第23卷 第3期
1999 年 5 月
大 气 科 学
Chinese Journal of Atmospheric Sciences
Vol. 23, No. 3
May 1999

一种新的统计回归模型及其建模方案

陶 云 曹 杰 严华生 谢应齐

（云南大学地球科学系，昆明 650091）

摘 要 根据经典门限自回归模型的基本思想，引入半截多项式变换，导出了一种新的统计回归模型，并提出了相应的一整套建模方案。这种模型和建模方案的特点是：（1）解决了统计回归分析中逐段线性化模型的检验问题；（2）在确定统计回归模型中各待估参数和变量——包括门限变量、门限值、阶数、时滞和回归系数时显得十分方便、快捷。试验结果表明，根据此方案建立的统计回归模型具有较高的拟合和预报精度，同时具有良好的稳定性。

关键词 统计回归模型 半截多项式 建模方案

1 数学模型

为了清楚起见，且不失一般性，取 2 个门限变量：x_{t-d_1}，x_{t-d_2}；门限值分别为：$x_{t-d_1,\min}<r_{1,1}<r_{1,2}<\cdots r_{1,n}\cdots<r_{1,L_1}<x_{t-d_1,\max}$；$x_{t-d_2,\min}<r_{2,1}<r_{2,2}<\cdots r_{2,m}\cdots<r_{2,L_2}<x_{t-d_2,\max}$；设最大时滞为 k，则经典门限自回归模型可写为

$$x_t=\begin{cases}
a_{0,0,0}+\sum_{i=1}^{k}a_{0,0,i}x_{t-i}+\varepsilon_{0,0,t} & x_{t-d_1,\min}\leqslant x_{t-d_1}<r_{1,1},\ x_{t-d_2,\min}\leqslant x_{t-d_2}<r_{2,1}\\
a_{0,1,0}+\sum_{i=1}^{k}a_{0,1,i}x_{t-i}+\varepsilon_{0,1,t} & x_{t-d_1,\min}\leqslant x_{t-d_1}<t_{1,1},\ r_{2,1}\leqslant x_{t-d_2}<r_{2,2}\\
a_{0,2,0}+\sum_{i=1}^{k}a_{0,2,i}x_{t-i}+\varepsilon_{0,2,t} & x_{t-d_1,\min}\leqslant x_{t-d_1}<r_{1,1},\ x_{2,2}\leqslant x_{t-d_2}<r_{2,3}\\
\vdots & \vdots \qquad \vdots\\
a_{0,m,0}+\sum_{i=1}^{k}a_{0,m,i}x_{t-i}+\varepsilon_{0,L_2,t} & x_{t-d_1,\min}\leqslant x_{t-d_1}<r_{1,1},\ x_{2,m}\leqslant x_{t-d_2}<r_{2,m+1}\\
\vdots & \vdots \qquad \vdots\\
a_{0,L_2,0}+\sum_{i=1}^{k}a_{0,L_2,i}x_{t-i}+\varepsilon_{0,L_2,t} & x_{t-d_1,\min}\leqslant x_{t-d_1}<r_{1,l},\ x_{2,L_2}\leqslant x_{t-d_2}<x_{t-d_2,\max}\\
a_{1,0,0}+\sum_{i=1}^{k}a_{1,0,i}x_{t-i}+\varepsilon_{1,0,t} & r_{1,1}\leqslant x_{t-d_1}<r_{1,2},\quad x_{t-d_2,\min}\leqslant x_{t-d_2}<r_{2,1}\\
a_{1,1,0}+\sum_{i=1}^{k}a_{1,1,i}x_{t-i}+\varepsilon_{1,1,t} & r_{1,1}\leqslant x_{t-d_1}<r_{1,2},\quad r_{2,1}\leqslant x_{t-d_2}<r_{2,2}\\
\vdots & \vdots \qquad \vdots\\
a_{n,m,0}+\sum_{i=1}^{k}a_{n,m,i}x_{t-i}+\varepsilon_{n,m,t} & r_{1,n}\leqslant x_{t-d_1}<r_{1,n+1},\quad r_{2,m}\leqslant x_{t-d_2}<r_{2,m+1}\\
\vdots & \vdots \qquad \vdots
\end{cases}\tag{1}$$

