第15卷 第1期 大气科学 Vol.15, No.1
1991年 1月 SCIENTIA ATMOSPHERICA SINICA Jan., 1991

多层递阶预报模型的一种改进方案

李邦宪
(浙江省金华市气象局)

提要

本文针对气象要素与其过去的数据自相关程度一般都较差，而与其自身的显著周期分量却存在较高的线性相关关系的特点，提出了一种改进的多层递阶预报模型.它将多层递阶方法与逐步回归双重分析相结合，用显著周期分量取代经典多层递阶预报模型中的自回归部分，使之能更好地反映气象要素自身的历史演变规律，从而稳定其预报效果.

关键词：天气预报；多层递阶方法；显著周期分量；逐步回归双重分析.

一、问题的提出

近年来，多层递阶方法在天气预报领域中的应用日益广泛，发挥着越来越重要的作用.然而，从目前所做的工作来看，天气预报中的多层递阶方法大多是采用了仅有预报因子的线性单输出系统模型[1]，其主要原因在于气象要素的自相关程度往往比较低，使得仅有预报因子的线性单输出模型反而要比含有自回归部分的一般线性单输出模型或仅有自回归部分的线性单输出模型预报精度高[2].但问题在于，当某些气象要素序列含有较强的周期性变化规律时，若仍然采用仅有预报因子的线性单输出多层递阶预报模型，就不可避免地会导致预报误差增大.我们在工作中发现，虽然气象要素的自相关性较差，然而当序列具有明显的周期性变化时，它与其自身的显著周期分量有着很好的相关性，其线性相关系数往往可以达到极显著水平[3].根据这一事实，作者曾将多层递阶方法用于周期分析，取得了一定的效果[4]，但该方法无法反映前期物理因素对气象要素变化的影响，因而实质上仍是时间序列分析方法.为此，本文在上述工作的基础上，提出了一种改进的多层递阶预报模型，它将多层递阶方法与逐步回归双重分析相结合，在经典多层递阶预报模型中加进了特殊的因子——预报对象的显著周期序列，从而有利于稳定多层递阶方法的预报效果，具有较高的实用价值.

二、基本数学模型

在某些情况下，气象要素与其过去的数据自相关性较差，而与自身的显著周期分量存

1988年11月16日收到，1989年4月1日收到修改稿.

在较高的相关关系.此时,如果我们能用某种方法提取气象要素序列的显著周期分量,并用这些显著周期分量去取代经典多层递阶模型中的自回归部分,就有可能进一步提高其预报精度.

已知在经典的多层递阶方法中,其基本数学模型为:

$$y(t)=\sum_{i=1}^{l}\alpha_i(t)y(t-i)+\sum_{i=1}^{k}\beta_i(t)X_i(t)+\varepsilon(t), \tag{1}$$

其中 $y(t)$ 为预报对象,$X_i(t)$ 为预报因子,$\varepsilon(t)$ 为白噪声,$\alpha_i(t)$ 及 $\beta_i(t)$ 为时变参数,l 为自回归阶数,k 为预报因子个数,t 为流动时间.

设 $y(t)$ 具有 h 个显著周期,其相应的显著周期分量记为 $u_1(t),u_2(t),\cdots\cdots u_h(t)$,若用显著周期分量取代(1)式中的自回归项,则(1)式改写为

$$y(t)=\sum_{i=1}^{h}\alpha_i(t)u_i(t)+\sum_{i=1}^{k}\beta_i(t)X_i(t)+\varepsilon(t), \tag{2}$$

该式就是改进的多层递阶预报模型.在实际工作中,可以根据需要只取周期项或预报因子项.为处理方便,不妨将 $\varepsilon(t)$ 作为预报误差处理,即 $\varepsilon(t)=y(t)-\hat{y}(t)$.

不难看出,在上述模型中,关键是要解决两个问题:一是气象要素序列显著周期分量 $u_i(t)$ 的提取,二是时变参数的确定和预报.气象上常用的周期分析方法有方差分析、谱分析、最大熵谱分析等,但是,在一般情况下,物理因子和周期性往往同时支配着气象要素的变化.因此,我们认为,采用逐步回归双重分析方法[5]同时进行周期分析和筛选因子将更为适宜.而时变参数的确定和预报在经典多层递阶方法中已有充分的阐述.

