第15卷 第2期
1991年 3 月

大　气　科　学
SCIENTIA ATMOSPHERICA SINICA

Vol.15, No.2
Mar., 1991

时间序列预报方法的推广

周家斌

（中国科学院大气物理研究所）

提　　要

文中讨论了车贝雪夫多项式值计算误差对作者以前提出的一种新的时间序列预报方法的预报结果的影响，并在此基础上对原来的计算方案作了推广，推广后的算法有利于提高预报准确率.

关键词: *时间序列; 预报方法; 车贝雪夫多项式; 非正交函数族.*

一、引　　言

作者在文献[1]、[2]中提出了一种新的时间序列预报方法. 几年来，我们用这一方法制作了北京4—8月逐月降水量预报和新安江流域年降水趋势(以新安江流域10站1—9月降水总量的加权平均值表示)预报，并提供北京市气象台和新安江水电厂使用. 表1给出预报检验结果. 由表可见，就降水变化趋势而言，大多数预报是准确的.

表1　时间序列方法预报效果检验

	新安江		北京									
			4月		5月		6月		7月		8月	
	Y	*S*	*Y*	*S*	*Y*	*S*	*Y*	*S*	*Y*	*S*	*Y*	*S*
气候平均		1486		18		32		76		228		167
1982			−2	20	−7	−19	−33	64	−89	−17	−108	−43
1983	388	541	17	61	−2	1	34	−9	19	−201	22	58
1984	−41	−48	22	18	18	−24	24	−32	−28	−168	53	111
1985	−127	−232	2	−4	8	−8	14	−44	−88	61	−17	131
1986	271	−303	12	−17	8	−27	4	127	−103	−65	63	−24
1987	140	93	10	24	7	33	−19	15	−96	−97	91	80
1988	−49	−115	8	1	11	1	−9	2	−63	−35	91	45

气候平均栏为平均降水量，其余为距平值. 单位: mm. *Y*为预测值，*S*为实况.

1989年3月22日收到，5月23日收到修改稿.

上述预报受到使用单位的欢迎，并曾就此做过鉴定．新安江电厂粟运华同志撰文介绍过我们提供预报的效果及他们应用这一方法做预报的情况[1]．目前，一些气象水文部门正在用这一方法制作业务预报[2]．

现在的问题是，如何进一步提高这一方法的预报准确率．我们在本文中，将给出解决这个问题的一个途径．为此，需要将预报方法进行推广．在第二节，首先简单介绍一下原先提出的时间序列预报方法的基本原理，第三节从分析车贝雪夫多项式值的误差对预报结果的影响入手讨论预报方法的推广问题，第四节给出计算实例．

二、时间序列预报方法简介

作者提出的时间序列预报方法的基本预报公式是[1,2]

$$Z^{(v)}(T_0+1)=\sum_{k=0}^{K_0} A_k^{(v)}\varphi_k(T_0+1),\quad (K_0<T_0), \tag{1}$$

$$A_k^{(v)}=\sum_{t=1}^{T_0} Z(t)\varphi_k(t)+Z^{(v-1)}(T_0+1)\varphi_k(T_0+1). \tag{2}$$

这是一组迭代公式，用于由 $t=1,2,\cdots,T_0$ 时刻的资料 $Z(1)$、$Z(2)$，$\cdots$，$Z(T_0)$ 计算 T_0+1 时刻 $Z(T_0+1)$ 的预报值 $Z^{(v)}(T_0+1)$．式中 v 为迭代次数，$\varphi_k(t)$ 为 k 阶规一化车贝雪夫多项式在格点 t 的值，K_0 为(1)式求和截止阶数，$A_k^{(v)}$ 为车贝雪夫系数 A_k 的第 v 次迭代结果．给定初值 $Z^{(0)}(T_0+1)$ 后，可以由(1)、(2)两式循环求得 $Z(T_0+1)$ 的各次近似值．

易证[2]

$$\varepsilon^{(v)}=\sigma^v\varepsilon^{(0)}+\frac{1-\sigma^v}{1-\sigma}\hat{\varepsilon}, \tag{3}$$

