第5卷 第3期 1981年9月
大气科学 SCIENTIA ATMOSPHERICA SINICA
Vol. 5, No. 3 Sept., 1981

大气气温和风场垂直分布的统计特征及其在设计数值天气预报模式中的应用

黄荣辉 李荣凤
(中国科学院大气物理研究所)

提要

本文用经验正交展开的方法分别展开了北京、上海、广州三地区的无线电探空得到的温度和风场的垂直廓线. 指出在中高纬地区用6个参数就可以描述气温垂直分布特征,而在低纬度地区则需用7—8个参数才行. 而对于风场,则中高纬和低纬均必须用9个参数才能描述风场的垂直分布特征. 从此统计特征可以看出,为了较准确反映大气的斜压性,最少得用6层,若用实测风场来算初始风场,最好用9层模式,这样才能比较准确地反映风场的垂直结构.

一、引言

在数值预报及大气环流的数值试验中所用模式的分层问题愈来愈引起人们的重视. 若模式层次取得太少,这样在模式中反映气象要素的垂直结构及大气的斜压性就不好,势必在数值预报或大气环流的数值实验中造成较大的误差, 例如文献[1]中就说明了由于在模式中垂直分层太少,不能很好地反映大气的斜压性,结果在阻塞高压的数值实验中造成较大的偏差. 目前,国际上均采用层次较多的数值模式,而且有层次愈分愈多的倾向. 但层次很多的数值模式,都是计算量很大的. 因此,有必要探讨一下较准确表达其大气斜压性的最小层次数目. 此外,在客观分析中,现正在应用 Hough 函数,在垂直方向的 Hough 函数与经验正交函数相近[2-3]. 在求经验正交函数展开时有必要知道到底需要用前几个特征值所对应的特征函数来展开才行.

鉴于上述原因,有必要知道数值模式中温度T(可以由静力学平衡关系来求出位势高度)及风场这两个最重要的变量的垂直分布的统计特征,这样使我们明确在数值预报及数值试验中在机器许可的条件下所用模式的层数及客观分析中使之较好地反映其气象要素的统计特征及斜压性,从而得到更好的预报及实验结果.

二、经验正交展开

设$F(z, t)$为某测站气象要素的垂直分布,其中z为高度,t为时间,又记$\bar{F}(z)$为该要素的平均分布,$F'(z, t)$为偏差,即

1979年8月7日收到.

$$F(z,t)=\bar{F}(z)+F'(z,t) \tag{2.1}$$

把整个大气分为N层，取M个样品，并简记 $F(z_j, t_m)$ 为 F_{jm}，即：

$$F_{jm}=\bar{F}_j+F'_{jm} \quad (j=1,2,\cdots N;\ m=1,2,\cdots M)$$

现将每个样品 F'_{jm} 按某种正交函数组展开，取 $\widetilde{N}<N$ 项：

$$F'_{mj}=\sum_{i=1}^{\widetilde{N}}(f_m)_i x_{ij} \quad (i=1,2,\cdots\widetilde{N};\ j=1,2,\cdots N) \tag{2.2}$$

上式中 $(f_i)_m$ 为 $F(z_j, t_m)$ 的展开系数，x_{ij} 为第 i 个标准化正交函数在点 z_j 上的值，展开系数 $(f_i)_m$ 就是：

$$(f_m)_i=\sum_{j=1}^{N}(F'_m)_j x_{ji} \tag{2.3}$$

x_{ji} 为正交函数 x_{ij} 所组成的矩阵的转置.

要使展开后所得的要素值与原来所要求分布的差值(展开误差)对所有样品来说为最小，从理论上可以证明，这样的正交函数组就是该要素的协方差矩阵的特征向量[4]. 由此可见，这种函数(特征向量)完全由样品全体特点所确定，因此，这种展开法即称经验正交函数展开，这种函数(向量)称经验正交函数(向量).

