全回归方程。

2.2 条件概率组合⁴⁾

一般多元线性回归方程为:

$$y = \sum_{m=1}^M A_m x_m \quad (1)$$

式中, $x_m (m=1, 2, \dots, M)$ 为子方法的预报结果, $A_m (m=1, 2, \dots, M)$ 是自变量 x_m 的集成权重系数。事实上,对于不同的样本,由于受内外条件和随机因素的影响,系数 A_m 应该不断变化。为此制定某一判别标准,对每个独立样本的自变量 x_m 重新进行排序,由排序结果依次决定权重系数 A_m 。而该判别标准则采用条件概率值为依据,对自变量进行排序。

条件概率公式:

$$P(A/B) = z_i'/z_i \quad (i = 1, 2, \dots, I) \quad (2)$$

式中, A 为实况出现在第 i 级条件; B 为子方法预报第 i 级条件; P 为预报 i 级而实况出现 i 级的条件概率; z_i 为该子方法预报 i 级的总次数; z_i' 为该子方法预报 i 级而实况出现 i 级的总次数。

对于每种子方法可求得 I 个条件概率,如有 M 种子方法,可从 M 个不同矩阵中找到 M 个相应的条件概率值,它组成 M 维条件概率向量

$$\bar{P}' = (P_1', P_2', \dots, P_M') \quad (3)$$

式中 $P_1' \geq P_2' \geq P_3' \geq \dots \geq P_{M-1}' \geq P_M'$ 关系成立,则式(3)为最后排序结果。

在建立集成方程时,对历史样本按上式原则做排序处理,然后用最小二乘法求出系数 A_m 。在业务预报时,按同样原则先对子方法的预报结果排序,决定其在方程中的位置,进行集成。

2.3 新回归系数估计法^{5)[2]}

该方法以加权回归为基础,通过不断修改权重系数,对回归系数的估计值加以订正。

回归系数估计值

$$B = (\bar{X}'W\bar{X})^{-1}\bar{X}'W\bar{Y} \quad (4)$$

其中,

$$\bar{X} = \begin{bmatrix} x_{11} - \bar{x}_1 & x_{12} - \bar{x}_2 & \dots & x_{1m} - \bar{x}_m \\ x_{21} - \bar{x}_1 & x_{22} - \bar{x}_2 & \dots & x_{2m} - \bar{x}_m \\ \vdots & \vdots & & \vdots \\ x_{n1} - \bar{x}_1 & x_{n2} - \bar{x}_2 & \dots & x_{nm} - \bar{x}_m \end{bmatrix} \quad (5)$$

$$\bar{Y} = (y_1 - \bar{y} \quad y_2 - \bar{y} \quad \dots \quad y_n - \bar{y})' \quad (6)$$

定义 Cook 距离

$$D_i = \frac{h_{ii}}{(m+1)(1-h_{ii}-\omega_i)^2} \left(\frac{\delta_i^2}{\delta^2} \right) \quad (7)$$

其中, h_{ii} 为帽子矩阵对角元,它刻划在空间第 i 个试验点试验中心 \bar{X} 的远近, δ_i 为偏差估计值, δ 为方差 σ 的无偏估计。

当 $\omega_i = 1 (i=1, 2, \dots, n)$ 按式(1)求得 b_0, b_1, \dots, b_n 的初始值,再按式(5)求出对应 Cook 距离,由此再修改权重系数,使强影响点所占比重较小。权重系数 $\omega_i = \exp(-D_i/4\bar{D})$, \bar{D} 为 D_i 的均值。按修改后的 ω_i ,再由式(4)求得系数值 b_0, b_1, \dots, b_m ,则为上次求得的 $b_0, b_1,$

4) 浙江省气象台,东海区域热带气路径客观综合集成方法。

5) 广州热带所,试用一种新的回归系数估计法制作南海责任区的热带气旋路径集成预报。

..., b_m 的订正值。如此迭代到 $\max |\Delta \hat{B}_j| < \epsilon$ 为止 (ϵ 为预定给出的精度), 此时求得的 b_0, b_1, \dots, b_m 即为回归方程系数的估计值。

这种估计方案较普通回归有明显优点, 它能够处理各强影响点, 同时也能克服 M-估计所不能克服的高杆点对回归方程的影响。为此称之为“新的回归系数估计方法”。

3 拟合结果分析

表 1 列出 3 个责任海区集成预报的拟合结果。从表中可见预报精度已达到考核目标, 历史拟合率较高。

为进一步检验方程的稳定性和可预报性, 对所建方程作一系列的统计分析, 同时对其预报效果作进一步对比分析, 复相关系数高达 0.90 以上, 方差缩减大于 0.80, 且预报因子和预报量相关均通过 $\alpha=0.05$ 的信度检验, 表明方程具有显著性(表略)。

表 1 集成预报拟合结果
(1982—1992 年, 单位: km)

责任海区	24h		48h		60h	
	样本	误差	样本	误差	样本	误差
西太平洋	361	113	361	245	361	340
东海	202	149	202	371	193	520
南海	151	154	151	327	114	546*

* 为 72 小时预报误差

4 1993—1994 年实例检验分析

通常拟合效果好, 往往在预报时并不理想, 为此对 1993 年的热带气旋 9302、9309、9306、9307、9313、9315、9316 按不同海区进行了预报试验(表 2), 其效果不错。在 1994 年台汛期, 将该项科研的初步成果投入了业务预报试验工作中, 对进入 3 个责任海区的热带气旋分别进行 28 次、26 次、31 次的计算