其中

$$\varepsilon^{(v)}=Z^{(v)}(T_0+1)-Z(T_0+1), \tag{4}$$

$$\varepsilon^{(0)}=Z^{(0)}(T_0+1)-Z(T_0+1), \tag{5}$$

$$\hat{\varepsilon}=\hat{Z}(T_0+1)-Z(T_0+1), \tag{6}$$

$$\hat{Z}(T_0+1)=\sum_{k=0}^{K_0} A_k\varphi_k(T_0+1), \tag{7}$$

$$A_k=\sum_{t=1}^{T_0+1} Z(t)\varphi_k(t), \tag{8}$$

$$\sigma=\sum_{k=0}^{K_0}\varphi_k^2(T_0+1). \tag{9}$$

以上各式中，$\varepsilon^{(v)}$ 为第 v 次迭代结果的预报误差，$\varepsilon^{(0)}$ 为初值误差，$\hat{\varepsilon}$ 为拟合误差，亦即 $Z(T_0+1)$ 已知时，(7)式的拟合误差，A_k 为车贝雪夫系数．

还可以证明[2]

$$0<\sigma<1. \tag{10}$$

由(10)式和(3)式可以推出，迭代计算是收敛的，其迭代终值为[2]

$$Z^{(\infty)}(T_0+1)=Z(T_0+1)+\frac{\hat{\varepsilon}}{1-\sigma}. \tag{11}$$

预报用两种方法算出：

(1) 应用理想初值，经少数几次迭代，由(1)，(2)两式算出预报误差较小的 $Z^{(\nu)}(T_0+1)$.

理想初值的条件是

(a)初值误差 $\varepsilon^{(0)}$ 同拟合误差 $\hat{\varepsilon}$ 反号；

(b) $|\varepsilon^{(0)}|>\dfrac{|\hat{\varepsilon}|}{\sigma}$.

(2) 利用以下两式由迭代终值计算预报

(a) $\hat{\varepsilon}>0$ 时，

$$Z_\infty(T_0+1)=Z^{(\infty)}(T_0+1)-\frac{\bar{\hat{\varepsilon}}_+}{1-\sigma}. \tag{12}$$

(b) $\hat{\varepsilon}<0$ 时，

$$Z_\infty(T_0+1)=Z^{(\infty)}(T_0+1)-\frac{\bar{\hat{\varepsilon}}_-}{1-\sigma}. \tag{13}$$

式中 $Z_\infty(T_0+1)$ 为预报值，$\bar{\hat{\varepsilon}}_+$ 和 $\bar{\hat{\varepsilon}}_-$ 分别为 $\hat{\varepsilon}>0$ 和 $\hat{\varepsilon}<0$ 时拟合误差 $\hat{\varepsilon}$ 的平均值.

具体做预报时，尚需制作一些预报工具图.

三、时间序列预报方法的推广

我们首先讨论一下车贝雪夫多项式值的误差对预报结果的影响，并据此对原有的预报方法加以推广.

假设由于某种原因造成了 $\varphi_k(t)$ 的误差.

设 $\varphi_{k_0}(t)$ 有错，令

$$\varphi'_{k_0}(t)=\varphi_{k_0}(t)+\delta_\varphi(t), \tag{14}$$

其中 $\varphi_{k_0}(t)$ 为真值，$\varphi'_{k_0}(t)$ 为有误差的值，$\delta_\varphi(t)$ 为误差，此时与(8)式相应的公式为

$$A'_{k_0}=\sum_{t=1}^{T_0+1}Z(t)\varphi'_{k_0}(t)=\sum_{t=1}^{T_0+1}Z(t)\varphi_{k_0}(t)+\sum_{t=1}^{T_0+1}Z(t)\delta_\varphi(t)=A_{k_0}+\delta_A. \tag{15}$$

同理可得(参见(2)式)

$$\begin{aligned}A^{(\nu)'}_{k_0}&=\sum_{t=1}^{T_0}Z(t)\varphi'_{k_0}(t)+Z^{(\nu-1)}(T_0+1)\varphi'_{k_0}(T_0+1)\\&=\sum_{t=1}^{T_0}Z(t)\varphi_{k_0}(t)+Z^{(\nu-1)}(T_0+1)\varphi_{k_0}(T_0+1)\\&\quad+\sum_{t=1}^{T_0}Z(t)\delta_\varphi(t)+Z^{(\nu-1)}(T_0+1)\delta_\varphi(T_0+1)\\&=A^{(\nu)}_{k_0}+\delta_A^{(\nu)}.\end{aligned} \tag{16}$$

式中 A_{k_0}'，$A_{k_0}^{(v)\prime}$ 为由 $\varphi_{k_0}'(t)$ 求得的车贝雪夫系数值，A_{k_0}，$A_{k_0}^{(v)}$ 为系数的准确值，δ_A 和 $\delta_A^{(v)}$ 为相应的误差.

