

亚洲500毫巴候平均环流的天气统计学分析

章基嘉 孙照渤 兰国明

提 要

本文利用自然正交函数分解气象场的方法,对亚洲地区500毫巴候平均环流进行了天气统计学的分析研究,以期在此基础上,对五—十天的形势预报提供可参考的依据和途径。

分析计算表明,自然正交函数前五项有一定的天气学意义,尤其是前三项的天气学意义更明显。在所对应的时空尺度上,它们主要反映了长波系统的状况。对前五个特征向量场的时间系数演变进行分析以后发现,广大台站常用的45天,60天,90天,180天,300天的韵律关系,在这里有所反映。最后,用“分段方差迭代法”分析了时间系数序列,并进行了历史序列的拟合尝试,结果也是令人满意的。

一、统计学的分析

取1964—1973年十年500毫巴候平均图资料(共720候),在每候平均图亚洲范围内选取29点(图2—6中各点)。所取资料构成矩阵 ${}_nH_m$:

$${}_nH_m = \begin{pmatrix} h_{11} & h_{12} & \cdots & h_{1m} \\ h_{21} & h_{22} & \cdots & h_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ h_{n1} & h_{n2} & \cdots & h_{nm} \end{pmatrix} \quad (1)$$

$$n=720, \quad m=29,$$

其中 h_{ij} 是经中心化处理的高度值,即为实际高度减去历年相对应的各候平均值的距平。

利用自然正交函数展开,把 ${}_nH_m$ 分解成只依赖时间与只依赖空间的函数^[1]:

$${}_nH_m = {}_nT_m X_m \quad (2)$$

用求实对称矩阵特征值与特征向量的雅可比方法计算协方差矩阵

$${}_m A_m = {}_n H_m' {}_n H_m \quad (3)$$

的特征值 $\lambda_1, \lambda_2, \cdots, \lambda_m$ 和特征向量即自然正交函数 ${}_m X_m$,则时间系数为

$${}_n T_m = {}_n H_m {}_m X_m^{-1} \quad (\text{或 } {}_n T_m = {}_n H_m {}_m X_m') \quad (4)$$

若特征值 $\lambda_1, \lambda_2, \cdots, \lambda_m$ 按降序排列,则

指标 $\frac{\sum_{i=1}^k \lambda_i}{\sum_{i=1}^m \lambda_i}$ 表示取前 k 项所能表示原始场的比例, 可用以衡量展开的精确度。

计算结果的前十项列在表 1 中:

表 1 特征向量在总方差中所占比例表

序 号	1	2	3	4	5	6	7	8	9	10
λ_1	431.78	247.17	145.62	124.41	92.53	47.82	41.65	39.28	30.23	22.81
$\lambda_1 / \sum_{i=1}^m \lambda_i$	32.3%	18.5%	10.9%	9.3%	6.9%	3.6%	3.1%	2.9%	2.3%	1.7%
$\frac{\sum_{i=1}^k \lambda_i}{\sum_{i=1}^m \lambda_i}$	32.3%	50.8%	61.7%	71.0%	77.9%	81.5%	84.6%	87.5%	89.8%	91.5%

由表 1 可见, 用自然正交函数逼近候平均高度场时, 它的前 5 项可表示场的总方差的 77.9%, 前 10 项表示出 91.5%, 前 13 项就可表示出 95%, 由此可见, 用自然正交函数逼近高度场, 其收敛性是相当快的。

在实际分析自然正交函数时, 没有必要取全部的特征向量。经过分析, 我们选取前五个向量。因为当特征值 λ_N 按降序排列时, 与序号 N 近似呈 $\lambda_N = aN^{-x}$ 的关系, 也就是说, 在对数座标中, λ_N 与 N 近于线性关系^[2] (图 1)。