某气象场 F' 的协方差矩阵为 $C_{F'}$

$$C_{F'}=\begin{pmatrix} \sum_{m=1}^{M}(F'_1F'_1)_m, & \sum_{m=1}^{M}(F'_1F'_2)_m, & \cdots\sum_{m=1}^{M}(F'_1F'_N)_m \\ \sum_{m=1}^{M}(F'_2F'_1)_m, & \sum_{m=1}^{M}(F'_2F'_2)_m, & \cdots\sum_{m=1}^{M}(F'_2F'_N)_m \\ \cdots\cdots & \cdots\cdots & \cdots\cdots \\ \sum_{m=1}^{M}(F'_NF'_1)_m, & \sum_{m=1}^{M}(F'_NF'_2)_m, & \cdots\sum_{m=1}^{M}(F'_NF'_N)_m \end{pmatrix} \tag{2.4}$$

因为上面协方差矩阵是一正定矩阵，其所有的特征值 $\lambda_i\geqslant 0$，若按其大小顺序排列，即 $\lambda_i\geqslant\lambda_{i+1}$ 等等. 若在展开时只取前 $\widetilde{N}$ 项，则均方根误差为

$$\sigma_N^2=\sum_{i=1}^{N}\lambda_i-\sum_{i=1}^{\widetilde{N}}\lambda_i=\sum_{i=\widetilde{N}+1}^{N}\lambda_i \tag{2.5}$$

由式 (2.5) 还可得到相对方差 ε_N 为

$$\varepsilon_N=1-\sum_{i=1}^{\widetilde{N}}\lambda_i\Big/\sum_{i=1}^{N}\lambda_i \tag{2.6}$$

很明显，若所取的 $\widetilde{N}$ 愈大，则相对方差愈小. 由于气象要素垂直分布的内在相关性，故在一定允许的精度范围内，$\widetilde{N}$ 不用取得很大，而展开的相对方差仍可达到允许的范围之内.

既然协方差矩阵是实对称的，所以我们用 Jacobi 方法或用收缩迭代方法求出所有特征值及特征向量.

三、大气中温度场的垂直分布特征

1. 大气温度场的垂直分布特征

大气温度场的垂直分布虽然随着时间和地点而异,但是从气候统计观点看,总有一定的共同性,例如在对流层温度随高度递减,在对流层与平流层之间存在着一个狭窄的过渡带——对流层顶,在对流层顶以上,温度开始是缓慢而后是比较快地随高度增加.所以,在一定的精度范围内,只需用有限个参数就可以反映出整层气温垂直分布特征.

2. 描述大气温度垂直廓线的参数个数

利用第二节所述的经验正交展开的方法分别统计北京、上海、广州三地区的温度垂直廓线,每个地区分别取 90 天探空温度廓线(从 1970 年 11 月 1 日到 1971 年 1 月 30 日)作为样品,每根探空温度廓线分成 20 层,由此得到 20 阶的协方差矩阵,需要分别求得这三个协方差矩阵的特征值与特征向量.表 1 表示这三个地区用前 4、6、8 个特征值与所有特征值总和之比.

表 1 T 的统计特性

比值% 项目 / 地区	$\sum_{i=1}^{4}\lambda_i/\sum_{i=1}^{20}\lambda_i$	$\sum_{i=1}^{6}\lambda_i/\sum_{i=1}^{20}\lambda_i$	$\sum_{i=1}^{8}\lambda_i/\sum_{i=1}^{20}\lambda_i$
北 京	93.3	96.3	98.0
上 海	90.5	94.2	96.5
广 州	82.6	89.1	93.6

图 1 表示北京气温垂直廓线前 3 个特征值所对应的特征向量. 从图 1 可以看出,最大特征值所对应的特征向量很类似于平均气温的垂直廓线.

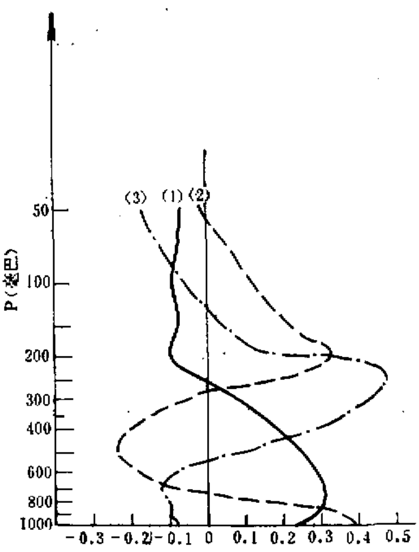


图 1 北京各层温度离差的相关矩阵前 3 个特征值所对应的特征向量. 括号中数字表示相应特征值的大小序号

从表 1 可以看出，对于北京、上海的气温垂直廓线来说，可用前 6 个特征值所对应的特征函数来展开，但对于广州来说，则用前 8 个特征值所对应的特征函数来展开才能得到比较好的结果. 这就是说，一般对于中高纬地带，气温垂直分布可以用 6 个参数来描述，而对于低纬度热带、亚热带地区，则必须用 7—8 个参数来描述.

