第 54 卷 第 6 期　　气 象 学 报　　Vol. 54, No. 6
1996 年 12 月　　ACTA METEOROLOGICA SINICA　　Dec., 1996

多维动态关联模型在台风路径、强度和风速同时预报中的应用研究*

吕纯濂　陈舜华　　朱永禔

（南京气象学院，南京，210044）（上海市气象局，上海，200030）

摘　要

从台风个例中采集时序资料，利用关联模型作台风路径、强度和风速的 24h、48h 和 72h 预报。在以往工作的基础上，对模型的计算方案作了改进，考虑了内生变量协方差阵的影响。正式提出"多维气象动态关联模型"的新概念，并做了 3 方面的工作：1. 3 种时效的回报和预报及其统计分析；2. 模拟观测误差产生均匀分布的随机数，作了若干次随机模拟试验，初步讨论了模型的稳定性；3. 就台风个例中反映出来的问题，提出了有关本模型的一些待解决的问题和建设性的意见。

关键词：多维动态关联模型，台风，两段最小二乘估计。

1　引　言

以往的气象统计预报常用多元线性回归模型，即由若干预报因子和一个预报对象建立预报模型。当需要对两个以上的预报对象同时进行统计预报时，需分别独立建立相应个数的预报方程，各预报方程中的预报因子各自和对应的预报量匹配，互不相干，且各预报量间彼此独立，互不关联。张尧庭、赵溱研究过多预报量多自变量回归模型，并在单预报量逐步回归的基础上，发展了多预报量多自变量双重筛选逐步回归[1]，但没有考虑到预报量之间的相互影响。多年来，线性回归模型在气象统计预报中，以及在与动力、数值、客观分析等方法结合的动力统计预报中，发挥了重要作用。人们为提高多元线性回归模型的预报水平，多致力于因子的筛选和资料的取舍工作，却没有注意到多元线性回归模型内在结构上的不合理性，例如：由理想气体状态方程 $P=\rho RT$，可知气压 P 和温度 T 之间存在一定的关系，如果我们用 P 和 T 作预报量，选定若干因子建立多元二重线性回归模型如下：

$$P=\beta_{10}+\beta_{11}X_{11}+\cdots+\beta_{1p}X_{1p}+u_1$$
$$T=\beta_{20}+\beta_{21}X_{21}+\cdots+\beta_{2q}X_{2q}+u_2$$

P 和 T 之间本来存在着的相互关系并没有被模型直接描述出来，当然，二者的相互作用可以由预报因子"内在地"加以传递体现，因为两个方程中的预报因子可以相同，即使不同也可能存在某种内在联系，但是，这种传递能否真正体现？这种内在联系是否真正存在？都不可得知。尤如一个"黑匣子"，只知其输入和输出两端，中间的过程是无法知晓的。以上模

* 初稿时间：1994 年 7 月 25 日；修改稿时间：1995 年 9 月 18 日。
资助课题："八五"国家科技攻关(85-906-05-02-06)项目"台风预报动力、统计释用新方法和新技术的研究"。
乔春晓同志参加了资料收集和整理工作。

型系统地忽略了预报量之间的相互作用，现将它改进为如下形式：

$$P = \gamma_{12}T + \beta_{10} + \beta_{11}x_{11} + \cdots + \beta_{1p}x_{1p} + u_1$$

$$T = \gamma_{21}P + \beta_{20} + \beta_{21}x_{21} + \cdots + \beta_{2q}x_{2q} + u_2$$

这样就可将 P 和 T 的关系明显地在模型中表现出来。这种模型的一般形式为：

$$\begin{cases} \gamma_{11}y_{1t} + \cdots + \gamma_{1g}y_{gt} = \beta_{10} + \beta_{11}x_{1t} + \cdots + \beta_{m1}x_{mt} + u_{1t} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \gamma_{1g}y_{1t} + \cdots + \gamma_{gg}y_{gt} = \beta_{g0} + \beta_{1g}x_{1t} + \cdots + \beta_{mg}x_{mt} + u_{gt} \end{cases} \quad t = 1,\cdots,T \tag{1}$$