$$\lg \lambda_N = \lg a - x \lg N \quad (5)$$

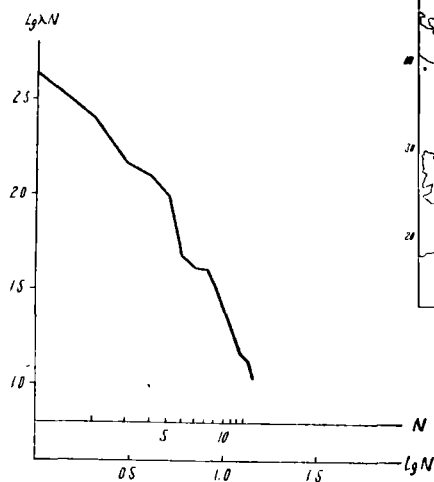


图 1、对数座标中不同序号的 λ

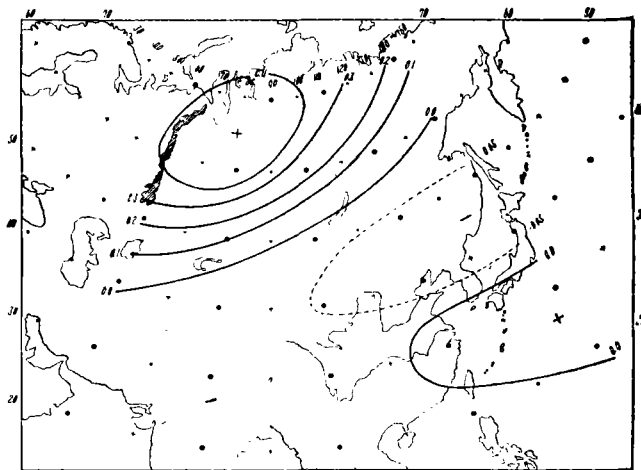


图 2、第一特征向量形势

由图 1 可见, 从 $N=5$ 到 $N=6$ 之间, 斜率比较大, 这意味着第 5 项以后的各项在高度场中所占比例明显减小, 是相对次要的。再结合对特征向量分析, 第 5 项以后, 系统比较零乱, 天气学意义不明显。因此, 选取前 5 个向量进行分析。这五个特征向量场所代表的环流形势是:

第一特征向量: 乌拉尔山阻高强烈发展, 付高主体位置偏北偏强, 极锋在我国华北

一带(图2)。

第二特征向量: 乌拉尔山阻高崩溃, 代之以高空低槽, 西伯利亚中部和东部高压脊异常发展, 付高减弱南退(图3)。

第三特征向量: 高空脊在亚洲西部发展, 东亚沿海高空槽加深(图4)。

第四特征向量: 在亚洲中高纬度范围内, 对流层中部呈现两脊一槽型(图5, 略)。

第五特征向量: 对流层中部呈现变形场型式(图6, 略)。

这五个向量表示了五种典型流场, 它们与时间系数(有正, 有负)相互配合, 基本上可以表示出亚洲500毫巴候平均环流的各种状况。

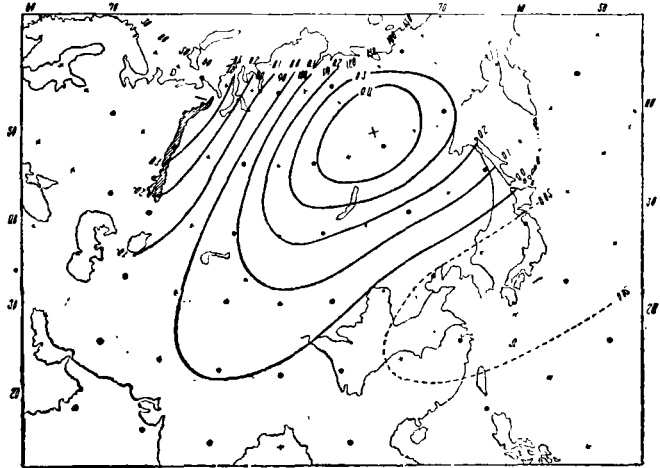


图3、第二特征向量形势

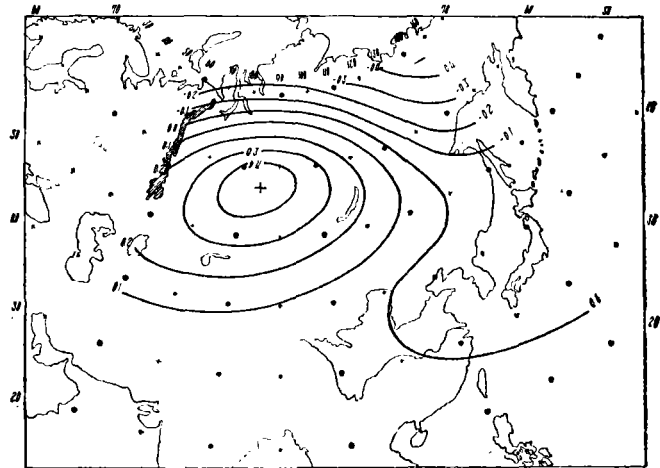


图4、第三特征向量形势

二、天气学的分析

时间系数从量上反映了亚洲500毫巴候平均环流典型流场随时间演变的规律。图7给出第一向量1964年各候的时间系数演变, 某候时间系数值越大, 表示该典型场所占的比例也越大。为了进一步从天气学意义上分析环流的演变规律, 我们选取部分比较典型的(即绝对值大的)时间系数所对应的候, 制作综合动态图, 由此出发分析该特征向量场所代表的典型流场的天气学意义。由于时间系数可正、可负, 一般地说, 它们所表示的环流形势的基本特征是相反的。分析结果表明:

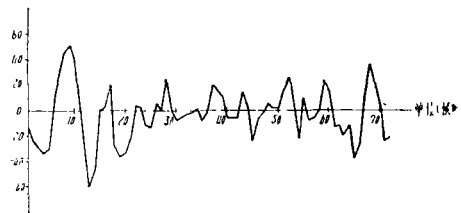
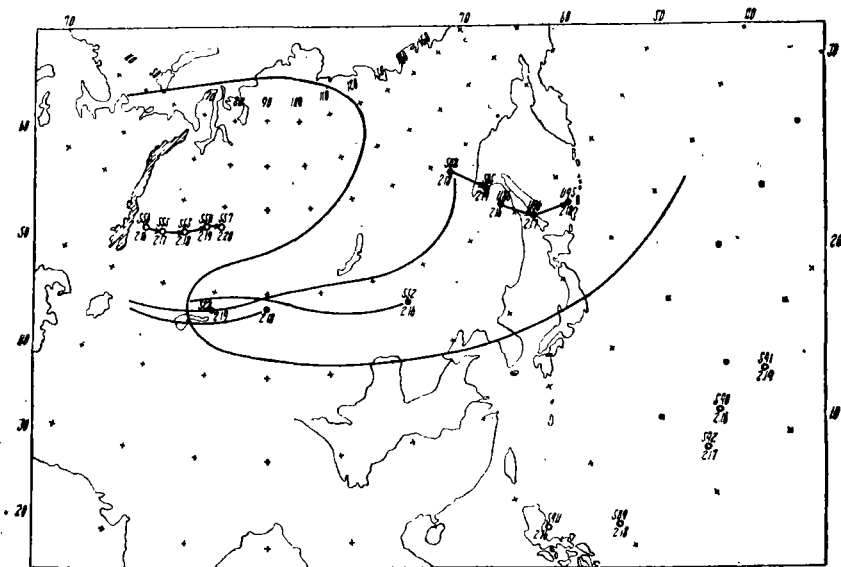


图7、第一向量1964年各候时间系数曲线

第一特征向量:

当时间系数为正时, 乌拉尔山有西南—东北向的高压脊区, 高压中心向东—东北方向缓慢移动; 从巴尔喀什湖到贝加尔湖一带为一横槽, 中纬度西风平直, 鄂霍茨克海为低压区, 低压中心自西北向东南移动。副高中心位置偏北、偏强。我国北方, 高空环流较平直, 时有小股冷空气不断南下。地面图上, 蒙古西部有冷高压, 冷空气活动路径偏北, 影响华北地区, 当横槽分裂或转向时, 也能侵入江淮流域及其以南地区, 导致一次寒潮爆发(图8)。



第二特征向量:

当时间系数为正时, 高压脊从中亚向东北方向伸展到西伯利亚东部, 从欧洲东部到乌拉尔山为低压槽控制, 日本海为东北—西南向的低压槽, 付高中心位置偏东(图10)。

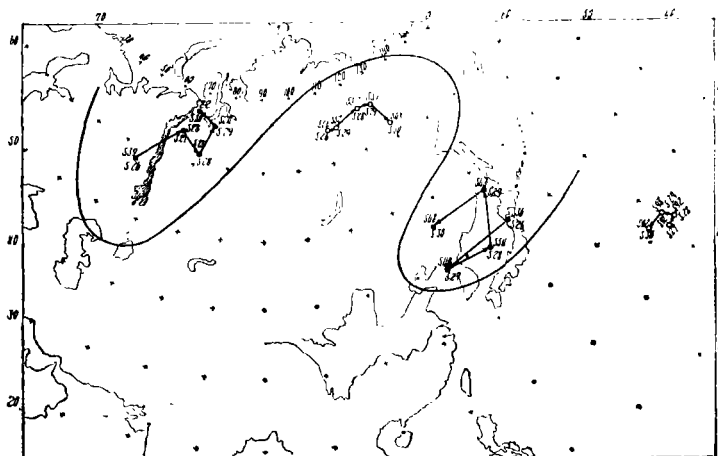


图10、1969年5月26—30日综合动态图

当时间系数为负时, 乌拉尔山地区为一高空脊, 从西伯利亚东部到蒙古为低压槽, 付高位置偏西, 中心在我国长江口附近(图11)。

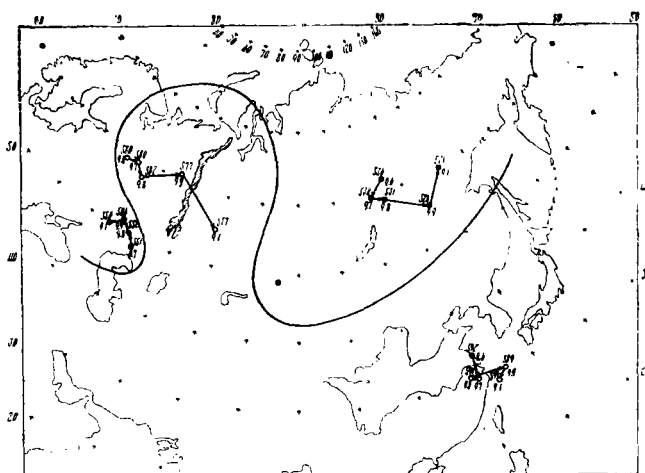


图11、1967年9月6—10日综合动态图

第三特征向量:

当时间系数为正时, 巴尔喀什湖到贝加尔湖有一弱高压脊, 鄂霍茨克海为一低压槽, 系统偏南。

当时间系数为负时, 乌拉尔山有一高脊区, 巴尔喀什湖以东直到沿海为一宽阔的低压区。

第四特征向量:

当时间系数为正时, 乌拉尔山到西伯利亚东部为一低压活动区, 欧洲和鄂霍茨克海到白令海一线分别为高脊区。

当时间系数为负时, 从贝加尔湖向北伸出一高压脊, 乌拉尔山以东和西伯利亚东部分别为一低压槽所控制。

第五特征向量:

基本上是由乌拉尔山高压脊、付高、西伯利亚东部低压区和中亚的切断低压所形成的变形场。

由分析看出, 在我们所使用资料的时空尺度上, 第一、第二两个向量场表示尺度较大的准静止性长波系统, 而第三、第四和第五向量场则分别表示尺度较小的移动性长波系统。

三、环流演变分析

由前面的天气统计分析可以看出, 研究环流演变规律, 只要研究时间系数随时间演变即可。

由图 7 可见, 时间系数序列是非平稳的, 为了使时间系数成为渐近平稳的, 我们进行了标准化处理, 即令

$$T'_{ti} = \frac{T_{ti} - \bar{T}_i}{\hat{\sigma}_i} \quad \begin{matrix} t=1, 2, \dots, 10; \\ i=1, 2, \dots, 72; \end{matrix} \quad (6)$$

其中 $\bar{T}_i$ 为历年各候平均值, $\hat{\sigma}_i$ 为历年各候样本均方差。

为揭露各特征向量场所代表的典型流场随时间演变过程中的周期性, 用标准化数据 (6) 计算了前五个向量的时间系数的标准化谱函数。样本标准化谱函数用公式^[4]:

$$S(\tau) = r(0) + 2 \sum_{\tau=1}^{100} r(\tau) \cdot \cos \frac{2\pi\tau}{t}$$

计算, 其中最大时间后延 $\tau=100$ (候)。

为了检验周期的显著性, 采用 χ^2 检验法^[6]:

$$\chi^2_{0.05} \approx k + 0.85 + 1.645 \sqrt{2k-1} \quad (7)$$

$$\chi^2_{0.01} \approx k + 2.20 + 2.326 \sqrt{2k-1} \quad (8)$$

其中 $K = \frac{2N-3}{2} m$ 。N 为资料总数, m 为最大时间后延。检验谱函数峰值显著性的

结果列于表 2 中。

表 2 谱函数峰值显著性检验结果

周期长度 α	向量 k_1	第一向量	第二向量	第三向量	第四向量	第五向量
$\alpha=0.05$		9, 14, 17, 21, 36, 60	10, 13, 15	16, 48	15, 19, 28, 36	5, 9, 14, 21, 37, 57
$\alpha=0.01$		9, 14, 17, 21, 36, 60	13, 15	16, 48	15, 19, 36	5, 14, 57