四、大气中风场的垂直分布特征

1. 大气中风场的垂直分布特征

与第三节中一样，大气中风场垂直分布虽然随时间和地点而异，但从统计角度看，也有一定的共同性. 在中高纬地区冬季风速随高度增加，到了对流层顶附近风速达到最大，这个高度称急流高度，而往上风速随高度递减. 在低纬平流层 20 公里以上经常出现东风，在近地面层经常是东风. 因此，在一定精度范围内只需用有限个参数也可以反映出整层风速分布特征.

2. 描述大气东西向风速分量与南北向风速分量的垂直廓线参数个数

我们分别统计了北京、上海、广州三地区的风场垂直廓线，每个地区分别取 90 天测风资料(从 1970 年11月 1 日到 1971 年 1 月 30 日)作为样品，每个地点的每天测风资料分成 18 层，每层的风根据风速、风向分解成东西方向分量 u 与南北方向分量 v. 分别计算这三个地区风速的东西方向分量与南北方向分量的协方差矩阵，并求得这三个协方差矩阵的特征值与特征向量. 表 2 与表 3 分别为这三个地区 u、v 分量的协方差矩阵前 6、9、12 个特征值与所有特征值总和之比.

表 2 u 的统计特性

比值% 项目 / 地区	$\sum_{i=1}^{6}\lambda_i/\sum_{i=1}^{18}\lambda_i$	$\sum_{i=1}^{9}\lambda_i/\sum_{i=1}^{18}\lambda_i$	$\sum_{i=1}^{12}\lambda_i/\sum_{i=1}^{18}\lambda_i$
北京	94.6	97.1	98.8
上海	90.1	96.0	98.2
广州	94.3	97.1	98.6

从表 2 可以看出，对于北京、广州东西方向风速的垂直廓线可用前 6 个特征值所对应的特征函数来展开，但对于上海则要用前 9 个特征值所对应的特征函数来展开才能收到比较好的结果. 这就是说，一般可以用 6 个参数来描述其东西方向的风速垂直分布，更精

表 3 v 的统计特性

比值% 项目 / 地区	$\sum_{i=1}^{6}\lambda_i/\sum_{i=1}^{18}\lambda_i$	$\sum_{i=1}^{9}\lambda_i/\sum_{i=1}^{18}\lambda_i$	$\sum_{i=1}^{12}\lambda_i/\sum_{i=1}^{18}\lambda_i$
北京	88.8	94.0	97.0
上海	89.0	94.4	97.1
广州	83.7	91.2	95.6

确一些则需用 9 个参数来描述.

从表 3 可以看出，对于南北方向风速分量的垂直分布，一般要用 9 个参数来描述.

图 2 表示北京地区风的 u 分量垂直廓线的协方差矩阵前 3 个特征值所对应的特征向量，可以看出最大的特征值所对应的特征向量与平均气温廓线形状很相似.

五、用经验正交函数展开大气中温度场、风场的垂直廓线

1. 对气温垂直廓线的展开结果

用第三节所得到的结果，分别挑选矩阵前 4、6、8 个特征值所对应的前几组特征向量展开，这样分别得到用这三种不同参数数目展开的气温垂直分布廓线.