将 A_{k_0}' 和 $A_{k_0}^{(v)\prime}(k=k_0$时)和其它准确的 A_k，$A_k^{(v)}(k\neq k_0)$一并分别代人(1)式和(7)式得到

$$Z^{(v)\prime}(T_0+1)=\sum_{k=0}^{K_0}A_k^{(v)\prime}\varphi_k'(T_0+1), \tag{17}$$

$$\hat{Z}'(T_0+1)=\sum_{k=0}^{K_0}A_k'\varphi_k'(T_0+1). \tag{18}$$

其中

$$A_k^{(v)\prime}=\begin{cases}A_{k_0}^{(v)\prime} & k=k_0,\\ A_k^{(v)} & k\neq k_0,\end{cases} \tag{19}$$

$$A_k'=\begin{cases}A_{k_0}' & k=k_0,\\ A_k & k\neq k_0,\end{cases} \tag{20}$$

$$\varphi_k'(t)=\begin{cases}\varphi_{k_0}'(t) & k=k_0,\\ \varphi_k(t) & k\neq k_0,\end{cases} \tag{21}$$

(17)—(20)式中上标“′”表示拟合值、迭代结果值、系数值，均系由有误差的多项式值 $\varphi_k'(t)$计算出来的.

应用与文献[1,2]同样的推导方法，可得

$$\varepsilon^{(v)\prime}=\sigma'^{v}\varepsilon^{(0)}+\frac{1-\sigma'^{v}}{1-\sigma'}\hat{\varepsilon}', \tag{22}$$

其中

$$\varepsilon^{(v)\prime}=Z^{(v)\prime}(T_0+1)-Z(T_0+1), \tag{23}$$

$$\hat{\varepsilon}'=\hat{Z}'(T_0+1)-Z(T_0+1), \tag{24}$$

$$\sigma'=\sum_{k=0}^{K_0}\varphi_k'^2(T_0+1). \tag{25}$$

若

$$0<\sigma'<1, \tag{26}$$

则可由(22)式推出

$$Z^{(\infty)\prime}(T_0+1)=Z(T_0+1)+\frac{\hat{\varepsilon}'}{1-\sigma'}. \tag{27}$$

亦即在满足(26)式的条件下，迭代计算式(17)仍是收敛的. 但需注意，在原来的算法[1,2]中，(10)式是必然满足的，而此处(26)式则需要作为约束条件引入.

与前一节类似，我们也有理想初值条件：

(a) $\varepsilon^{(0)}$与$\hat{\varepsilon}'$反号.

(b) $|\varepsilon^{(0)}|>\dfrac{|\hat{\varepsilon}'|}{\sigma}$.

同样，有从迭代终值出发的预报公式：

(a) $\hat{\varepsilon}'>0$ 时，

$$Z_{\infty}'(T_0+1)=Z^{(\infty)\prime}(T_0+1)-\frac{\bar{\hat{\varepsilon}}_+'}{1-\sigma'}. \tag{28}$$

(b) $\hat{\varepsilon}'<0$ 时,

$$Z_{\infty}'(T_0+1)=Z^{(\infty)\prime}(T_0+1)-\frac{\bar{\hat{\varepsilon}}_-'}{1-\sigma'}. \tag{29}$$

以上二式中, $Z_{\infty}'(T_0+1)$是由迭代终值$Z^{(\infty)\prime}(T_0+1)$出发求得的预报值, $\bar{\hat{\varepsilon}}_+'$和$\bar{\hat{\varepsilon}}_-'$分别是$\hat{\varepsilon}'>0$和$\hat{\varepsilon}'<0$时拟合误差$\hat{\varepsilon}'$的平均值.

在上面的讨论中, 假定了某一阶$(k\neq k_0)$的车贝雪夫多项式值$\varphi_{k_0}(t)$有误差. 容易看出, 对不只一阶(即k_0有多个)多项式值有误差的情况也是对的. 因此, 若$\varphi_k'(t)$尽管有误差, 但仍能满足(26)式, 则我们可以得到与应用准确的多项式值$\varphi_k(t)$类似的结果.