1997-12-29 收到，1998-10-05 收到再改稿

$$\begin{cases} a_{L_1,0,0}+\sum_{i=1}^{k}a_{L_1,0,i}x_{t-i}+\varepsilon_{L_1,0,t} & r_{1,L_1}\leqslant x_{t-d_1}<x_{t-d_1,\max},\quad x_{t-d_2,\min}\leqslant x_{t-d_2}<r_{2,1}\\ a_{L_1,1,0}+\sum_{i=1}^{k}a_{L_1,1,i}x_{t-i}+\varepsilon_{L_1,1,t} & r_{1,L_1}\leqslant x_{t-d_1}<x_{t-d_1,\max},\quad r_{2,1}\leqslant x_{t-d_2}\leqslant r_{2,2}\\ a_{L_1,2,0}+\sum_{i=1}^{k}a_{L_1,2,i}x_{t-i}+\varepsilon_{L_1,2,t} & r_{1,L_1}\leqslant x_{t-d_1}<x_{t-d_1,\max},\quad r_{2,2}\leqslant x_{t-d_2}<r_{2,3}\\ \vdots & \qquad\vdots\qquad\qquad\qquad\vdots\\ a_{L_1,L_2,0}+\sum_{i=1}^{k}a_{L_1,L_2,i}x_{t-i}+\varepsilon_{L_1,L_2,t} & r_{1,L_1}\leqslant x_{t-d_1}<x_{t-d_1,\max},\quad r_{2,L_2}\leqslant x_{t-d_2}<x_{t-d_2,\max}. \end{cases}$$

式中x_t是随机序列，且$\{\varepsilon_{n,m,t}\}$是$(L_1+1)(L_2+1)$个相互独立的正态白噪声，$n=0, 1, 2, \cdots, L_1$；$m=0, 1, 2, \cdots, L_2$。

在（1）式中引入半截多项式

$$u_+=\begin{cases}u, & u>0\\ 0, & u\leqslant 0\end{cases}\tag{2}$$

式中 u 为线性多项式。从半截多项式出发，得到两组分别以r_{1,j_1}和r_{2,j_2}为跳跃点的半截多项式：$(x_{t-d_1}-r_{1,j_1})_+$　和　$(x_{t-d_2}-r_{2,j_2})_+$，其中$j_1=1, 2, \cdots, L_1$；$j_2=1, 2, \cdots, L_2$．把它们加到一个一次多项式$\sum_{i=1}^{k}c_{0,i}x_{t-i}+c_0$中去，则在整个分段区间$r_{l,j_l}\in[x_{t-d_l,\min}, x_{t-d_l,\max}]$，$l=1, 2$可获得方程

$$s_t=\sum_{i=1}^{k}c_{0,i}x_{t-i}+\sum_{j_1=1}^{L_1}c_{1,j_1}(x_{t-d_1}-r_{1,j_1})_++\sum_{j_2=1}^{L_2}c_{2,j_2}(x_{t-d_2}-r_{2,j_2})_++c_0.\tag{3}$$