设 σ_t 为按经验正交展开所得各层近似气温值与探空廓线各层气温值的均方根误差

$$\sigma_t = \sqrt{\frac{\sum_{m=1}^{90}\sum_{j=1}^{20}[T_{jm} - (T_{jm})_0]^2}{90 \times 20}} \quad (5.1)$$

并设 σ_{T_0} 为原气温垂直廓线的离散度

$$\sigma_{T_0} = \sqrt{\frac{\sum_{m=1}^{90}\sum_{j=1}^{20}[(T_{jm})_0 - (\bar{T}_j)_0]^2}{90 \times 20}} \quad (5.2)$$

上两式中 T_{jm} 为每天用经验正交展开得到的各层气温近似值，$(T_{jm})_0$ 为每天各层气温的实测值，对于上述三地区所得结果如下：

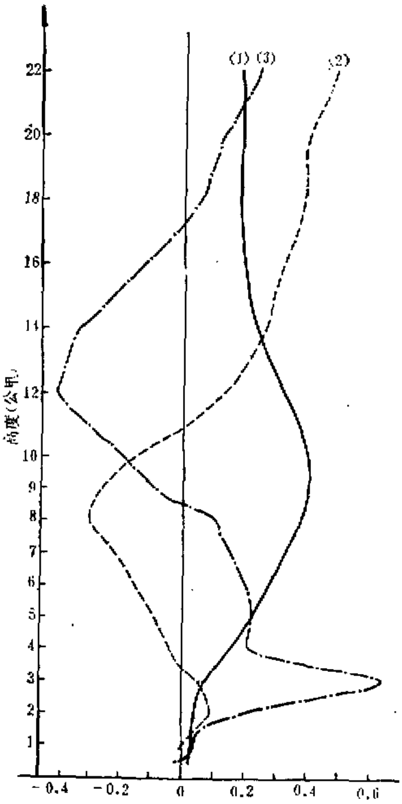


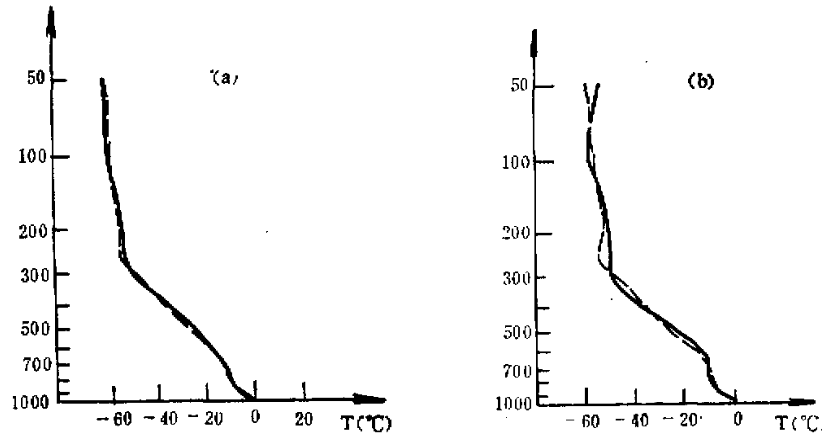
图 2　北京各层风速 u 分量离差的相关矩阵前 3 个特征值所对应的特征向量
括号中数字表示相应特征值的大小序号.

表 4

地区 \ 展开结果 (℃)	4 个参数		6 个参数		8 个参数		σ_{T_0} (℃)
	σ_T(℃)	σ_T/σ_{T_0} (%)	σ_T(℃)	σ_T/σ_{T_0} (%)	σ_T(℃)	σ_T/σ_{T_0} (%)	
北 京	1.7	34	1.0	19	0.5	10	5.0
上 海	2.2	48	1.1	24	0.8	17	4.6
广 州	1.6	54	1.0	33	0.6	20	3.0

从表 4 可看出，用 4 个参数来展开气温廓线显得粗糙些，不能很好地反映其垂直分布特征，展开后得到的各层气温值与各层气温实测值的均方根误差均超过 1.5℃，有的甚至

超过了 2℃；与其离散度比较，则相对比值均超过 30%，有的甚至超过 50%，这不满足气象上的要求．由表 4 可清楚看出用 6 个参数来展开气温垂直廓线一般是可以的，展开后所得各层气温值与各层气温实测值的均方根误差均在 1℃ 左右，一般来说这是气象工作所允许的，在探空仪精度范围之内．但是从表中也可以看出，对于纬度较低的广州，用 6 个参数来描述气温垂直分布，虽然绝对误差在 1℃ 左右，但相对离差仍达到 33%，这仍然是不能令人满意的．我们发现用 8 个参数来描述，对于低纬度显然效果好得多，其均方根误差只有 0.6℃，与其原气温离散度比较，其相对量已减少到 20% 了．



