综合评述

气象要素场的显著性检验

黄嘉佑

（北京大学地球物理系）

提　　要

本文介绍在气象要素场显著性检验过程中所存在的问题及其解决的方法，其中包括考虑序列持续性影响的单点检验及考虑空间相关性的Monte　Carlo等近代气象统计检验方法。

一、前　　言

在气候变化或长期天气过程分析研究中，常常会碰到关于气象要素场的显著性检验问题。一种是关于在两种不同气候背景下，两组气象要素场有无实质性的差别。例如要检验在厄尔尼诺现象发生的年份与反厄尔尼诺现象发生的年份我国降水场有无显著的不同。类似的检验还可使用在数值试验的“敏感性”试验中，以比较在两种不同的外力（例如热源）条件下，两组试验结果是否有显著的差异。另一种是要检验某一气象要素（例如南方涛动指数）与另一气象要素场（降水场）之间有无显著相关关系。

上述两种气象要素场的显著性检验的过程均有一个共同的特点，即通过对构成场的单点（站点或网格点）进行检验，寻找通过显著性检验水平的点形成一个或几个显著区域然后作出对问题的判断（是否有显著差异或相关存在）。问题是如何进行逐点的检验，以及有多少点通过检验才能判断该场是显著的。

本文着重介绍这类检验问题在气象中是如何进行的及其新的进展。由于这两类检验在逐点检验时所使用的统计量略有差别，所以在第二及第三部份分别介绍差异性及相关性的逐点检验过程，第四部份介绍关于场的显著性检验。

二、单点差异性检验

假设在两种不同气候背景下，抽取两组样本，容量分别为N_1及N_2。分别计算这两组样本的平均值为$\overline{x_1}$及$\overline{x_2}$。在单点情况下，要检验在不同气候背景的气候平均状态有无显著差异，可归结为检验这两组平均值有无显著差异。这种检验有两种做法，一种是当两组样本容量相等时，另一种是当两组样本容量不等时使用不同的统计量。

当两组样本容量相等时，即$N_1=N_2=N$。使用统计量[1]

$$t_1=\frac{\Delta\overline{x}}{\hat{s}} \qquad (1)$$

该统计量在假设两样本对应的总体期望值之差无显著差异的条件下，遵从自由度为$2N-2$的t分布。式中$\Delta\overline{x}=\overline{x_1}-\overline{x_2}$称为气候信号，$\hat{s}$为平均值差值的标准差估计，它表示为

$$\hat{s}=[(s_1^2+s_2^2)/N]^{1/2} \qquad (2)$$

$$s_1^2=\frac{1}{N-1}\sum_{i=1}^{N}(x_{1i}-\overline{x_1})^2$$

$$s_2^2=\frac{1}{N-1}\sum_{i=1}^{N}(x_{2i}-\overline{x_2})^2$$

x_{1i}，x_{2i}分别为两组样本中第i次观测值，$\hat{s}$亦称为气候噪音。统计量t_1又称为信号噪音比，检验可用一般的显著性检验方法进

行[2]。

当两组样本容量不等时，可用统计量[3]

$$t_2=\frac{\bar{x}_1-\bar{x}_2}{\left[\frac{N_1 s_1^2+N_2 s_2^2}{N_1+N_2-2}\left(\frac{1}{N_1}+\frac{1}{N_2}\right)\right]^{1/2}} \qquad (3)$$

该统计量在类似（1）式的假设前提下遵从自由度为（N_1+N_2-2）的t分布。显然，利用（3）式的检验过程实质上与用（1）式的信噪比检验过程是类似的。只是对平均值差值的标准差估计有所不同，也即是对气候噪音振幅水平的估计上。