在每个子区间内，（3）式可写为如下线性多项式形式：

$$\begin{cases} S_{1,1,t}=\sum_{i=1}^{k}c_{0,i}x_{t-i}+c_0 & x_{t-d_1}<x_{t-d_1}\leqslant r_{1,1},\quad x_{t-d_2,\min}<x_{t-d_2}\leqslant r_{2,1}\\ S_{1,2,t}=S_{1,1,t}+c_{2,1}(x_{t-d_2}-r_{2,1})_+ & x_{t-d_1,\min}<x_{t-d_1}\leqslant r_{1,1},\quad x_{t-d_2}\leqslant r_{2,2}\\ S_{1,3,t}=S_{1,1,t}+c_{2,1}(x_{t-d_2}-r_{2,1})_++c_{2,2}(x_{t-d_2}-r_{2,2})_+ & \\ & x_{t-d_1,\min}<x_{t-d_1}\leqslant r_{1,1},\quad x_{t-d_2}\leqslant r_{2,3}\\ \vdots & \qquad\vdots\qquad\qquad\vdots\\ S_{1,(L_2+1),t}=S_{1,1,t}+\sum_{j_2=1}^{L_2}c_{2,j_2}(x_{t-d_2}-r_{2,j_2})_+ & \\ & x_{t-d_1,\min}<x_{t-d_1}\leqslant r_{1,1},\quad r_{2,L_2}<x_{t-d_2}<x_{t-d_2,\max}\\ S_{2,1,t}=S_{1,1,t}+c_{1,1}(x_{t-d_1}-r_{1,1})_+ & \\ & x_{t-d_1}\leqslant r_{1,2},\quad x_{t-d_2,\min}<x_{t-d_2}\leqslant r_{2,1}\\ S_{2,2,t}=S_{1,1,t}+c_{1,1}(x_{t-d_1}-r_{1,1})_++c_{2,1}(x_{t-d_2}-r_{2,1})_+ & \\ & x_{t-d_1,1}\leqslant r_{1,2},\quad x_{t-d_2}\leqslant r_{2,2}\\ \vdots & \qquad\vdots\qquad\qquad\vdots\\ S_{n,m,t}=S_{1,1,t}+\sum_{j_1=1}^{n-1}c_{1,j_1}(x_{t-d_1}-r_{1,j_1})_++\sum_{j_2=1}^{m-1}c_{2,j_2}(x_{t-d_2}-r_{2,j_2})_+ & \\ & x_{t-d_1}\leqslant r_{1,n},\quad x_{t-d_2}\leqslant r_{2,m}\\ \vdots & \qquad\vdots\qquad\qquad\vdots \end{cases}\tag{4}$$

$$
\begin{cases}
S_{(L_1+1),1,t}=S_{1,1,t}+\sum\limits_{j_1=1}^{L_1}c_{1,j_1}(x_{t-d_1}-r_{1,j_1})_+ \\
\qquad\qquad r_{1,L_1}<x_{t-d_1}<x_{t-d_1,\max},\quad x_{t-d_2,\min}<x_{t-d_2}\leqslant r_{2,1} \\
S_{(L_1+1),2,t}=S_{1,1,t}+\sum\limits_{j_1=1}^{L_1}c_{1,j_1}(x_{t-d_1}-r_{1,j_1})_+ + c_{2,1}(x_{t-d_2}-r_{2,1})_+ \\
\qquad\qquad r_{1,L_1}<x_{t-d_1}<x_{t-d_1,\max},\quad x_{t-d_2}\leqslant r_{2,2} \\
\qquad\vdots\qquad\qquad\vdots\qquad\qquad\vdots \\
S_{(L_1+1),(L_2+1),t}=S_{1,1,t}+\sum\limits_{j_1=1}^{L_1}c_{1,j_1}(x_{t-d_1}-r_{1,j_1})_+ + \sum\limits_{j_2=1}^{L_2}c_{2,j_2}(x_{t-d_2}-r_{2,j_2})_+ \\
\qquad\qquad r_{1,L_1}<x_{t-d_1}<x_{t-d_1,\max},\quad r_{2,L_2}<x_{t-d_2}<x_{t-d_2,\max}.
\end{cases}
$$