(a) 为 1970 年 11 月 22 日 12 时 00 分（世界时）

(b) 为 1970 年 12 月 2 日 12 时 00 分（世界时）

图 3 北京探空温度廓线（实线）与用经验正交展开得到的温度廓线（虚线）的比较．

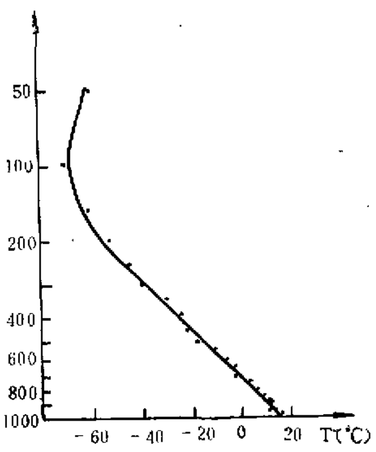


图 4 1970 年 11 月 7 日 12 时 00 分（世界时）上海探空温度廓线（实线）与用经验正交展开得到的温度廓线（点线）的比较

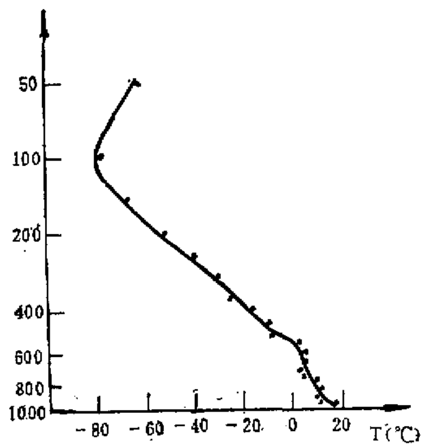


图 5 1970 年 12 月 11 日 12 时 00 分（世界时）广州探空温度廓线（实线）与用经验正交展开得到的温度廓线（点线）的比较

图 3(a)、4、5 分别为北京、上海、广州三地区用 6 个参数做经验正交展开所得的某一天近似气温廓线，其结果还是很好的，有的几乎重合。当然，我们所用的方法毕竟是统计方法，因此也有逼近较差的情况，如图 3(b) 就是这种情况，与实测气温分布误差较大，特别对于锋面逆温层这种细微结构则往往反映不出来。

2. 对风速垂直廓线的经验正交展开

用第三节所得的结果，分别挑选矩阵前 6、9、12 个特征值所对应的前几组特征向量展开。这样，就可以分别得到用这三种不同参数数目展开的风速 u 及 v 的垂直廓线。

设 σ_u（或 σ_v）为按经验正交展开所得各层近似气温值与探空廓线各层气温值的均方根误差

$$\sigma_u = \sqrt{\frac{\sum_{m=1}^{90}\sum_{j=1}^{18}[u_{jm} - (u_{jm})_0]^2}{90 \times 18}} \tag{5.3}$$

并设 σ_{u_0}（或 σ_{v_0}）为原风速 u（或 v）垂直廓线的离散度

$$\sigma_{u_0} = \sqrt{\frac{\sum_{m=1}^{90}\sum_{j=1}^{18}[(u_{jm})_0 - \bar{u}_j]^2}{90 \times 18}} \tag{5.4}$$

上两式中 u_{jm} 为每天用经验正交展开得到各层的风速 u 的值，则对于上述三地区所得的结果如下：

表 5

展开结果 / 地区	6个参数		9个参数		12个参数		σ_{u_0} (米/秒)
	σ_u (米/秒)	σ_u/σ_{u_0} (%)	σ_u (米/秒)	σ_u/σ_{u_0} (%)	σ_u (米/秒)	σ_u/σ_{u_0} (%)	
北 京	2.5	23	1.8	17	1.2	11	10.7
上 海	2.9	32	1.9	20	1.3	14	9.3
广 州	1.9	24	1.8	17	1.0	12	8.0

表 6

展开结果 / 地区	6个参数		9个参数		12个参数		σ_{v_0} (米/秒)
	σ_v (米/秒)	σ_v/σ_{v_0} (%)	σ_v (米/秒)	σ_v/σ_{v_0} (%)	σ_v (米/秒)	σ_v/σ_{v_0} (%)	
北 京	2.7	33	2.0	25	1.4	17	8.0
上 海	2.4	33	1.7	24	1.2	17	7.3
广 州	2.3	40	1.7	30	1.2	21	5.6