另外, 若计算车贝雪夫系数有误差δ_A, 则相当于相应的车贝雪夫多项式值有误差$\delta_\varphi(t)$. 简要证明如下:

设

$$A_k'=A_k+\delta_A=\sum_{t=1}^{T_0+1}Z(t)\varphi_k(t)+\delta_A. \tag{30}$$

令

$$\delta_A=Z(t_0)\delta_\varphi(t_0), \tag{31}$$

$$\varphi_k'(t_0)=\varphi_k(t_0)+\delta_\varphi(t_0). \tag{32}$$

则有

$$A_k'=\sum_{t=1}^{T_0+1}Z(t)\varphi_k'(t), \tag{33}$$

其中

$$\varphi_k'(t)=\begin{cases}\varphi_k'(t_0) & t=t_0,\\ \varphi_k(t) & t\neq t_0.\end{cases} \tag{34}$$

当然, 一个δ_A可以对应不同的$\delta_\varphi(t)$. 一般地, 可令[1)]

$$\delta_A=\sum_{t=1}^{T_0+1}Z(t)\delta_\varphi(t), \tag{35}$$

则有

$$A_k'=\sum_{t=1}^{T_0+1}Z(t)\varphi_k'(t). \tag{36}$$

因而

$$\varphi_k'(t)=\varphi_k(t)+\delta_\varphi(t). \tag{37}$$

证毕.

对$A_k^{(v)}$的计算误差也可做类似讨论.

因此, 我们得到如下结论: 当采用有误差的车贝雪夫多项式值或车贝雪夫系数值用

1) 满足此式的$\delta_\varphi(t)$可有无穷多组.

(17)式进行迭代运算时，只要满足约束条件(26)式，就可以得到与应用准确的车贝雪夫多项式值(此时车贝雪夫系数也是准确的)用(1)式进行迭代运算时类似的结果．从本质上看，采用带有误差的 $\varphi_k'(t)$，相当于引进了满足约束条件(26)式的一个新的系数系．这个函数系往往不再是车贝雪夫多项式，且往往是非正交的．从这个意义上讲，本节的结果是对作者以前提出的方法的推广．

由于 $\varphi_k(t)$ 有误差时在一定的约束条件下也能得到合理的计算结果，因此，若采用准确的 $\varphi_k(t)$ 值难以获得较好的预报效果时，我们可以适当地将 $\varphi_k(t)$ 做些改变(以满足(26)式为前提)，重新进行计算，并通过多种方案的比较设法取得理想的结果．这就为改进预报提供了广阔的途径．

需要说明，我们这里所说的误差，不是因计算机字长限制而产生的微小的舍入误差(这种误差在实际应用中可以忽略)，而是由于其它原因产生的较大的误差．我们在用原来的时间序列预报方法[1,2]进行计算时，由于偶然的原因(程序错误)在计算车贝雪夫多项式值时产生了差错．但分析结果时发现与由准确的多项式值算出的结果有类似规律．这一情况启发我们研究了上述关于车贝雪夫多项式值的误差对计算结果的影响．这个成果的积极意义，正如上面所述，我们可以在计算时有目的地将车贝雪夫多项式值作出适当调整，以改进预报效果．这点在下节将看到它的优越性．

四、计 算 实 例

本节给出一个计算济南异常旱涝预报的例子．从济南五百年旱涝资料1)，可以求得异常旱涝年的时间间隔．将此间隔的序列看成时间序列，则可以用时间序列预报方法做出未来异常旱涝年出现年份的预报2)．本例为旱年时间间隔预报．取 $T_0=8, K_0=2$.

有误差的车贝雪夫多项式为 $\varphi_1'(t)$，而 $\varphi_0(t)$ 与 $\varphi_2(t)$ 正确(即 $k_0=1$)．

表2给出有误差的 $\varphi_1'(t)$ 及正确的 $\varphi_1(t)$．

表2 $\varphi_1'(t)$ 和 $\varphi_1(t)$ 值

t	1	2	3	4	5	6	7	8	9
$\varphi_1'(t)$	−0.5681	−0.4418	−0.3156	−0.1894	−0.0631	0.0631	0.1894	0.3156	0.4418
$\varphi_1(t)$	−0.5164	−0.3873	−0.2582	−0.1291	0	0.1291	0.2582	0.3873	0.5164

还可算出 $\varphi_0(T_0+1)=\varphi_0(9)=0.3333$，$\varphi_2(T_0+1)=0.5318$．因此 $\sigma'=0.5891$，$\sigma=0.6606$．显然，σ' 满足(26)式．