不难证明（4）式和（1）式具有等价性。

首先，在第一区间内等价；其次，在第 j 区间内对（4）式中 s_t 的半截多项式展开，合并同类项，与（1）式比较系数，即可证明其等价性。不失一般性，以 $L_1=1$、$L_2=1$ 为例，对（4）式中的半截多项式展开，得

$$c_{1,1}(x_{t-d_1}-r_{1,1})_+ + c_{2,1}(x_{t-d_2}-r_{2,1})=c_{1,1}x_{t-d_1}+c_{2,1}x_{t-d_2}-(c_{1,1}r_{1,1}+c_{2,1}r_{2,1}),$$

于是有

$$
\begin{aligned}
s_{2,2,t}&=c_0+\sum_{i=1}^{k}c_{0,i}x_{t-i}+c_{1,1}(x_{t-d_1}-r_{1,1})_+ + c_{2,1}(x_{t-d_2}-r_{2,1}) \\
&=c_0+\sum_{i=1}^{k}c_{0,i}x_{t-i}+c_{1,1}x_{t-d_1}+c_{2,1}x_{t-d_2}-(c_{1,1}r_{1,1}+c_{2,1}r_{2,1}).
\end{aligned}
$$

与（1）式比较系数，得

$$
\begin{aligned}
&a_{1,1,0}=c_0-(c_{1,1}r_{1,1}+c_{2,1}r_{2,1}), \\
&a_{1,1,i}=\begin{cases} c_{0,i} & i\neq d_1,\ \ i\neq d_2 \\ c_{0,i}+c_{1,1} & i=d_1,\ \ i\neq d_2 \\ c_{0,i}+c_{2,1} & i\neq d_1,\ \ i=d_2. \end{cases}
\end{aligned}
\tag{5}
$$

对于其他情形可仿此获得类似结论。可见，本文提出的新的统计回归模型（3）或(4)式与经典门限自回归模型的实质是完全一致的，只是形式上存在差异而已。

2　建模方案

(1) 给定回归模型时滞最大上限 k，样本长度 n，则生成基本的预报因子集 A 和预报对象集 B。

$$A=\begin{bmatrix} x_k & x_{k-1} & \cdots & x_{n-1} \\ x_{k-1} & x_k & \cdots & x_{n-2} \\ \vdots & \vdots & \vdots & \vdots \\ x_1 & x_2 & \cdots & x_{n-k} \end{bmatrix},\qquad B=[x_{k+1}\ \ x_{k+2}\ \ \cdots\ \ x_n]^T,$$

简记$A=[x_{n-1} \quad x_{n-2} \quad \cdots \quad x_{n-k}]^T$。

经过步骤（1），我们获得了预报对象和k个基本的预报因子，样本数为$n-k$。

（2）确定分点数。

对给定的预报因子x_i的$n-k$个样本数据，对x_i按从小到大的顺序排列，得到新x'_i，于是可得到样本区间所有可能的分点L，考虑到预报因子x_i中有重值的情况存在，有$L \leqslant n-k-1$。

（3）当分点确定后，按下式作替换，可得新的备选因子集，即令

$$\begin{cases} Z_1 = x_{t-1}, & Z_2 = x_{t-2}, & Z_3 = x_{t-3}, \\ \cdots & \cdots & \cdots \\ Z_k = x_{t-k}, & Z_{k+1} = (x_{t-1} - r_{1,1})_+, & Z_{k+2} = (x_{t-1} - r_{1,2})_+, \\ \cdots & \cdots & \cdots \\ Z_{k+L} = (x_{t-1} - r_{1,L})_+, & Z_{k+L+1} = (x_{t-2} - r_{2,1})_+, & Z_{k+L+2} = (x_{t-2} - r_{2,2})_+, \\ \cdots & \cdots & \cdots \\ Z_{k(L+1)} = (x_{t-k} - r_{k,L})_+, \end{cases} \tag{6}$$