式(1)称为系统方程式，其中包含 g 个结构方程。$t = 1,\cdots,T$ 表示有 T 次独立观测；用矩阵形式可写为：

$$Y\Gamma = XB + U \tag{2}$$

其中 Y：$T \times g$ 内生变量矩阵，Γ：$g \times g$ 内生变量系数矩阵，X：$T \times (m+1)$ 外生变量矩阵，B：$(m+1) \times g$ 外生变量系数矩阵，U：$T \times g$ 误差矩阵，称式(2)为“关联模型”。它首先由 Klein[2]根据经济现象提出(故原称“经济计量模型”)。在发现气象中的多元(自变量)多重(预报量)线性回归模型的结构缺陷后(主要是未考虑各因变量间的相互影响)，将该模型引入气象研究领域[1]。

所谓“关联”，就是在该模型中，不仅外生变量(即自变量或预报因子)作用于内生变量(即因变量或预报量)，而且内生变量(预报量)之间也存在着相互作用。一般多元线性回归模型系统忽略的预报量之间的相互作用，在“关联模型”中得到了明确的数学描述，保证了一些物理过程和相互作用的再现，预报量(内生变量)之间不再是象多元多重回归模型中那种互不相干的孤立关系。

由于外生变量个数人为确定，且数据间的相关程度不同，则出现能否用所给数据估计模型参数的问题。若对系统方程中的参数能从独立观测数据求得其唯一估计量，称该模型可识别，只有对可识别模型讨论其参数估计才有意义。因可对系统方程组(1)中的结构方程的次序作适当调整，假定

$$\Gamma = \begin{array}{c} \begin{pmatrix} \gamma_{(0)} & \Gamma_0 \\ 0 & \Gamma_1 \end{pmatrix} \begin{array}{l} g_1\text{ 行} \\ g - g_1\text{ 行} \end{array} \\ \begin{array}{cc} 1\text{ 列} & g-1\text{ 列} \end{array} \end{array} \qquad B = \begin{array}{c} \begin{pmatrix} \beta_{(0)} & B_0 \\ 0 & B_1 \end{pmatrix} \begin{array}{l} m_1 + 1\text{ 行} \\ m - m_1\text{ 行} \end{array} \\ \begin{array}{cc} 1\text{ 列} & g-1\text{ 列} \end{array} \end{array}$$

$$D = \begin{array}{c} \begin{pmatrix} \Gamma_1 \\ B_1 \end{pmatrix} \begin{array}{l} g - 1\text{ 行} \\ m - m_1\text{ 行} \end{array} \\ g - 1\text{ 列} \end{array} \qquad r = Rank(D)$$

则有以下可识别性准则：(1) $r = g - 1$ 且 $m - m_1 > g_1 - 2$：过分可识别，(2) $r = g - 1$ 且 $m - m_1 = g_1 - 2$：恰好可识别，(3) $r < g - 1$：不可识别。

因为关联模型中变量之间的相关性，不满足一般回归满足的 Gauss-Markov 定理的条件，故不能用一般回归所用的方法得到最优线性无偏估计，假定 $T > (g + m + 1)$，对系数矩阵 Γ 和 B 有许多估计方法，两段最小二乘估计(2SLS)，三段最小二乘估计(3SLS)，全部信息极大似然估计(FIML)和局部信息极大似然估计(LIML)等[4]。

现简要介绍模型参数的两段最小二乘估计(2SLS)基本思路，设系统方程组(1)可识别，且

(1)外生变量 $(x_1,\cdots,x_m)$ 与干扰变量 U 不相关；

(2) $E(u)=0$, $\mathrm{cov}(U,U)=\sum\otimes I_T$, $\sum$ 为 g 阶正定矩阵($\otimes$ 为 Kronecker 乘积)，系统方程组(1)中的每个结构方程中都有一个 γ 系数等于 1，设第一结构方程的 $\gamma_{11}=1$，故可写为