根据（2）式，气候噪音方差的估计依赖于各组样本平均值的方差估计，例如当样本容量为N时，样本平均值的方差可表示为

$$\sigma_{\bar{x}}^2=\sigma^2/N \qquad (4)$$

但是，上式成立的前提条件是抽样样品是相互独立的。实际上，气象观测序列的抽样不可能完全满足相互独立的条件，因为一般气象要素时间序列存在较强的持续性。Leith[4]在70年代初就注意到这一问题，提出用有效自由度来修正原有的独立抽样个数，有效自由度为

$$N_{eff}=\frac{N}{T_0} \qquad (5)$$

式中T_0称为特征时间，即有效独立观测时间间隔，它依赖于样本自相关系数$r(\tau)$，即

$$T_0=1+2\sum_{\tau=1}^{N}\left(1-\frac{\tau}{N}\right)r(\tau) \qquad (6)$$

在非独立观测下，样本平均值表示为

$$\sigma_{\bar{x}}^2=\frac{\sigma^2}{N}\left[1+2\sum_{\tau\ 1}^{N}\left(1-\frac{\tau}{N}\right)r(\tau)\right] \qquad (7)$$

但是，在用样本来估计平均值的过程中，当落后步长τ增大时，$r(\tau)$的估计很不可靠。而且这一计算式是假定序列是红色噪音情况下导出的，因此Katz[5]认为对具体序列作持续性的具体估计，即用p阶自回归模型作估计，阶数p可用他提出的Bayes信息判据定出。平均值方差的估计可表示为

$$\hat{\sigma}_{\bar{x}}^2=\frac{1}{N}\hat{v}_p^2 \qquad (8)$$

其中

$$\hat{v}_p^2=\frac{\hat{\sigma}^2(p)}{\left[\sum_{R\ 0}^{p}\hat{\phi}_k(p)\right]^2}$$

$\hat{\sigma}^2(p)$为用p阶自回归模型$\hat{x}_t(p)$拟合原序列x_t的拟合残差方差，$\hat{\Phi}_k$为自回归模型中的系数，对于两组不同容量样本平均值差异是否显著可用统计量

$$z_1=\frac{\bar{x}_1-\bar{x}_2}{\left[\frac{1}{N_1}\hat{v}_{p1}^2+\frac{1}{N_2}\hat{v}_{p2}^2\right]^{1/2}} \qquad (9)$$

式中$\hat{v}_{p1}^2$及$\hat{v}_{p2}^2$分别为两组序列自回归模型拟合的方差估计，当N_1及N_2中最小的一个充分大时，统计量z_1遵从正态分布，就可以用相应的方法进行检验。

Jones[6]提出用功率谱作平均值方差的估计，并导出平均值方差与接近于频率为0时序列功率谱的关系为

$$\sigma_{\bar{x}}^2\sim\frac{1}{N}f_x(0) \qquad (10)$$

$f_x(0)$为序列x_t在频率为0附近的功率谱，实际计算时可用所有频率上的功率谱的平均来估计，即其估计功率谱为

$$F_x(0)=\frac{1}{n}\sum_{i=1}^{n}A(\omega_i)A^*(\omega_i) \qquad (11)$$

式中$A(\omega_i)$为在圆频率ω_i上复傅氏振幅谱，n为在$0<\omega<\omega_{max}$区间上抽样点数。在假设两组序列样本均来自同一总体的前提下，统计量可用[7]

$$t_3=\frac{\bar{x}_1-\bar{x}_2}{\left[\left(\frac{1}{N_1}+\frac{1}{N_2}\right)F_x(0)\right]^{1/2}} \qquad (12)$$

该统计量遵从自由度为2n的t分布。

在上述两组样本平均值差异的检验中，当其中某一组的样本容量退化为1个样品时，即检验来自同一总体中某一种特殊样品与一组容量为N的序列平均值$\overline{x}$有无显著性差异时，可以使用Trenberth[8]使用的统计量，即用统计量

$$t_4=\frac{x_0-\overline{x}}{s}\left(\frac{N-1}{N+1}\right)^{1/2} \qquad (13)$$

该统计量遵从自由度为（$N-1$）的t分布。式中x_0为特殊样品值，$\overline{x}$及s分别为另一组序列的平均值与标准差。对于非独立序列的情形也可以类似上面的介绍，对平均值的方差作出相应的修正或估计。