旱年时间序列长度为61，对序列进行滑动分组，以序号1—9为第1个例，序号2—10为第2个例，依次类推，序号53—61为第53个个例．亦即，用 $Z(1)$，$Z(2)$，…，$Z(8)$ 预报 $Z(9)$；用 $Z(2)$—$Z(9)$ 预报 $Z(10)$；…；用 $Z(53)$—$Z(60)$ 预报 $Z(61)$．

1) 1470—1979年的资料取自文献[4]，1980—1987年的资料由沈长泗同志提供，作者谨致谢意．

2) 周家斌，一个异常旱涝预报方法，1989年长期预报研讨会文件．

计算前对序列进行标准化处理，其公式为:

$$Z=\frac{Y-Y_{\min}}{Y_{\max}-Y_{\min}}\times 100. \tag{38}$$

式中 Y 为旱年时间间隔，$Y_{\max}$ 与 $Y_{\min}$ 分别为其最大、最小值. 标准化处理后，

$$0\leqslant Z\leqslant 100.$$

表 3 给出计算结果示例. 表3a 是用准确的多项式值 $\varphi_0(t)$, $\varphi_1(t)$, $\varphi_2(t)$计算的，表 3b 是用 $\varphi_0(t)$, $\varphi_1'(t)$, $\varphi_2(t)$计算的，其中 $\varphi_1'(t)$有误差. 表中序号指预报对象的序号，Z 为实际值，Z_i 为理想初值预报结果，Z_∞ 为由迭代终值求得的预报结果，ε_i 和ε_∞ 为相应的误差. 由于个例太多，只列出部分结果作为示例，绝对误差的平均值为: 对表 3a, $|\overline{\varepsilon_i}|=24.9$, $|\overline{\varepsilon_\infty}|=25.9$; 对表 3b, $|\overline{\varepsilon_i}|=19.2$, $|\overline{\varepsilon_\infty}|=16.6$. 易见，两者误差相差不大，表 3b 的误差还要小一些. 由此可知，用经过调整的 $\varphi_1'(t)$参加运算，不仅能得到可用的预报结果，还可使预报误差有所减小. 图 1 是两种情况下理想初值的预报结果. 图1a 是用准确的 $\varphi_k(t)$计算的，图 1b 是用 $\varphi_0(t)$, $\varphi_1'(t)$, $\varphi_2(t)$计算的.

表 3　旱年序列计算结果

a 序号	Z	Z_i	ε_i	Z_∞	ε_∞	b 序号	Z	Z_i	ε_i	Z_∞	ε_∞
9	33.3	39.0	5.7	38.3	5.0	9	33.3	33.5	0.2	48.3	15.0
10	3.3	0.3	−3.0	−5.7	−9.0	10	3.3	0.0	−3.3	0.3	−3.0
11	0.0	0.0	0.0	−36.4	−36.4	11	0.0	0.0	0.0	−9.3	−9.3
12	10.0	41.1	31.1	40.5	30.5	12	10.0	33.5	23.5	48.3	38.3
13	96.7	44.5	−52.2	44.2	−52.5	13	96.7	34.3	−62.3	49.2	−47.5
14	0.0	62.2	62.2	61.8	61.8	14	0.0	13.6	13.6	29.8	29.8
15	0.0	1.2	1.2	−4.6	−4.6	15	0.0	0.0	0.0	3.8	3.8
16	40.0	40.4	0.4	39.7	−0.3	16	40.0	36.7	−3.3	51.7	11.7
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
61	83.3	79.5	−3.8	82.4	−0.9	61	83.3	55.3	−28.0	71.5	−11.8

本文关于预报方法的推广，对资料插补问题[5] 也是适用的. 此时，相应的公式化为

$$Z'^{(v)}(i_0)=\sum_{k=0}^{K_0}A_k'^{(v)}\varphi_k'(i_0), \tag{39}$$

$$A_k'^{(v)}=\sum_{\substack{i=1\\ i\neq i_0}}^{I_0}Z(i)\varphi_{k(i)}'+Z^{(v-1)}(i_0)\varphi_k'(i_0). \tag{40}$$

约束条件为

$$0<\sigma'(i_0)<1, \tag{41}$$

其中

$$\sigma'(i_0)=\sum_{k=0}^{K_0}\varphi_k'^2(i_0). \tag{42}$$

以上各式中 i_0 为待插补的点.