$$y_{(1)}=-Y_*\gamma_{(*)}+X_0\beta_{(0)}+u_{(1)} \tag{3}$$

其中

$$y_{(1)}=\begin{bmatrix}y_{11}\\ \vdots\\ y_{1T}\end{bmatrix},\qquad \gamma_{(1)}=\begin{bmatrix}1\\ \gamma_{(*)}\\ 0\end{bmatrix}\begin{matrix}1\text{ 行}\\ g_1-1\text{ 行},\\ g-g_1\text{ 行}\end{matrix}\qquad \beta_{(1)}=\begin{pmatrix}\beta_{(0)}\\ 0\end{pmatrix}\begin{matrix}m_1\text{ 行}\\ m-m_1\text{ 行}\end{matrix}$$

$Y_{(1)}=(y_{(1)}|Y_*)$—— 内生变量观测矩阵，X_0—— 外生变量矩阵。故 (Y_0,X_0) 是 (Y,X) 中除去第一结构方程中系数为零的那些变量的观测矩阵，现求 $\gamma_{(*)}$ 与 $\beta_{(0)}$ 的估计，由于 Y_* 与 $u_{(1)}$ 相关，不满足 Gauss-Markov 定理的条件，故不能得到 $\gamma_{(*)}$ 的最优线性无偏估计，因 X_0 与 $u_{(1)}$ 不相关，故可将 X_0 作为回归变量而对 Y_* 作第一段回归估计：$\hat{Y}_*=X_0(X_0'X_0)^{-1}X_0'Y_*$，有了 $\hat{Y}_*$，则可进行第二段估计，即再用 $\hat{Y}_*$ 对 $\gamma_{(*)}$ 与 $\beta_{(0)}$ 作出最终估计，故两段最小二乘估计法(2SLS)实质是两次最小二乘估计的复合，将式(3)改写为

$$\begin{aligned}y_{(1)}&=-\hat{Y}_*\gamma_{(*)}+X_0\beta_{(0)}+[u_{(1)}-(Y_{(*)}-\hat{Y}_*)\gamma_{(*)}]\\&=-\hat{Y}_*\gamma_{(*)}+X_0\beta_{(0)}+\tilde{u}_{(1)}\\&=\hat{Z}b+\tilde{u}_{(1)}\end{aligned}$$

得 b 的回归估计，即 $\gamma_{(*)}$ 与 $\beta_{(0)}$ 的 2SLS 估计：

$$\hat{b}=(\hat{Z}'\hat{Z})^{-1}\hat{Z}'y_{(1)} \tag{4}$$

其中

$$\hat{b}=\begin{bmatrix}-\gamma_{(*)}\\ \beta_{(0)}\end{bmatrix},\qquad Z=(-\hat{Y}_*\ X_0);$$

类似可得第 i 个结构方程中 $\gamma_{(*)}$ 与 $\beta_{(0)}$ 的 2SLS 估计：

$$\hat{b}=\begin{bmatrix}-\gamma_{(*)}\\ \beta_{(0)}\end{bmatrix}=(\hat{Z}'\hat{Z})^{-1}\hat{Z}'y_{(i)},\qquad i=2,\cdots,g \tag{4'}$$

式(4′)中的 $\gamma_{(*)}$，$\beta_{(0)}$ 与 $\hat{Z}$ 皆由各自结构方程($i=2,\cdots,g$)确定，最后求得 Γ 和 B 的 2SLS 估计。

一般不要求系统方程(2)中的系数矩阵 Γ 非奇异，但在实际应用中经常是非奇异的，若非奇异，则式(2)简化为模型的简约形式：

$$Y=XB\Gamma^{-1}+U\Gamma^{-1}=X\Pi+V \tag{5}$$

Zellner 提出了用于简约形式的第一段估计的 Zellner 法[4]，考虑式(5)的简约结构方程程

$$y_{(i)}=X_i\pi_{(i)}+v_{(i)}\qquad i=1,\cdots,g \tag{6}$$

其中 X_i 为 $T\times m_i$ 阵，是第 i 个简约结构方程中 T 次观测矩阵，$\sum_{i=1}^{g}m_i=m$，$\pi_{(i)}$ 为 m_i 维向量，将 $y_{(i)},\cdots,y_{(g)}$ 按列拉直为 Tg 维向量 y，参数和干扰向量也相应拉直：$\pi'=(\pi_{(1)}',\cdots,\pi_{(g)}')$；$v'=(v_{(1)}',\cdots,v_{(g)}')$；以 $X_1,\cdots,X_g$ 为主对角线建立 $Tg\times m$ 阶设计矩阵 W，有

$$y = W\pi + v,\ E(v) = 0,\ \mathrm{cov}(v,v) = \Omega = \sum \otimes I_T \tag{7}$$