则得因子集$Z=[Z_i]$，$i=1,2,\cdots,k(L+1)$，总的因子数为$k(L+1)$个，于是模型的形式可写为

$$B = CZ^T. \tag{7}$$

（4）运用逐步引入剔除紧凑变换技术对$[Z_i]$进行筛选，则可以合并某些不必要的分段。例如：若$(x_{t-i}-r_4)_+$这一因子在逐步回归中没有人选，则可以认为分点r_4可以去掉，即将第i个因子的第4个子区间$[r_3, r_4]$与第5个子区间$[r_4, r_5]$合并，这样即可从中自动挑选出最优门限变量和门限值，最终建立统计回归模型

$$\hat{B} = DZ'^T, \tag{8}$$

其中$Z' \subset Z$。

（5）当给出新的观察资料后，按上述步骤生成新的预报因子集，即可作外推预报。

3 应用实例

考虑到预报的时效，分别以昆明、蒙自、腾冲和沾益4站1951年5月～1990年4月的逐月雨量序列作为预报因子，并以该4站1952～1990年汛期雨量（6～8月）作为预报对象，按照前述建模步骤分别建立该4站汛期的统计回归模型如下:

$$\text{昆明: } \hat{x}_t = -6.17878(x_{t-11} - 314.1)_+ - 1.18362(x_{t-10} - 263.0)_+ \\ + 1.73658(x_{t-8} - 150.4)_+ - 9.73443(x_{t-4} - 23.2)_+ \\ + 6.11626(x_{t-3} - 35.7)_+ + 613.498, \tag{9}$$

$$\text{蒙自: } \hat{x}_t = 0.82105x_{t-6} + 2.21485(x_{t-5} - 44.7)_+ - 3.19972(x_{t-3} - 46.1)_+ \\ - 0.86031(x_{t-2} - 0.4)_+ + 433.853, \tag{10}$$

$$\text{腾冲: } \hat{x}_t = 1.74637(x_{t-7} - 67.9)_+ - 4.25085(x_{t-6} - 12.4)_+$$

$$+ 8.78522(x_{t-6} - 44.9)_+ + 790.314, \tag{11}$$

沾益：
$$\hat{x}_t = 1.39268(x_{t-12} - 225.8)_+ - 1.68375(x_{t-4} - 9.8)_+ + 3.68995(x_{t-2} - 62.6)_+ + 524.328, \tag{12}$$

其 F 检验值分别为：4.9、4.4、9.1、6.8，均通过了 $\alpha=0.01$ 的显著性检验。

为了进行比较，根据门限自回归模型的建模方案，分别建立该 4 站的经典门限自回归模型如下：

昆明：

$$\hat{x}_t = \begin{cases} 157.7424 + 0.1588x_{t-10} - 0.1122x_{t-9} - 0.2757x_{t-8} - 0.3337x_{t-7} \\ \quad + 0.2159x_{t-6} - 0.09102x_{t-5} - 0.32698x_{t-4} - 0.1598978x_{t-3} \\ \quad + 0.0566774x_{t-2} + 0.19643x_{t-1}, \qquad x_{t-6} \leqslant 45.4 \\ 21.213 + 0.3081x_{t-12} + 0.2382x_{t-11} + 0.03085x_{t-10} - 0.01499x_{t-9} \\ \quad - 0.07164x_{t-8} - 0.00144x_{t-7} + 0.00178x_{t-6} - 0.03786x_{t-5} - 0.0089x_{t-4} \\ \quad - 0.07904x_{t-3} + 0.06321x_{t-2} + 0.12758x_{t-1}, \qquad x_{t-6} > 45.4 \end{cases} \tag{13}$$

