资料同化中背景场位势高度误差统计分析的研究

庄照荣[1,2] 薛纪善[1] 庄世宇[1] 朱国富[1]

1 中国气象科学研究院，北京 100081
2 中国科学院研究生院，北京 100049

摘 要 在客观分析中，背景误差协方差对观测信息的传播和平滑、反映不同变量之间的关系有着非常重要的作用。构造合理的背景误差协方差矩阵对于同化系统至关重要，甚至会决定同化分析的好坏。作者主要利用观测余差方法，用 T213 预报资料和无线电探空观测资料统计我国区域的背景位势高度误差协方差样本，分析背景误差协方差场的结构特征和拟合误差场的空间分布。

关键词 背景误差协方差 特征尺度 变分同化 观测余差方法

文章编号 1006－9895(2006)03－0533－12 **中图分类号** P413 **文献标识码** A

A Study of the Statistical Analysis of the Geopotential Height Background Errors in the Data Assimilation

ZHUANG Zhao-Rong[1, 2], XUE Ji-Shan[1], ZHUANG Shi-Yu[1], and ZHU Guo-Fu[1]

1 *Chinese Academy of Meteorological Sciences*, *Beijing* 100081
2 *Graduate University of Chinese Academy of Sciences*, *Beijing* 100049

Abstract Background error covariance is very important to govern the amount of smoothing and spreading of the observed information and to decide the relationships between different variables in variational data assimilation. Because of the existence of a balance in the reality and in the model state, there is a version of the balance that exists in the background error covariances. Background error covariances depend on the uncertainty of the previous analysis and forecast. To a large extent, the form of this background error covariance governs the resulting objective analysis.

With the development of data assimilation, the methods to estimate the forecast error correlation structure have been reported in many literatures. However there is a little work about background error covariance in our country and the work is needed in the operational data assimilation system and GRAPES (Global and Regional Assimilation and PrEdiction System) 3D Var (three-Dimensional Variational data assimilation) research. So the statistical structure of background error covariance is studied in this paper.

It is difficult to directly get error covariances, which can only be estimated in a statistical sense. In order to get the height background error covariance, the innovation vector method is used in this paper. The data consist of innovation data (12 h and 24 h predicted height of T213 model minus radiosonde measurements) at 0000 UTC and 1200 UTC. Horizontal characteristic length, prediction error variance and observation error variance are obtained using Gauss correlation function approximation in a particular level. The straightforward way and the empirical thickness method are used to get the approximate function in interlevel values. In the vertical direction, vertical covariance approximation is obtained by the second-order autoregressive (SOAR) correlation function and distance transformation method. The resulting three-dimensional approximation function is partially separable, which is the product of the

收稿日期 2005－03－14，2005－07－11 收修定稿
资助项目 国家自然科学基金资助项目 40233036，国家“十五”科技攻关计划项目(2004BA607B)
作者简介 庄照荣，女，1978 年出生，博士生，主要从事数值预报的研究。E-mail：zhuangzr@126.com

horizontal covariance function and the vertical correlation function.

The major products of the analysis include: (i) In regions of sufficiently dense data coverage, the statistical analysis of innovation vectors can be employed. With homogeneous and isotropic assumption, it is a reasonable approximation to fit the horizontal covariances with Gauss function. (ii) It may be better way to calculate the covariances between the levels with empirical thickness method. (iii) The range of correlation distance parameter for prediction error is from 500 to 700 km and for synoptic scale prediction error is from 450 to 650 km in the troposphere. It shows that the range of the influence of large-scale prediction error is large. (iv) The prediction and observation error standard deviations of the values obtained by the T213 data are different from that of LAFS. Observation error is a little smaller than prediction error for the lower troposphere and is larger for the middle and upper levels of the troposphere. (v) The resulting three-dimensional covariances are approximated by a combination of Gauss function and SOAR function.

Key words background error covariance, characteristic length, variational data assimilation, the innovation vector method

1 引言

为了给数值预报模式提供恰当的初值，资料同化往往利用背景信息和观测信息得到最优的初值，其中作为权重的背景误差协方差非常重要，然而却很难得到。背景误差协方差决定着观测值订正到背景场的程度，以及观测点上的信息传播到分析格点上的方式。在观测资料稀疏的地区，分析结果主要由背景场决定，在观测资料稠密的地区，背景误差协方差矩阵中的相关性控制着观测信息的平滑。同化中可能出现的计算病态问题也往往和背景误差协方差矩阵有关。在类似大气这样的系统中，真实的背景误差依赖于天气形势，因此背景误差应和天气形势相协调①。