即 v 的分量间相关，故 π 的广义二乘估计

$$\hat{\pi}^* = (W'\Omega^{-1}W)^{-1}W' \tag{8}$$

是 π 的最优线性无偏估计，比 π 的 LS 估计 $\hat{\pi} = (W'W)^{-1}W'$ 为优，其中

$$\hat{\pi}_{(i)} = (X_i'X_i)^{-1}X'_{Y_{(i)}},\qquad i = 1,\cdots,g \tag{9}$$

故可得 $y_{(i)}$ 的估计

$$\hat{y}_{(i)} = X_i(X_i'X_i)^{-1}X_i'Y_{(i)},\qquad i = 1,\cdots,g \tag{10}$$

若式(7)中 $\sum$ 未知，由式(6)，(9)和(10)知

$$\hat{v}_{(i)} = y_{(i)} - \hat{y}_{(i)} = y_{(i)} - X_i\hat{\pi}_{(i)} = [I_T - (X_i'X_i)^{-1}X_i']y_{(i)},\qquad i = 1,\cdots,g$$

而有 $\hat{\sum} = (\hat{\sigma}_{ij})_{g\times g}$，$E(\hat{\sum}) = \sum$，其中 $\hat{\sigma}_{ij} = \hat{v}'_{(i)}\hat{v}_{(i)}/(T-m_i)^{1/2}(T-m_j)^{1/2}$，$i,j = 1,\cdots,g$，从而有 $\hat{\Omega} = \hat{\sum} \otimes I_T$ 代入式(8)中得 $\hat{\pi}^* = (W'\hat{\Omega}^{-1}W)^{-1}W'$ 因而得 y 的估计 $\hat{y} = W\hat{\pi}^* = (\hat{y}'_{(1)},\cdots,\hat{y}'_{(g)})'$，由此作第二段估计，将 $(Y_{(2)},\cdots,Y_{(g)})$ 看作式(3)中的 Y_*，用式(4)作出 $\gamma_{(*)}$ 与 $\beta_{(0)}$ 的估计，求出 $\gamma_{(1)}$ 与 $\beta_{(1)}$ 的估计值；同理可得 $\gamma_{(i)}$ 与 $\beta_{(i)}$，$i = 2,\cdots,g$ 的估计，最终得到 Γ 和 B 的估计。

2 建模与资料

台风路径、强度和风速的多维动态关联预报模型如下：

$$\begin{cases} y_1 = \gamma_{12}y_2 + \gamma_{13}y_3 + \gamma_{14}y_4 + \beta_{10} + \sum\limits_{i=1}^{10}\beta_{1i}x_i + \sum\limits_{i=11}^{15}\beta_{1i}x_i \\ y_2 = \gamma_{21}y_1 + \gamma_{23}y_3 + \gamma_{24}y_4 + \beta_{20} + \sum\limits_{i=1}^{10}\beta_{2i}x_i + \sum\limits_{i=16}^{20}\beta_{2i}x_i \\ y_3 = \gamma_{31}y_1 + \gamma_{32}y_2 + \gamma_{34}y_4 + \beta_{30} + \sum\limits_{i=1}^{10}\beta_{3i}x_i + \sum\limits_{i=21}^{25}\beta_{3i}x_i \\ y_4 = \gamma_{41}y_1 + \gamma_{42}y_2 + \gamma_{43}y_3 + \beta_{40} + \sum\limits_{i=1}^{10}\beta_{4i}x_i + \sum\limits_{i=26}^{30}\beta_{4i}x_i \end{cases} \tag{1'}$$