蒙自：

$$\hat{x}_t = \begin{cases} 86.691 + 0.25358x_{t-12} + 0.20488x_{t-11} - 0.03543x_{t-10} - 0.1003x_{t-9} \\ \quad - 0.0648x_{t-8} - 0.1553x_{t-7} - 0.09989x_{t-6} - 0.1339x_{t-5} - 0.2101x_{t-4} \\ \quad + 0.0383x_{t-3} + 0.00233x_{t-2} + 0.06373x_{t-1}, \qquad x_{t-6} \leqslant 66.2 \\ 18.65799 + 0.11438x_{t-12} + 0.1247x_{t-11} + 0.08863x_{t-10} - 0.0499x_{t-9} \\ \quad - 0.01664x_{t-8} - 0.00297x_{t-7} + 0.01406x_{t-6} - 0.04796x_{t-5} - 0.0764x_{t-4} \\ \quad + 0.01185x_{t-3} + 0.1246x_{t-2} + 0.1265x_{t-1}, \qquad x_{t-6} > 66.2 \end{cases} \tag{14}$$

腾冲：

$$\hat{x}_t = \begin{cases} 156.2783 + 0.3148x_{t-12} + 0.0662x_{t-11} - 0.01989x_{t-10} + 0.0134x_{t-9} \\ \quad - 0.04789x_{t-8} - 0.227x_{t-7} - 0.2439x_{t-6} - 0.1155x_{t-5} - 0.0514x_{t-4} \\ \quad - 0.1719x_{t-3} + 0.1032x_{t-2} + 0.055x_{t-1}, \qquad x_{t-6} \leqslant 182.2 \\ 50.52937 + 0.0588x_{t-10} + 0.03587x_{t-9} - 0.018977x_{t-8} \\ \quad - 0.03454x_{t-7} - 0.03471x_{t-6} - 0.03662x_{t-5} - 0.01852x_{t-4} \\ \quad - 0.02013x_{t-3} + 0.00599x_{t-2} + 0.083x_{t-1}, \qquad x_{t-6} > 182.2 \end{cases} \tag{15}$$

沾益：

$$\hat{x}_t = \begin{cases} 29.5447 + 0.1145x_{t-11} + 0.03949x_{t-10} - 0.00213x_{t-9} \\ \quad - 0.0326x_{t-8} - 0.0133x_{t-7} - 0.0588x_{t-6} - 0.0265x_{t-5} \\ \quad + 0.0268x_{t-4} + 0.0049x_{t-3} - 0.0274x_{t-2} + 0.1025x_{t-1}, \qquad x_{t-12} \leqslant 30.7 \\ 117.1107 + 0.1997x_{t-12} + 0.0206x_{t-11} + 0.0212x_{t-10} - 0.0138x_{t-9} \\ \quad - 0.02497x_{t-8} - 0.2663x_{t-7} - 0.1857x_{t-6} - 0.1257x_{t-5} \\ \quad - 0.1066x_{t-4} - 0.1897x_{t-3} + 0.0906x_{t-2} + 0.1711x_{t-1}. \qquad x_{t-12} > 30.7 \end{cases} \tag{16}$$

按文献[3]中的检验标准分别对（9）、（10）、（11）和（12）式及建立的经典门限自

回归模型（13）、（14）、（15）和（16）式进行历史拟合检验，其拟合准确率为：32／39、32／39、34／39和34／39、29／39、29／39、33／39和31／39。

分别对（9）、（10）、（11）、（12）和（13）、（14）、（15）、（16）式进行1991～1995年5年汛期雨量的外推预报检验，结果见表1和表2。

表1、表2中 x 为实测值，x'为外推预报值；"+"为正距平，"－"为负距平。显然，若 x 与 x'同号，则预报判定为正确。于是从表1、表2可看出，（9）、（10）、（11）和（12）式汛期雨量的外推预报准确率分别为：3／5、4／5、5／5和4／5；（13）、（14）、（15）和（16）式则分别为：3／5、3／5、3／5和4／5。4站平均拟合和外推预报准确率新模型为：84.6%和80.0%，两者相差在5%之内；而经典门限自回归模型为：78.2%和65%，两者差值为13.2%。可见，新模型及其建模方案具有较高的准确率和良好的稳定性。