三、单点相关性检验

对于气象要素场中单点上与另一气象要素之间的相关性检验可以通过这两个序列的相关系数的显著性检验来实现。在原假设两个序列所对应的总体相互无关的情形下，对样本容量为N的两组序列相关系数r的检验可用统计量[9]

$$t_5=\sqrt{N-2}\frac{r}{\sqrt{1-r^2}} \qquad (14)$$

该统计量遵从自由度为（$N-2$）的t分布。

在相关系数的检验中，也有的使用Fisher的z变换进行[10]。在N较大时，统计量

$$z_2=\frac{1}{2}\log\frac{1+r}{1-r} \qquad (15)$$

遵从正态分布，其数学期望的估计值与（5）式相同，方差为（$N-3$）$^{-1}$。

当气象要素序列具有较强的持续性时，对相关系数的显著性检验同样也要考虑用有效自由度N_{eff}来代替独立抽样的自由度N。因此要使用相关序列的特征时间尺度[11]

$$T_1=1+2r_1r_1'+2r_2r_2'+\cdots \qquad (16)$$

来代替（5）式中的T_0，式中r_i及r_i'（i=1，2，…）分别为两个序列的落后自相关系数。

当自相关系数在阶数较大时，用样本估计会有较大的误差，因而使（16）式的T_1值估计发生问题。这时可以用修正序列的非独立性方法[12]，即把原序列转换成独立序列。既然序列存在持续性，可以借助自回归模型把持续性部分（由p阶自回归模型描述）和独立性部分（序列白噪音）分离。把序列x_t表示成

$$x_t=\hat{x}_t+a_t=\sum_{k=1}^{p}\Phi k x_{t-k}+x_t' \qquad (17)$$

式中$\hat{x}_t=\sum\limits_{k=1}^{p}\Phi k x_{t-k}$为持续性部分，$x_t'=a_t$为白噪音部分，阶数$p$可用Akaike的AIC准则定出。从（7）式可求出独立性部分x_t'。另一序列也作同样处理，然后用两序列白噪音序列求相关系数，这个相关系数即反映两序列独立抽样时的相关性，对它作显著性检验即可。

研究相关性时，落后交叉相关也常被研究。对落后交叉相关系数，可使用刘家铭[13]提出的一种检验方法。对以样本容量为N的两序列x_t及y_t，落后τ时刻的交叉相关系数为

$$r_{XY}(\tau)=\frac{c_{XY}(\tau)}{[c_{XX}(0)\,c_{YY}(0)]^{1/2}} \qquad (18)$$

其中$c_{XY}(\tau)$为两序列落后τ时刻的交叉协方差

$$c_{XY}(\tau)=\frac{1}{N-|\tau|}\sum_{t=1}^{N-|\tau|}(x_{t+\tau}-\overline{x})(y_t-\overline{y}) \qquad (19)$$

$\overline{x}$，$\overline{y}$分别为两序列平均值。在N较大情形下，交叉相关系数遵从均值为0，均方差为

$$SD[r_{XY}(\tau)]\doteq\left[\frac{2}{(N-|\tau|)\{1-\exp[1-(\lambda_X+\lambda_Y)]\}}\right]^{1/2} \qquad (20)$$

的正态分布，式中λ_X及λ_Y为假定序列为一阶马尔柯夫过程中的e折叠时间，即满足

$$\begin{cases} r_{XX}(\tau)=e^{-\lambda_X|\tau|} \\ r_{YY}(\tau)=e^{-\lambda_Y|\tau|} \end{cases} \qquad (21)$$