从表 5、表 6 可以看出：用 6 个参数来表示东西方向的风速廓线还是可以的，利用 6 个展开系数所得到的各层风速 u 值的均方根误差为 2—3 米/秒。虽然，用 6 个参数来表示南北方向的风速 v 廓线的均方根误差也在 2—3 米/秒范围内，但由于 v 分量风速本身就较小，所以其相对误差大，因此，用 6 个参数来描述 v 分量是不够的。而用 9 个参数来表

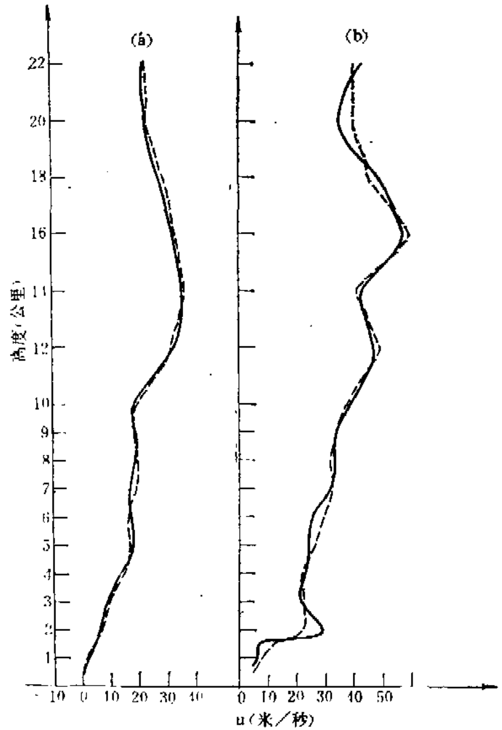


图 6　北京实测风速 u 分量廓线(实线)与用经验正交展开得到的风速廓线(虚线)比较
(a) 为 1970 年 12 月 5 日，
(b) 为 1971 年 1 月 11 日。

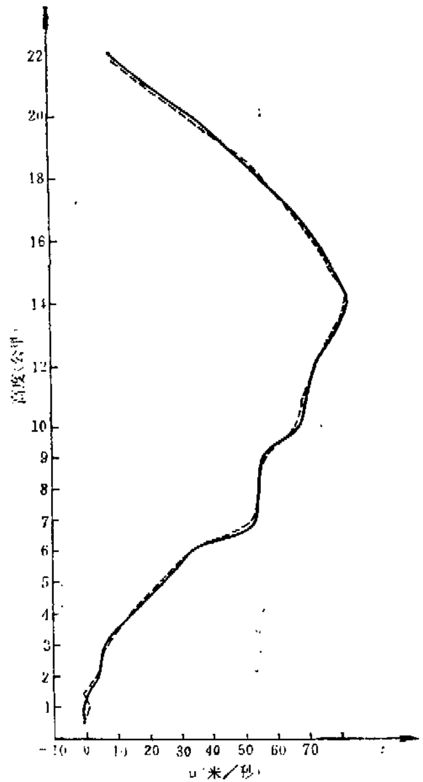


图 7　1970 年 12 月 28 日上海实测风速 u 分量廓线(实线)与用经验正交展开得到的 u 分量廓线(虚线)的比较

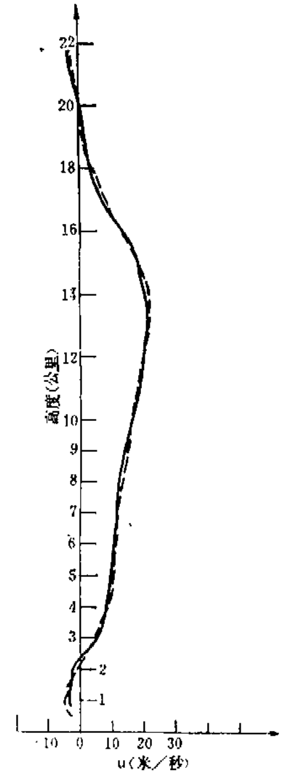


图 8　1970 年 11 月 27 日广州实测风速 u 分量廓线(实线)与用经验正交展开得到的 u 分量廓线(虚线)的比较

示东西方向与南北方向风速分量廓线，其结果更好，均方根误差在1—2米/秒内，与其离散度相比只有百分之十几.