表1 新模型外推预报检验结果

年		1991	1992	1993	1994	1995
昆明	x	－	－	－	+	+
	x'	－	－	－	－	－
蒙自	x	－	－	－	+	+
	x'	－	+	－	+	+
腾冲	x	+	－	－	－	+
	x'	+	－	－	－	+
沾益	x	+	－	－	+	－
	x'	+	－	－	+	+

表2 经典门限自回归模型外推预报检验结果

年		1991	1992	1993	1994	1995
昆明	x	－	－	－	+	+
	x'	－	－	－	－	－
蒙自	x	－	－	－	+	+
	x'	－	－	－	－	－
腾冲	x	+	－	－	－	+
	x'	－	－	－	－	－
沾益	x	+	－	－	+	－
	x'	－	－	－	+	－

5 总结与讨论

（1）本文在遵循门限自回归模型的基本原理基础上，引入半截多项式变换，导出了一种新的模型形式，并提出了相应的一整套建模方案。

（2）由 $x_t=\sum\limits_{i=1}^{k}c_{0,i}x_{t-i}+\sum\limits_{j_1=1}^{L_1}c_{1,j_1}(x_{t-d_1}-r_{j,j_1})_{+}+\sum\limits_{j_2=1}^{L_2}c_{2,j_2}(x_{t-d_2}-r_{2,j_2})_{+}+c_0$ 可清楚地看出，由于把等式右边包含门限变量和门限值的第二、第三项分别当作 L_1 和 L_2 个变量来处理，则对存在于上式中各个待估参数，包括门限变量、门限值、回归系数和模型阶数的确定和检验就可依托于经典的多元分析理论。其中尤其是对其运用逐步引入（或剔除）的求解逆紧凑方案作消去变换后，则可保证人选模型的变量均是对回归方程有显著贡献的变量，于是就解决了门限变量和门限值的统计显著性检验问题。当整个建模过程完成后，又可运用经典的对回归方程的 F 检验或复相关系数检验等对整个模型的统计特性作出检验。从而彻底解决了经典的分段线性化模型的统计检验问题。实例计算结果也表明，新模型及其建模方案与经典的门限自回归及其建模方案相比较具有一定的优越性。

(3) 由$x_t=\sum_{i=1}^{k}c_{0,i}x_{t-i}+\sum_{j_1=1}^{L_1}c_{1,j_1}(x_{t-d_1}-r_{1,j_1})_++\sum_{j_2=1}^{L_2}c_{2,j_2}(x_{t-d_2}-r_{2,j_2})_++c_0$可以看出，当所有包含分点的变量均入选回归方程时，则其蜕变为$x_t=\sum_{i=1}^{k}c_{0,i}x_{t-i}+c_0$。此式即为一般的自回归形式。

(4) 这种基于半截多项式的新建模方案，使门限变量和门限值的确定转化为求解一般自回归方程中回归系数问题，这就在处理多个门限变量和多个门限值等方面显得十分方便、快捷。

参 考 文 献

1 Tong, H. and Lim, K. S., 1980, Threshold autoregression, limit cycles and cyclical data, *J. Roy. Statist. Soc.*, 42(B), 245～292.

2 项静恬等，1991，动态和静态数据处理——时间序列和数据统计分析，北京：气象出版社.

3 尤卫红等，1995，月雨量的相空间相似统计预报试验，高原气象，14(2), 237～242.

A New Statistical Regression Model and Its Building Method

Tao Yun,　Cao Jie,　Yan Huasheng　and　Xie Yingqi

(Department of Earth Science, Yunnan University, Kunming 650091)

Abstract　On the basis of principle of the classical threshold auto-regression model, we advanced a new statistical regression model and a relative building method via introducing the semi-polynomial transformation. Its characteristics includes: First, solving the test problem of stagewise linear model in the statistical regression analysis. Second, determing the unestimated parameters and variables conveniently and promptly. The results show that the new statistical regression model is more efficient and stable than the classical one.

Key words　statistical regression model　semi-polynomial　method build-up