故有

$$\Gamma = \left(\begin{array}{c|ccc} 1 & -\gamma_{21} & -\gamma_{31} & -\gamma_{41} \\ -\gamma_{12} & 1 & -\gamma_{32} & -\gamma_{42} \\ -\gamma_{13} & -\gamma_{23} & 1 & -\gamma_{43} \\ -\gamma_{14} & -\gamma_{24} & -\gamma_{34} & 1 \end{array}\right) = (\gamma_{(0)} | \Gamma_0)$$

$$B' = \left(\begin{array}{ccccccc|ccccccccc} \beta_{10} & \beta_{11} & \cdots & \beta_{1,10} & \beta_{1,11} & \cdots & \beta_{1,15} & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ - & - & - & - & - & - & - & - & - & - & - & - & - & - & - & - \\ \beta_{20} & \beta_{21} & \cdots & \beta_{2,10} & 0 & \cdots & 0 & \beta_{2,16} & \cdots & \beta_{2,20} & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \beta_{30} & \beta_{31} & \cdots & \beta_{3,10} & 0 & \cdots & 0 & 0 & \cdots & 0 & \beta_{3,21} & \cdots & \beta_{3,25} & 0 & \cdots & 0 \\ \beta_{40} & \beta_{41} & \cdots & \beta_{4,10} & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & \beta_{4,26} & \cdots & \beta_{4,30} \end{array}\right)$$

$$= \left(\begin{array}{c|c} \beta'_{(0)} & 0' \\ - & - \\ B_0' & B_1' \end{array}\right)$$

$D = B_1, r = \text{Rank}(D) = 3,\ g = 4, g_1 = 4, m = 30, m_1 = 15, r = g - 1 = 3$ 和 $m - m_1 = 15 > g_1 - 1 = 3$。故该模型(过分)可识别。必须 $T > g + m + 1 = 35$，取 $T = 36$。模型中变量为：

纬　　度：$y_1, x_{11}, x_{12}, x_{13}, x_{14}, x_{15}$；经　　度：$y_2, x_{16}, x_{17}, x_{18}, x_{19}, x_{20}$；

中心气压：$y_3, x_{21}, x_{22}, x_{23}, x_{24}, x_{25}$；最大风速：$y_4, x_{26}, x_{27}x_{28}, x_{29}, x_{30}$。

另据文献[5]，选用海口站为指标站，选 x_1, x_2, x_3, x_4, x_5：为海口站地面相对湿度；$x_6, x_7, x_8, x_9, x_{10}$：为海口站地面温度；每个结构方程有 19 个系数，用系数向量 $b_i(i=1,2,3,4)$ 表示。即：

$$
\begin{aligned}
b_1 &= (\gamma_{12}, \gamma_{13}, \gamma_{14}, \beta_{10}, \beta_{11}, \cdots, \beta_{1,10}, \beta_{1,11}, \cdots, \beta_{1,15})' \\
b_2 &= (\gamma_{21}, \gamma_{23}, \gamma_{24}, \beta_{20}, \beta_{21}, \cdots, \beta_{2,10}, \beta_{2,16}, \cdots, \beta_{2,20})' \\
b_3 &= (\gamma_{31}, \gamma_{32}, \gamma_{34}, \beta_{30}, \beta_{31}, \cdots, \beta_{3,10}, \beta_{3,21}, \cdots, \beta_{3,15})' \\
b_4 &= (\gamma_{41}, \gamma_{42}, \gamma_{43}, \beta_{40}, \beta_{41}, \cdots, \beta_{4,10}, \beta_{4,26}, \cdots, \beta_{4,30})'
\end{aligned} \qquad (1'')
$$