有了关于r_{XY} (τ) 的分布，就可用经典的正态分布检验方法作出有关的判断。

四、场的显著性检验

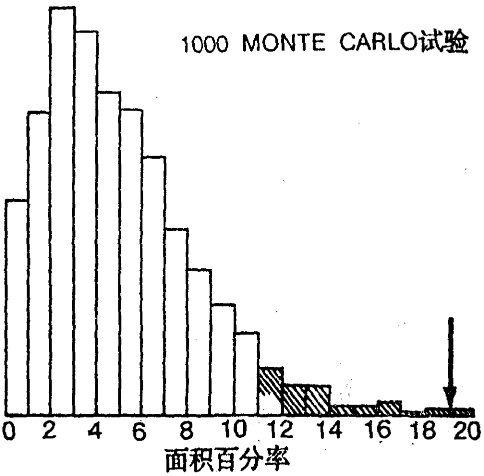
在差异场或相关场中，逐点的显著性检验，可能有若干个站点或网格点上的差异性或相关性被认为是显著的，它们组成一部分显著区域。但可否认为该要素差异场或相关场是显著的呢？显然，这要依赖于显著区域的大小，或者依赖于显著区域所包含的站点（或网格点）的数目。然而究竟多少个点通过显著性检验并由它们组成的显著域才能认为该场是显著的呢？这些点子的数目，或者点子数占全部格点数的百分率的大小本身也是一个要作检验的问题。通常检验有两种方法，一种是假定场中各格点的逐点检验是相互独立的。另一种则是真实场，即保留原场中各格点的空间依赖性，下面介绍以相关场为例的场检验过程。

Livezey等人[14]以概率观点把含 N 个格点的相关场检验看作为N次掷硬币试验，即每个格点只有两种可能结果：通过检验或不通过检验。这是一个二项分布的检验问题。设通过相关系数显著性检验的显著水平为 5 %，那么对某一格点来说，成功地通过检验的概率为p＝ 5 %，不成功的概率为q＝95%。根据二项分布，可以算出在N次试验中有M次事件发生（即有M点通过检验）的概率。例如当N＝30时，有 4 点通过检验的概率为0.045；有5点通过的概率为0.016，…。这样一来，可以算出“至少 4 点通过检验”事件的概率为0.045＋0.016＋…，约等于0.062。如果以5%作为场的显著性检验水平，那么可以计算出否定域的点子数为4.24个。即要求至少4.24个点通过检验，该场可认为是显著的。显著域占全场的百分率为14.1%（＝4.24/30）可作为临界显著域，类似地可算出对N＝80的场，临界显著域百分率为10%；对N＝500则为 7 %等。

当实际气象场格点间存在空间依赖关系时，独立性的假定就不能满足，而且它们的空间关系不象持续性那样可用某种模式模拟，它们随场不同而不同，没有统一模式来解决非独立性检验问题。因此，Livezey等人提出用统计模拟（Monte Carlo）方法来作实际相关场的检验。首先用遵从已知分布（例如正态分布）的随机数序列代替研究对象，这种随机数序列可以用迭代取中法、中平方法等方法由机器产生。然后计算它与实际要素场各格点的相关系数，产生一个相关场，从中得到一个显著相关区域的百分比；再由机器产生另一种随机数序列，又可得到第二个显著域的比例；如此作数百次或上千次的试验，最后得到一个显著域比例的经验概率分布，由它可定出某一显著水平下场的显著性判据。以Erickson[15]工作来说明，他计算700hPa高度与美国不同季节季平均气温的相关，得到相关场（图略），凡计算出相关系数绝对值大于0.35时为显著，得到显著相关区域，区域内包含有102个点，全场点数为528个，显著面积百分率为102/528＝19.3%。这一比例是否显著可作Monte Carlo检验，用随机数序列替代季平均温度，并作1000次试验，得到显著区域比例的经验概率分布（见附图）。实际相关场的比例用箭头表示其位置，它落在 5 %的否定域内，故否定相关场不是来自由随机数序列所产生的，即相关场显著。

Gorden[16]提出稍为不同的Monte Carlo检验方案，他根据实际相关场的各点相关系数，制作关于相关系数的经验概率分布密度曲线。然后用随机数序列代替研究对象，用大量Monte Carlo试验产生随机数序列相关系数的经验概率分布密度曲线，用χ^2统计量检验实测的与随机的概率分布是否有显著差异来判定相关场是否显著。