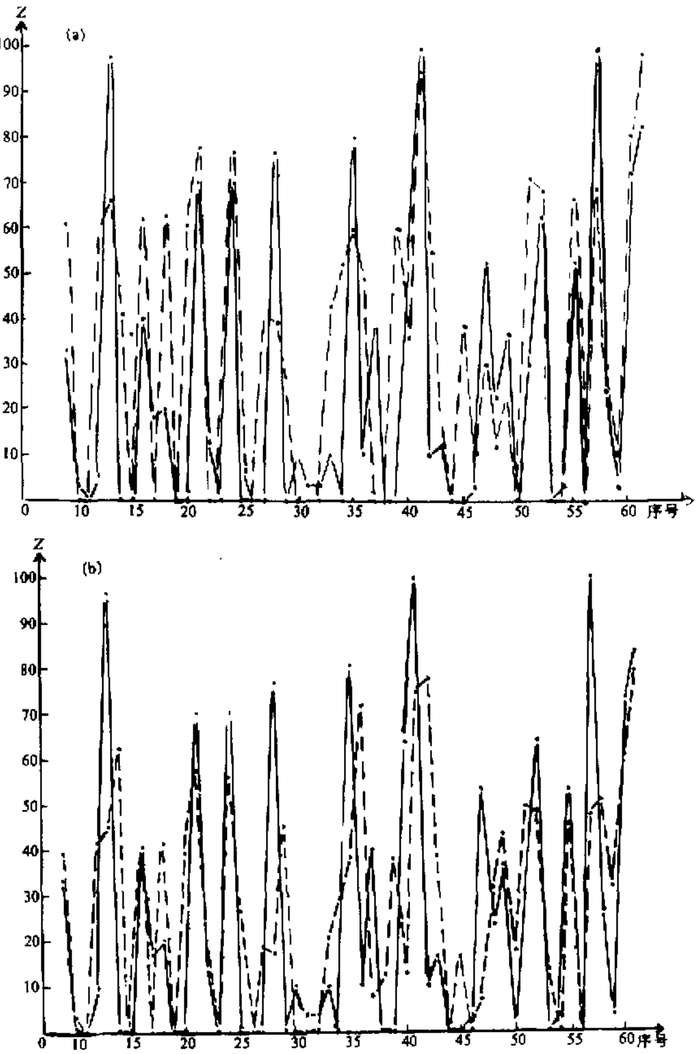

图1　济南旱年序列计算结果

a：用 $\varphi_k(t)$计算，b：用 $\varphi_k'(t)$计算． ——— 实际值， －－－－ 计算值．

二维情况下的公式是

$$Z'^{(v)}(i_0,j_0)=\sum_{k=0}^{K_0}\sum_{s=0}^{S_0}A_{ks}'^{(v)}\varphi_k'(i_0)\psi_s'(j_0), \tag{43}$$

$$A_{ks}'^{(v)}=\sum_{\substack{i=1 \\ (i,j)\neq(i_0,j_0)}}^{I_0}\sum_{j=1}^{J_0}Z(i,j)\varphi_k'(i)\psi_s'(j)+Z^{(v-1)}(i_0,j_0)\varphi_k'(i_0)\psi_s'(j_0), \tag{44}$$

$$\sigma'(i_0, j_0) = \sum_{k=0}^{K_0} \sum_{s=0}^{S_0} \varphi_k'^2(i_0)\psi_s'^2(j_0) < 1. \tag{45}$$

此时 i_0, j_0 为待插补的点.

参 考 文 献

[1] 周家斌，1985，一种新的时间序列预报方法，大气科学，9，No.1，27 — 35.

[2] 周家斌，1987，关于用车贝雪夫多项式做时间序列预报的几个问题，车贝雪夫多项式及其在气象水文中的应用文集，辽宁科学技术出版社，74 — 87.

[3] 粟运华，1987，关于使用新的时间序列预报方法做降水预报的一些体会，车贝雪夫多项式及其在气象水文中的应用文集，辽宁科学技术出版社，97 — 99.

[4] 中央气象局气象科学研究院，1981，中国近五百年旱涝分布图集，地图出版社.

[5] 周家斌，1987，一种气象资料插补方法，科学通报，32，No.15，1199 — 1200.

A GENERALIZATION TO THE FORECASTING METHOD OF TIME SERIES

Zhou Jiabin

(Institute of Atmospheric Physics, Chinese Academy of Sciences)

Abstract

It is discussed that how the errors of values of Chebyshev polynomials influence the predicted value obtained from the new forecasting method of time series proposed previously by the author. Based on this result the forecasting method has been generalized. The forecast skill is able to be improved by using the generalized scheme.

Key words: Time series; Forecasting method; Chebyshev polynomials; Set of nonorthogonal functions.