表 2 中 α 为信度, 周期长度 K_1 单位为候。

在检验的基础上, 选用若干最显著的周期拟合时间系数序列, 但拟合误差与原序列方

差很接近, 结果不理想。然而, 从表 2 中, 可以看到一个很有意义的事实, 即广大台站常用的一些韵律关系, 如 45 天, 60 天, 90 天, 180 天, 300 天左右的韵律都有所反映。在如此长的样本中得到这样一些统计上显著的关系, 是很值得令人注意的。

为了进一步寻找预报途径, 我们对文献〔5〕中提出的方差迭代法作了改进, 即采用分段方差迭代法进一步分析时间系数序列。因为一般方差分析法求得的周期函数是在不同周期相互影响的情况下求得的, 为了消去不同周期的相互影响, 采用分段方差迭代法进行分析。其主要作法如下:

假设原序列为: $T_1, T_2, \dots, T_L$, 用 $\{T_i\}$ 表示。

①第一段方差迭代: 用方差分析求得的长度为 K_1 的周期, 得到 K_1 周期的第一近似值 $\{T_i^{(1)}\}_{K_1}$, 对原序列 $\{T_i\}$ 减 $\{T_i^{(1)}\}_{K_1}$ 的差值, 用方差分析求得长度为 K_2 的周期, 得到 K_2 周期的第一近似值 $\{T_i^{(1)}\}_{K_2}$ 。

然后用原序列与 $\{T_i^{(1)}\}_{K_2}$ 的差值, 按 K_1 周期排列, 求得 K_1 的第二近似值 $\{T_i^{(2)}\}_{K_1}$; 又得原序列与 $\{T_i^{(2)}\}_{K_1}$ 的差值, 按 K_2 周期排列, 求得 K_2 的第二近似值 $\{T_i^{(2)}\}_{K_2}$ 。

这样往复多次, 直到剩余方差百分比不再显著减小为止。所谓剩余方差百分比, 就是把每个计算值与实际值的差值的平方和除以原序列总方差, 这里的计算值是指两个周期函数的迭加值, 称为第一段计算值。

②第二段方差迭代: 用原序列减第一段计算值之余差, 继续用①的办法选两个新的周期函数迭代, 可以得到这两个新的周期函数的迭加值, 再与第一段计算值迭加, 就得到第二段计算值。

③继续下去, 直到计算值与实际值之余差小于某个事先给定的允许误差而没有必要再进行迭代为止。

用上面简单介绍的分段方差迭代法, 对中心候各特征向量场的时间系数序列进行分析。所谓中心候时间系数是指两候时间系数的平均值作为该旬的中心候时间系数。我们取 69—73 年 5 年的中心候时间系数用分段方差迭代法分析, 结果列成表 (表 3)。

表 3 中心候时间系数分段方差迭代法计算结果

项 目 时间系数	第 一 段 迭 代			第 二 段 迭 代			原序列 均方差
	周期长度	R_D	S_a	周期长度	R_n	S_a	
第一向量	90, 74,	0.10	5.63	85, 89,	0.22	2.61	17.46
第二向量	90, 84	0.08	3.64	60, 74	0.19	1.58	12.94
第三向量	80, 90	0.09	3.15	89, 66	0.22	1.48	10.54
第四向量	88, 80	0.07	2.47	68, 87	0.25	1.23	9.46
第五向量	85, 59	0.14	2.81	89, 74	0.34	1.63	7.57