为了便于在气象中应用，称为："多维气象关联模型"，在该模型中，又引进变量之间存在着的时间序列关系，故正式命名为"多维气象动态关联模型"。本文就是用此模型对台风的路径、强度和风速同时作 24h，48h 和 72h 预报。从《台风年鉴》的原始资料中按照一定的时序要求采集时序资料，每采集一次作为一次观测样本：$y_i, x_{11+5(i-1)}, x_{12+5(i-1)}, x_{13+5(i-1)}, x_{14+5(i-1)}, x_{15+5(i-1)}$ 对每个 $i(=1,2,3,4)$，它是一个时间序列；如将 $x_{11+5(i-1)}$ 视作 t 时刻台风某要素 F 的值，则有：y_i：F 在 $(t+h)$ 时的值；$x_{11+5(i-1)}$：F 在 t 时的值；$x_{12+5(i-1)}$：F 在 $(t-6)$ 时的值；$x_{13+5(i-1)}$：F 在 $(t-12)$ 时的值；$x_{14+5(i-1)}$：F 在 $(t-18)$ 时的值；$x_{15+5(i-1)}$：F 在 $(t-24)$ 时的值；其中 h 是预报时效，h 分别为 24，48，72h。

又据文献[5]，所取指标站要素对应的时刻如下：x_1, x_6 对应时刻 t；x_2, x_7 对应时刻 $(t-6)$；x_3, x_8 对应时刻 $(t-12)$；x_4, x_9 对应时刻 $(t-18)$；x_5, x_{10} 对应时刻 $(t-24)$；收到台风个例的资料后，每定位一次 t，就取得各要素的一次观测样本，共定 $T(=36)$ 次；要保证每次"观测样本"的各时刻都要在同一台风个例内，但并不要求 T 次"观测样本"都在同一台风个例内。各变量的单位仍用常规单位(只将相对湿度去掉百分号，放大 100 倍)。

用 8303，8302，8304 和 8309 号台风资料及相应的海口地面温度、相对湿度采集时序资料(24h 建模用了 8303 和 8302 号台风资料；48h 建模用了 8304 号台风资料；72h 建模用了 8309 号台风资料)；用两段最小二乘法(2SLS)估计系数矩阵 Γ 和 B，但分为：

方法 1. 考虑内生变量协方差阵的算法，称为模型 cov$h$ (h =24，48，72)；

方法2. 不考虑内生变量协方差阵的算法，称为模型 wcov$h$ (without covariance)；

还用 $(y_i - x_{11+5(i-1)})$ 代替 y_i，即对预报量的 hh 差用方法1，建立 hh 的预报模型，称为模型 dcov$h$。

预报用的是8406号台风资料(1984年7月25日08时到1984年8月3日08时；由于用的是动态模型，充分利用了8406号台风个例的信息量，故24h 预报的样本容量29个，48h 预报的为25个，72h 预报的为21个)。

下面给出由2SLS 估计法计算出来的台风路径(纬度：y_1，经度：y_2)，气压(y_3)和风速(y_4)的24h，48h，72h 多维动态关联预报模型(1)中各结构方程的系数：(48h，72h 的略)

模型 cov24预报模型中的系数*

b_1:	0.0007306	6.138E−005	−0.0008886	−6.884	0.02239	0.04910	0.02583	0.03707	−0.01339
0.04289	0.05832	−0.003438	0.005826	−0.1176	1.454	−1.266	1.043	−0.5866	0.2415
b_2:	0.006149	−1.504E−006	0.002817	35.61	0.02271	−0.05335	−0.04848	−0.02728	−0.05082
−0.1326	−0.3852	−0.2484	−0.2116	−0.2136	1.742	−0.6908	−0.2396	0.4200	−0.1755
b_3:	0.003144	−0.009037	−0.04671	2055	−0.5287	−0.7518	−0.6273	−0.6765	0.1723
0.9546	−0.1535	0.8277	1.029	1.048	3.275	0.3410	−0.6968	0.1949	−1.839
b_4:	−0.001152	0.0002138	0.0006127	10.51	0.3429	0.4481	0.3623	0.5089	−0.06929
0.4063	−0.5386	−0.9248	−0.3864	−2.081	0.6356	−0.7954	−0.4949	−1.228	1.180