当然，上述关于相关场的显著性检验方法完全可以适用于差异场的显著性检验。



附图 700hPa 纬圈平均高度与随机数序列所得相关场显著区域比例经验概率分布

五、讨　论

气象要素差值场及与另一气象要素的相关场显著性检验，由于使用广泛而逐步被重视和深入研究。不同的考虑和方法还会进一步深入。例如，为了满足场中各格点的独立性要求，还可以对场作经验正交函数展开，由于展开后的分量相互正交，可以满足空间相互独立条件，对分量逐一检验，然后用二项分布检验即可。也有考虑场的空间结构，用多维检验方法进行场的显著性检验。可以期望，这方面的研究将会有更大的进展。

参 考 文 献

〔1〕 Hayashi. Y., Confidence intervalues of a climatic signal, J. Atmos. Soc., 39, 1895—1905, 1982.

〔2〕 陈家鼎，刘婉如，汪仁官，概率统计讲义，人民教育出版社，1980。

〔3〕 Lau Ka-Ming and P.H.Chen, Short-term climate variability and atmospheric teleconnections from satellite-observed outgoing long-wave radiation. Part I: Simultaneous relationships, J.Atmos. Sci. 40, 2735-2750, 1983.

〔4〕 Leith, C.E., The standard error of time-average estimates of climatic means, J.Appl.Meteor., 12, 1066-1075, 1973.

〔5〕 Katz, R.W., Statistical evaluation of climate experiments with general circulation model: A parametric time series modeling approach, J.Atmos. Sci., 39, 1446-1455, 1982.

〔6〕 Jones, R. H., Estimating the variance of time averages. J.Appl.Meteor., 14. 159-163, 1975.

〔7〕 Hannoschock, G. and C.Frankigonul Multivariate statistical analysis of a sea surface temperature anomaly experiment with the GISS general circulation model I.J.Atmos.Sci., 42, 1430-1450, 1985.

〔8〕 Trenberth, K.E., Interannual variability of the Southern Hemisphere circulation: Representativeness of the year of the global weather experiment, Mon Wea, Rev., 112, 108-123, 1984.

〔9〕 黄嘉佑 气象统计预报讲义. 北京大学地球物理系，1979。

〔10〕 Gutzler D.S., and K.C.Mo, Autocorrelation of Northern Hemisphere geopotential heights Mon. Wea. Rev., 111, 155-164, 1983.

〔11〕 Ananthakishnan, R. and B.Parthasarathy, Indian rainfall in relation to the sunspot cycle, J Climatololgy, 4, 149-169, 1984.

〔12〕 Chin W-C, A.Lo and D.H.Weidler, Jr, A study of the possible statistical relationship between the tropical Pacific sea surface temperature and atmospheric circulation. Mon. Wea. Rev., 109, 1013—1020, 1981.

〔13〕 Lau ka-Ming and P.H.Chen, Short-term climate variability and atmospheric teleconnections from satellite-observed outgoing long-wave radiation. part II: Lagged correlations. J.Atmos.Sci., 40., 2751-2767, 1983.

〔14〕 Livezey., R.E. and W.Y.Chen, Statistical field significance and its determination by Monte Carlo techniques, Mon. Wea. Rev., 111., 46-59, 1983.

〔15〕 Erickson. C.O., Hemispheric anomalies 700mb height and sea level pressure related to mean summer temperatures over the States. Mon.Wea.Rev., 111. 545-561, 1983.

〔16〕 Gordon, N.D., The southern oscillation and New Zealand weather. Mon.Wea. Rev., 111., 46-59, 1986.

The significant test for meteorological element field

Huang Jiayou

(Department of Geophysics, Peking University)

Abstract

The problems and their resolutions in testing the significance for meteorological element field are introduced in this paper. The influence of the persistence in series for the testing at single station and the correlation relationships of grid points in the field are considered and resolved by means of Monte Carlo and modern methods of statistical test.