图12. 第一向量中心候时间系数曲线拟合比较

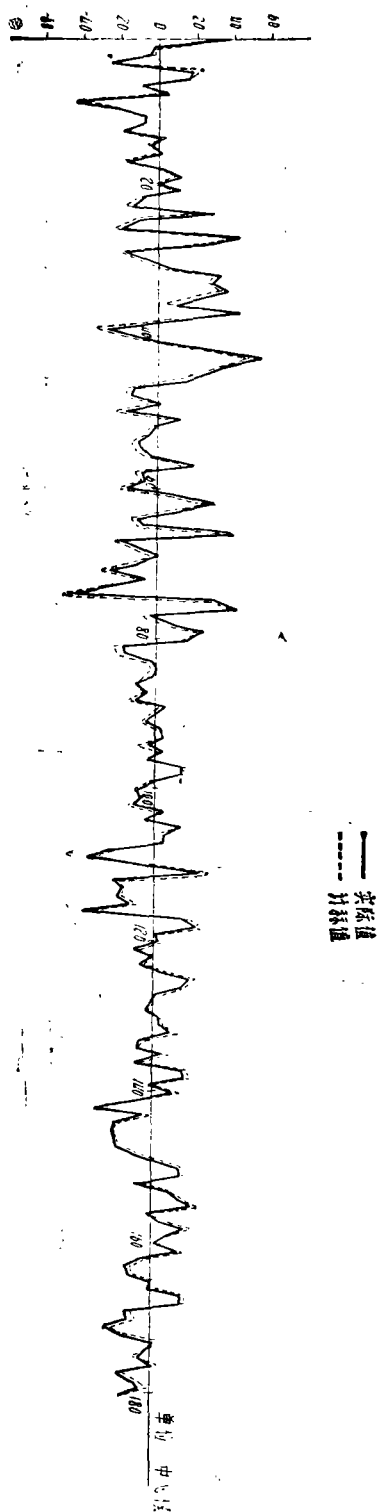


表3中周期长度的单位是中心候； R_D 为当段序列的剩余方差与当段序列的方差比值，用于检查当段的迭代精度； S_a 为分段方差迭代拟合误差的均方差。由表3可见，取两段迭代后，结果已是令人满意的了。在图12中给出第一特征向量的中心候时间系数序列实际值与拟合值曲线。由图可见，不仅趋势一致，而且误差也是很小的。

预报时，用周期函数外推迭加，就可报出以后某些指定时刻的中心候的时间系数。但是，应该注意到，时间系数序列并非平稳的，因此不能无限制地外推下去。这种状况涉及到气象预报的一个根本问题，即所有的气象要素时间序列几乎都是非平稳的时间序列。

四、几个问题讨论

1. 用自然正交分解方法，把气象场表示成随时间和空间变化的两部分，有利于研究气象场的演变。最大的优点是这种方法比另外一些按正交函数展开气象场的方法收敛快，少数几项就可比较好地表示出气象场的基本状况，而且天气学意义也清楚。从统计学观点看，实际天气过程可以看成是一些基本环流的演变与随机干扰的迭加，而随机干扰部分的天气学意义至少在目前还没有被认识。

2. 自然正交函数代表了一些典型的气象场的状况，它并不是固定不变的。而是随展开对象和资料长短而变化的，因此所谓自然正交函数不依时间变化也仅仅是对所展开的资料而言。对时间系数的预报是在假定自然正交函数以后时刻“稳定”不变的条件下进行的。但是，这种“稳定”在所展开的时空尺度上能维持多长时间仍是一个问题，这个问题与时间系数所允许预报的时效范围有密切关系，无论在实践上还是在理论上都是一个有待解决的问题。

3. 在我们所展开的时空尺度上，自然正交函数特征向量场表示了长波活动的空间形式，而其时间系数则表示长波系统随时间演变的规律。我们分别用谱分析方法和分段方差迭代法作了分析研究。分

段方差迭代法比方差分析求得的周期函数要好一些。而且用分段的方法可以灵活地按需要选取若干个周期，计算量也不大，方差迭代一般有6步左右就可以有相当令人满意的结果。而用谱分析方法没有取得令人满意的结果，原因是多方面的，从方法上讲，用分段方差迭代法选取的周期比较自然，而谱函数选取峰值的方法人为性大一些，这可能是其中的一个原因。

参 考 文 献

- 〔1〕 Н.А. Багров, Аналитическое представление последовательности метеорологических полей посредством естественных ортогональных составляющих Труды ЦИП вып74 1959.
- 〔2〕 М.И. Юдин, А.В. Мещерская Некоторые оценки естественных составляющих как предикторов и предиктантов Труды ГГО вып 273. 1972.
- 〔3〕 周华章编，工业技术应用数理统计学(下册)，人民教育出版社，1964.
- 〔4〕 Д.И. 卡扎凯维奇著，随机函数论原理及其在水文气象学中的应用，章基嘉译，科学出版社，1974.
- 〔5〕 北京大学地球物理系气象教研室编，天气分析和预报，科学出版社，1976.
- 〔6〕 章基嘉编著，中长期天气预报基础，南京气象学院油印本，1977.