图6(a)、7、8分别为北京、上海、广州三地区用9个参数作经验正交展开所得到某一天 u 分量风速近似垂直廓线，其结果还是相当好的，有的几乎重合. 当然，我们所用的方法毕竟是统计方法，因此也有逼近较差的情况，图6(b)就是这种情况.

表7、表8分别表示北京、上海、广州三地区由展开系数所得的风速 u、v 分量廓线最大均方根误差出现的层次，其结果如下：

表7　u 的最大均方根误差

地　　区	最大均方根误差	
	层次(公里)	$\sigma_{u\max}$ (米/秒)
北　京	18	2.5
上　海	7	2.4
广　州	3	1.8

表8　v 的最大均方根误差

地　　区	最大均方根误差	
	层次(公里)	$\sigma_{v\max}$ (米/秒)
北　京	9	12.1
上　海	7	10.9
广　州	5	10.1

从表7、表8可以看出，由9个参数作经验正交展开所得到的风速 u 分量最大误差为2—3米/秒，但 v 分量最大误差在10米/秒以上，这就是在风场客观分析中，不易准确的原因所在. 由此可见，在风场的客观分析与数值预报中，南北向的风速分量误差较大.

六、结　束　语

从上面计算与分析中可以得出结论：在中高纬地区用6个参数就可以描述气温垂直分布特征，而在低纬度地区则需用7—8个参数才行. 对风场说，则中高纬和低纬地区均必须用9个参数才能描述风场的垂直分布特征. 从这些统计特征可看出，在数值预报模式中，若初始风场是由高度场来计算的话，这样在模式中所用的经验正交函数只用前6个特征值所对应的6组特征函数来描述气象要素的垂直结构就足够精确. 这样，若有6层资料，则可由这6层资料及6组特征函数由(2.3)式来算得6个系数，由预报方程可以预报这6个系数的预报值，从而可得到任何层次的气象要素的预报值. 但是，如果所用的初始风场是由实测风算得，则必须用前9个特征值所对应的9组特征函数才可以比较准确地反映风场的垂直结构. 因此，这种情况必须用9层资料，从而得到9个系数. 这样，通过预报方程来预报这9个系数，则可得到任何层次的气象要素的预报值. 但，即使是这样，对

于南北方向的风速分量 v，有时误差还是比较大．即，在数值预报及客观分析中，v 分量的误差会比较大．

至于如何应用经验正交函数来作数值天气预报，我们还需作进一步的研究．

参 考 文 献

[1] 张镜福、刘瑞芝，动力气象学论文集(二)，科学出版社，1963．

[2] 曾庆存，数值天气预报的数学物理基础(第一卷)，科学出版社，1979．

[3] Kasahara A., Normal modes of ultralong wave in the atmosphere, *Mon. Wea. Rev.*, **104**, 669—690, 1976.

[4] 曾庆存，大气红外遥测原理，科学出版社，1974．

STATISTICAL FEATURES OF THE VERTICAL DISTRIBUTION OF TEMPERATURE AND WIND OF THE ATMOSPHERE AND THEIR APPLICATION TO THE DESIGN OF NUMERICAL WEATHER PREDICTION MODEL

Huang Rong-hui Li Rong-feng

(*Institute of Atmospheric Physics, Chinese Academy of Sciences*)

Abstract

In this paper the empirical orthogonal function is used to expand the observed vertical temperature profiles and the observed vertical wind profiles for Beijing, Shanghai, and Guangzhou. The results show that the vertical temperature profiles at the temperate and high latitude can be described by 6 parameters, the number of eigenvalues and corresponding eigenvectors, while 7—8 parameters are needed for low latitude. On the other hand, the vertical wind profiles can be represented by 9 parameters for the temperate and high latitude as well as low latitude. It is shown from these facts that in order to accurately describe the baroclinic characteristic of the atmosphere, at least 6 levels are needed. If the initial wind fields are computed from the observed wind data, it is adequate to use 9 levels in a primitive equation model.