三、计 算 步 骤

(1) 计算试验周期序列,按方差分析求试验周期的方法,将时间序列 $y(t)$ 依次按长度 $L(2<L<m)$ 进行分组:

$$\left.\begin{aligned}&y(1),\ \cdots\cdots,\ y(i),\ \cdots\cdots,\ y(L),\\&y(L+1),\ \cdots\cdots,\ y(L+i),\ \cdots\cdots,\ y(2L),\\&\cdots\cdots\\&y[(n_0-1)L+1],\ \cdots\cdots,\ y[(n_0-1)L+i],\ \cdots\cdots,\ y(n),\end{aligned}\right\} \tag{3}$$

其中 n 为原序列样本长度,n_0 为满足 $(n_0-1)L+i<n$ 的最大整数,m 为不大于 $n/2$ 的最大整数.

(2) 对以上各组分别求平均,即

$$u_j(i)=\frac{1}{n_0}\sum_{k=1}^{n_0}y\,[(k-1)L+i], \tag{4}$$

$$(i=1,2,\cdots\cdots,m-1;j=2,3,\cdots\cdots m)$$

则可得到 $m-1$ 个样本长度为 L 的平均值序列，称为周期长度为 L 的试验周期序列.将各试验周期序列按其周期性外延，使该序列的样本长度与 $y(t)$ 相等，并将各外延后的试验周期序列视为预报因子 $X_1, X_2, \cdots\cdots, X_{m-1}$.

(3) 在给定信度 α 下，对大量的候选因子进行相关普查，即计算

$$R=\left[\sum_{i=1}^{n}(y_i-\bar{y})(X_i-\bar{X})\right]\Big/\sqrt{\sum_{i=1}^{n}(y_i-\bar{y})^2\cdot\sum_{i=1}^{n}(X_i-\bar{X})^2}, \tag{5}$$

选取 $|R|>R_\alpha$ 的 k 个高相关因子，记为 $X_m, X_{m+1}, \cdots\cdots, X_{m+k-1}$，并将 $y(t)$ 视为 X_{m+k}.若预报因子已经过人工筛选，则此步工作可省略.

(4) 按照逐步回归的计算步骤，在给定的 F 显著性检验水平下，对 X_1、X_2、$\cdots\cdots$、X_{m+k-1} 进行变量的逐个引入和剔除，直至既无变量可引入，又无变量可剔除为止，记下被选变量的序号 i.显然，当 $i<m$ 时，则该变量就是预报对象序列 $y(t)$ 所含的显著周期分量；当 $i>m$ 时，则该变量就是前期预报因子[6].显然，这里我们把提取周期和因子筛选结合在一起，实质上是寻找前期因子和周期分量的最佳组合，以达到较好的预报效果.

(5) 根据多层递阶方法确定和预报时变参数，若置

$$\varphi^T(t)=[u_1(t),\ u_2(t),\ \cdots\cdots,\ u_h(t),\ X_1(t),\ X_2(t),\ \cdots\cdots X_k(t)],$$

$$\theta^T(t)=[\alpha_1(t),\ \alpha_2(t),\ \cdots\cdots,\ \alpha_h(t),\ \beta_1(t),\ \beta_2(t),\ \cdots\cdots\beta_k(t)],$$

则(2)式可简写为

$$y(t)=\varphi^T(t)\theta(t)+\varepsilon(t). \tag{6}$$

引进时变参数递推算法公式：

$$\hat{\theta}(t)=\hat{\theta}(t-1)+\frac{1}{\|\varphi(t)\|^2}\cdot\varphi(t)\cdot[y(t)-\varphi^T(t)\hat{\theta}(t-1)], \tag{7}$$

其中 $\hat{\theta}(t-1)$ 为 $\theta(t-1)$ 的估值，$\|\varphi(t)\|^2=\varphi^T(t)\cdot\varphi(t)$.

由历史资料，利用(7)式则可得到一系列时变参数跟踪估值序列.对上述参数估值序列进行分析，根据其不同的变化特点，通过适当的数学手段，建立其所满足的参数预报模型，常用的有均值近似法、多层 AR 模型递阶法、分段周期变量法等，作出时变参数预报值.然后运用 $y(t)$ 的向前一步预报公式：

$$\hat{y}(n+1/n)=\varphi^T(n+1)\cdot\hat{\theta}_i^*(n+1), \tag{8}$$

即可作出气象要素的向前一步预报[7].