（*：模型 wcov24、wcov48、wcov72—不考虑协方差阵的预报模型略；

模型 dcov24、dcov48、dcov72--考虑协方差阵的预报量 hh 差的预报模型略）

3　随机模拟试验

任何观测资料都存在误差，资料的随机误差对模型参数的估计结果会产生影响，为考察这种影响，给出和资料误差量级等同的均匀分布的随机数加到原始资料的数据上建模进行回报，以研究模型的稳定性，随机数具体范围如下：经纬度：−0.1—+0.1(°)；气压：−2—+2(hPa)；风速：−3—+3(m/s)；对应24h，48h，和72h 的模型，各作了100次随机模拟试验进行回报，各得到100组4个预报量的36个回报值，再求得其与实况值的绝对偏差，然后就此偏差，对样本容量为36，4个变量的100组(母体)作了多元方差分析[6]，结果表明，与资料误差量级等同的均匀分布的随机误差对模型没有显著影响，模型是稳定的。

4　回报和预报结果分析

表1　模型 cov24的4要素回报结果统计分析表

Variable	Mean	Std Dev	Variance	Minimum	Maximum	Valid	Missing
$y1c$	0.9087	0.7204	0.5189	0.0227	2.4074	36	0
$y2c$	1.2315	1.2530	1.5699	0.0066	5.1125	36	0
$y3c$	11.9847	9.3006	86.5009	0.2049	34.3583	36	0
$y4c$	4.8775	3.8624	14.9185	0.2394	16.3176	36	0

上述数据中 $y1c$，$y2c$，$y3c$，$y4c$ 分别为实况值 y_1，y_2，y_3，y_4与估计值 $y1g$，$y2g$，$y3g$，$y4g$ 的绝对偏差，以下均相同。根据回报的统计分析，实况与回报值的绝对偏差平均提供了预报时的参考修正值，现将经适当修正后的预报结果及统计分析如下(见表2)：

表2　模型 COV24的4要素实况与预报结果及统计分析表

No	y_1	y_{1g}	y_{1c}	y_2	y_{2g}	y_{2c}	y_3	y_{3g}	y_{3c}	y_4	y_{4g}	y_{4c}
1	26.90	27.01	0.1081	140.8	135.6	5.155	982.0	975.0	7.043	25.00	29.19	4.188
2	27.70	26.57	1.134	140.4	136.1	4.302	978.0	986.9	8.872	25.00	28.08	3.077
……	……	……	……	……	……	……	……	……	……	……	……	……
28	37.40	35.05	2.352	118.8	116.1	2.668	1000.	949.8	50.24	15.00	51.24	36.24
29	38.50	35.17	3.329	119.5	115.8	3.718	1000.	950.9	49.07	15.00	44.73	29.73

Variable	Mean	Std Dev	Variance	Minimum	Maximum	Valid	Missing
y_{1c24}	0.7450	0.7265	0.5277	0.0216	3.3294	29	0
y_{2c24}	1.9417	1.2734	1.6215	0.0396	5.1549	29	0
y_{3c24}	17.7538	17.7267	314.2347	0.4200	50.2396	29	0
y_{4c24}	13.3781	10.7988	116.6146	0.4732	36.2429	29	0

模型 cov48,cov72的回报和预报结果统计分析详见表3。

5　结语和设想

(1)"气象关联模型"将多元线性回归模型中系统忽略的预报量(内生变量)间的相互作用明确地加以描述,使得过去认为在统计模型中忽略物理机制的问题得到解决。若根据一定的物理机制,某内生变量对另一内生变量不应有任何作用,则可令模型参数矩阵 Γ 的相应元素为零;若不能确定内生变量间是否有相互作用,也可从模型参数矩阵 Γ 的相应元素是否接近零,而了解它们间的相互作用的大小。如模型中气压对经纬度的影响非常小,可忽略;尽管相对来说,预报量之间的相互影响是比预报因子对预报量的影响要小,但是否直接用一般多元回归模型就比用"气象关联模型"更好呢?为此,分别用相同的预报因子,用与"关联"模型完全相同的样本资料,分别作了4个独立的关于各预报量的一般多元回归模型进行比较,现将"关联"模型与多元回归分析模型对台风4要素的3种时效的回报和预报的结果与实况值间的平均绝对误差列表如下:

表3　台风4要素平均绝对误差3种时效回报和预报的"关联"模型和回归模型的结果对比表

时效	24h				48h				72h			
	回报		预报		回报		预报		回报		预报	
模型	关联	回归	关联	回归	关联	回归	关联	回归	关联	回归	关联	回归
纬度	0.9087	0.2798	0.7450	3.6470	0.9202	0.7717	1.5148	1.7070	1.3014	0.7189	2.5127	4.8480
经度	1.2315	0.6988	1.9417	2.0130	0.9596	0.8929	2.6950	3.2960	1.7216	1.2910	2.9749	7.7510
气压	11.985	4.69901	7.7541	3.080	6.3539	6.22502	1.2452	6.8701	9.7921	5.4101	3.9833	4.010
风速	4.8775	3.22801	3.378	11.500	4.1532	4.100	15.814	23.790	9.0630	8.7860	6.6165	12.770

从表3可知,用多元回归方法虽然回报的效果略好些,但预报的效果却差得较多,即"气象关联模型"的预报效果比一般多元回归模型的预报效果要好得多,这可能正是预报量之间的相互作用改进了预报效果所致。

另外必须说明的是,在其他的气象问题中,不可避免的会有预报量之间的相互作用相当明显的情况,如文献[7]中,台风受灾面积(y_1)对伤亡人数(y_2)和倒损房屋(y_3)的影响就很大,但人员伤亡(y_2)对受灾面积(y_1)的作用小到可以忽略,对倒损房屋(y_3)的影响也较小,而倒损房屋(y_3)对人员伤亡的影响就比对受灾面积的影响大,这些系数的大小完全能反映实际情况,"气象关联模型"将会起到它更有效的作用.

(2)由于条件限制,没能在模型中引进更多更有意义的适当的物理量因子,这一点在今后的实际业务中必须注意考虑。

(3)实际业务使用时,建议根据需要,分区域或分类型分别建模,以进一步提高预报的准确率和精度。

(4)如何在模型本身中引进逐步筛选因子的理论和方法,也是一个有意义的课题。

参考文献

[1] 张尧庭. 方开泰. 多元统计分析引论. 北京:科学出版社. 1982. 288—297.

[2] Klein L R. A Textbook of Econometrics. NewYork. Prentice—Hall. 1974. 198—312.

[3] Schneeweiss H. Oekonometrie 3 Auflage. Wuerzburg Wien. Physica—Verlg. 1986. 242—331.
[4] Zellner A. Three-stage-least-squares: simultaneous estimation of simultaneous equations. Econometrica. 1962. 30: 54—78.
[5] 台风会议文集(1976. 1981. 1985). 上海科技出版社.
[6] 吕纯濂. 陈舜华. 矩阵语言与多元分析(专著). 北京: 气象出版社. 1994. 193—204.
[7] 吕纯濂. 陈舜华. 福建省台风灾害损失估算试验. 中国减灾. 1994. 4(3): 31—34.

RESEARCH AND APPLY MULTIPLE DYNAMIC INTER-DEPENDENT MODEL (MDIM) TO PREDICT TYPHOON TRACK, INTENSITY AND WIND-SPEED, SIMULTANEOUSLY

Lu Chunlian Chen Shunhua

(Nanjing Institute of Meteorology, Nanjing, 210044)

Zhu Yongti

(Shanghei Meteorological Bureau, Shanghei, 200030)

Abstract

The Multiple Dynamic Interdependent Model (MDIM) to predict the typhoon track, intensity and wind-speed in 24h, 48h and 72h, respectively were researched and applied. Based on the previous work, the algorithm with considering covariance of endogenous variables was improved . New term "Multple Meteorological Dynamic Interdependent Model" was formaly named. And three aspects of the work were given: 1. Backward test and prediction in 24h, 48h and 72h and its statistical analysis. 2. Random simulation with adding uniform distributed random number to observation and analysis of its stability. 3. According to the problem of the typhoon in the model, some questions to solve and constructive suggestions to use are raised.

Key words: Multiple dynamic interdependent model, Typhoon, Two-stage least two-square estimation.