表 2 集成预报 1993 年试报误差/km

责任海区	24h		48h		60h	
	样本	误差	样本	误差	样本	误差
西太平洋	12	114.4	12	317.2	8	439.6
东海*	/	/	/	/	/	/
南海	12	153	9	348	6	550**

* 1993 年进入该区的热带气旋不多, 未能计算;

** 72 小时预报

预报, 结果为 24h 的 ≤ 160 km、48h 的 ≤ 350 km、60h 的 ≤ 460 km, 效果令人满意(见表 3)。

表 3 集成预报 1994 年台风期试报误差/km

责任海区	24h		48h		60h	
	样本	误差	样本	误差	样本	误差
西太平洋	28	139.2	23	308.8	18	361.9
东海	26	156	26	279	18	407
南海	31	154	21	341	9	451*

* 为 72 小时预报

图 1、2 以 9403、9406 号热带气旋为例, 说明集成预报与实况路径对照其预报是成功的。

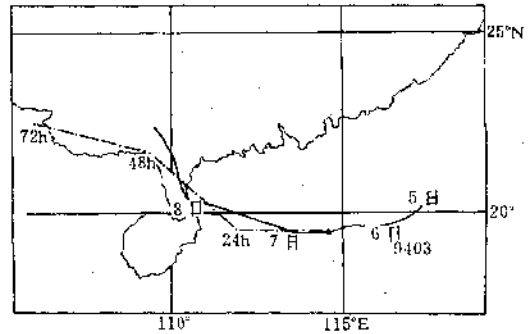


图 1 9403 号台风路径实况(实线)与预报(虚线)对照图

5 社会经济效益和推广前景

5.1 1994 年登陆我国 12 个热带气旋中除一个在山东以外, 有 6 个在广东阳江—海南三亚一带沿海登陆, 4 个在浙闽沿海登陆。在我们参加试验的台风中, 对登陆地点的预报基本准确。多数在 36—48 小时前报出, 及时为业务部门提供预报参考依据, 效果较好。

5.2 本课题完成后, 将有一个计算机自动化系统投入日常预报业务中运行, 它综合了现代科技成果和资料并与气象专家宝贵经验相结合, 较明显提高对台风灾害性天气的监测预报水平, 以便于有关领导做好决策指挥, 加强各条生产战线上的管理、保护。

5.3 本成果可应用于我国气象部门外, 还可以在海洋、航运、渔业及军事等部门推广应用。

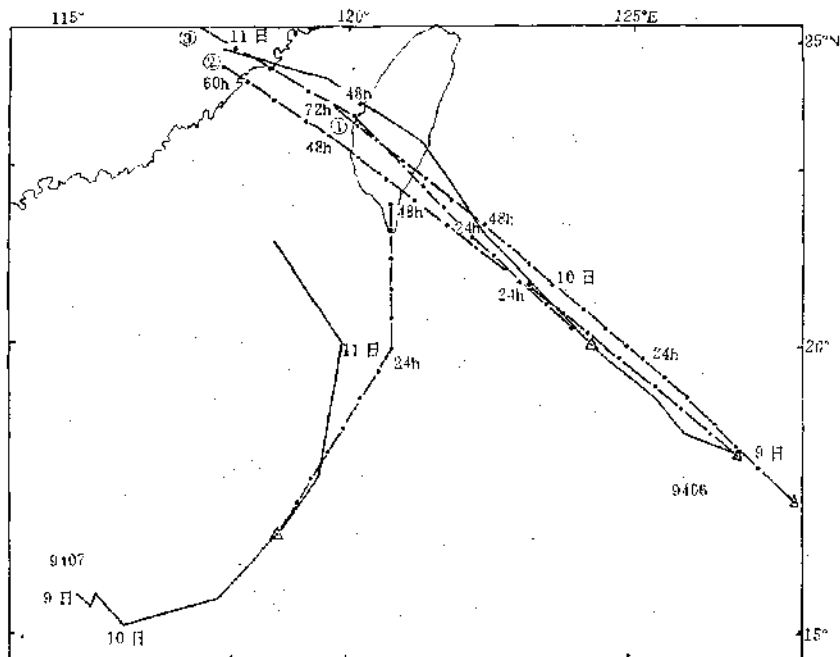


图2 9406、9407号台风路径实况(实线)与集成预报(虚线)对照图
①、②、③为9406号台风路径的3次集成预报

6 小结

综上所述,本模式具有一定的可预报性和稳定性,预报精度的进一步提高,对业务预报具有较好的参考价值。

由于模式建立不久,尚需在实践中进一步检验、不断修正,以完善方案。

参考文献

- 1 史久恩,任福民,西北太平洋热带气旋路径预报的客观评估和非线性集成方法的研究,待出版。
- 2 么枕生,气候学研究——统计气候学,北京:气象出版社,104—112。

An Objective Integrated Prediction Model of Tropical Cyclone Track

Xie Lingjuan

(Guangzhou Institute of Tropical and Oceanic Meteorology, 510080)

Abstract

Following correct assessment of seven objective prediction methods used in China, employing a number of statistic models such as the typical correlation, conditional combination of probability and updated method for estimating regressive coefficients, a package of composite procedures for the tropical cyclone track forecasting (including NWP product) are formulated for areas in the northwestern pacific, East China Sea and South China Sea.

Key Words: tracks of tropical cyclone NWP product integrated prediction model