四、实 例 计 算

本文以汛期降水量预报为例，每年五月至七月上旬为浙江省金华地区的梅汛期，其降水量年际变化很大，最多年与最少年之比达 4 倍之多，因而预报难度较大.选取北太平洋

37个格点的月平均海温和高空环流特征量作为候选因子，经相关普查，选出线性相关系数$|R_i|>0.40$的30个因子.取$F=3.0$作为显著性检验水平，经逐步回归双重分析计算，分别选入3个预报因子和11年显著周期分量，其中3个预报因子分别是：X_1为上年3月下旬副热带高压面积指数，X_2为上年8月中旬20°N，120°E与35°N，135°E两点的500hPa高度差，X_3为上年7月30°N，135°E处月平均海温.这样，可建立金华汛期降水量预报模型为

$$\hat{y}(t)=\alpha_{11}(t)u_{11}(t)+\beta_1(t)X_1(t)+\beta_2(t)X_2(t)+\beta_3(t)X_3(t) \tag{9}$$

取初始样本27个，时变参数初值为零，应用时变参数递推算法公式(7)，得到参数跟踪估值序列$\{\hat{\theta}(t)\}$.根据$\{\hat{\theta}(t)\}$的变化特点，我们采用多层AR递阶算法求得1984至1987年的时变参数预报值（见表1），然后应用(9)式则可求得状态预报值，其预报结果列在表2.

表1 时变参数预报值

年 份	α_{11}	β_1	β_2	β_3	年 份	α_{11}	β_1	β_2	β_3
1984	0.8083	0.0034	−0.0093	0.0502	1986	0.9273	0.0068	−0.0085	0.0575
1985	0.9672	0.0062	−0.0051	0.0613	1987	1.1034	0.0057	−0.0104	0.0626

表2 三种方法的预报效果比较

年份	实测值	改进的多层递阶预报模型			经典多层递阶预报模型			逐步回归双重分析		
		预报值	绝对误差	相对误差(%)	预报值	绝对误差	相对误差	预报值	绝对误差	相对误差
1984	569	612	43	7.6	437	175	28.6	692	123	21.6
1985	301	436	135	44.9	573	272	90.4	335	34	11.3
1986	385	415	30	7.8	563	178	46.2	640	255	66.2
1987	543	584	41	7.6	456	87	16.0	673	130	23.9
平均误差		—	62.3	17.0	—	178.0	45.3	—	135.5	30.8

取同样的资料，分别用仅有预报因子的经典多层递阶方法和逐步回归双重分析方法建立预报模型，其降水量预报值也列在表2.比较三种方法的预报值，可以看出改进后的多层递阶预报模型其精度有明显提高.

五、结 语

(1) 改进后的多层递阶模型是用显著周期分量取代原模型中的自回归部分，使之既反映了物理因子的支配作用，又体现了本身的周期性变化规律，因而有利于减少预报误差.

(2) 本文采用逐步回归双重分析方法提取显著周期，它把提取周期分量和筛选因子

进行有机的结合，不是单纯地提取周期，而是寻找预报效果较好的周期分量，这对于保证总的预报效果是有益的.

(3) 在多层递阶模型中，时变参数与因子值成正比，因子量值越大，时变参数也就越大.若因子间量级差异较大时，量值小的因子贡献就很小，此时，最好能对预报因子作预处理，使其处在同一量级水平上.对此我们将作进一步的研究.另外，选取合适的时变参数预报模型，对预报效果影响很大，应予足够重视.

(4) 从本文计算结果看，改进的多层递阶模型其预报精度有明显提高.但由于所计算的个例较少，还不足以证明其可靠性，本文的目的仅在于提供一种可供继续试验的预报方案，以期起到抛砖引玉的作用.

参　考　文　献

[1] 姚嘉玲，吴钟浚，1986，用多层递阶方法作中期气温预报，浙江气象科技，7，NO.4，20—23.
[2] 蔡惠芳，1986，试用多层递阶预报方法做夏季降水量预报，气象科技，NO.1，26—30.
[3] 李邦宪，1986，试用逐步回归周期分析作月降水预报，浙江气象科技，7，NO.4，11—12.
[4] 李邦宪，1988，多层递阶周期分析，气象，14，NO.11，44—46.
[5] 李邦宪，逐步回归双重分析及其在长期天气预报中的应用，大气科学(待发表).
[6] 李邦宪，1988，因子筛选与周期分析相结合的逐步回归双重分析预报模型，气象，14，NO.6，41—43.
[7] 韩志刚等，1985，黑龙江省冬季平均气温的多层递阶长期预报模型，大气科学，9，NO.2，171—177.

IMPROVED SCHEME OF MULTISTAGE RECURSIVE ESTIMATE FORECASTING MODEL

Li Bangxian

(Meteorological observatory of Jinhua City, Zhejiang Province)

Abstract

In this paper, a improved forecasting scheme is presented. In this new forecasting model, the obvious periodic components of time series already replaced the autoregression component in classical recursive estimate model. This model not only eonsiders the important action of various predictors, but also fairly reflects its periodical variable rules of prediction; therefore, its forecast effect is more stable.