背景误差协方差的研究随着资料同化而发展。在最初的多项式插值客观分析方法中，只用了分析区域的观测信息。后来，Bergthorsson 等[1]提出的逐步订正法引入了背景信息，用观测值逐步修正背景场。Cressman[2]用简单的距离二次函数确定逐步订正法中的权重，而 Sasaki[3]和 Barnes[4, 5]用高斯函数确定逐步订正法中的权重。在这些方法中均没有对背景场误差做进一步的研究，在以后发展的统计插值和变分同化方法中，对背景误差协方差的研究才增多起来。在统计插值和三维变分同化方法中，一般假定背景误差协方差是均匀、各向同性的，而且在一定时段内不变，可以利用估计方法得到模型化的背景误差协方差[6~8]。在四维变分同化方法中，背景误差协方差可以随时间隐式地发展。在卡尔曼滤波方法中，预报误差协方差是随时间演变的诊断方程，预报值作为下一次同化窗的背景场，预报误差协方差作为下一次同化窗的背景误差协方差②。

最早利用观测资料统计背景误差场的是 Drozdov 等[9]和 Rutherford[10]，他们用密集无线电探空观测网上的观测值来统计背景误差相关。20 世纪 80 年代中期，Hollingsworth 和 Lonnberg[11]用观测余差方法对 ECMWF 全球同化系统的短期预报风场误差的结构进行统计，随后又分析了位势高度与风场的误差相关[12]。Franke[13]和许秦等[14]分别用二阶自回归函数和贝塞尔函数对 NOGAPS 数据进行了背景误差协方差统计分析研究。由于观测余差方法受到观测资料的限制，经验公式方法、NMC(National Meteorological Center)方法[15, 16]、集合预报等方法也得到进一步发展。

近二十多年，随着资料同化的发展，尽管背景误差协方差的估计方法有较大的改进，但目前我国的变分同化研究起步较晚，适合我国地形和天气形势的背景误差协方差的统计工作还很少开展，无论是我国现有的业务同化系统，还是最近正在发展的 GRAPES (Global and Regional Assimilation and PrEdiction System) 3D Var (three-Dimensional Variational data assimilation) 都缺乏这样的工作。

① Bouttier F, Courtier P. Data assimilation concepts and methods March 1999. Meteorological training course lecture series. ECMWF, 2002

② Thomas M Hamill. Ensemble-based data assimilation: a review. 2002, 9, 24

如前所述，背景场误差的统计特征是直接影响同化效果的重要因子，因此，研究我国资料同化的背景场误差是我国数值预报研究一项十分重要的基础性工作。本文主要利用观测余差方法统计分析我国区域的背景位势高度误差协方差场的结构特征和拟合误差场的空间分布。

2　背景误差协方差的估计方法和模型

假定观测互不相关，通过观测场和背景场可以得到较好的平均背景误差协方差(观测余差方法)，这种方法是以观测值作为参考的一个最简单、可靠的估计方法。它通常要求观测网稠密，能够提供足够多的信息。由于本文统计的是我国区域陆地上的背景误差协方差，观测资料较多，文中主要是用此方法来对背景误差协方差的结构进行研究。

如果定义 z 为预报值，z_o 为观测值，z_t 为大气真实状态值，则预报误差为 $z'=z-z_t$，观测误差为 $z''=z_o-z_t$，观测增量为 $z_d=z_o-z=z''-z'$。在三维坐标系中，用 i、j 代表位置的序号，m、n 代表垂直层的序号，则两点间的背景误差协方差可以定义为

$$\langle z'_{im}z'_{jn}\rangle=\langle(z_{im}-z_{t,im})(z_{jn}-z_{t,jn})\rangle, \quad (1)$$

其中，$\langle\cdot\rangle$代表数学期望。假设观测误差和预报误差都是无偏的，而且两者互不相关，则观测增量的误差协方差可以分离为观测误差协方差和背景误差协方差，即

$$\langle z_{d,im}z_{d,jn}\rangle=\langle z''_{im}z''_{jn}\rangle+\langle z'_{im}z'_{jn}\rangle. \quad (2)$$

如果进一步假定观测误差只在垂直方向是相关的(对多数观测而言，这是一个合理的假定)，而且背景场误差在水平方向是均匀、各向同性的，由(2)式得：

$$\langle z_{d,im}z_{d,jn}\rangle=\begin{cases}C_B(r,p_m,p_n), \\ \qquad i\neq j\ (r\neq 0), \\ C_O(p_m,p_n)+C_B(0,p_m,p_n), \\ \qquad i=j\ (r=0),\end{cases} \quad (3)$$

其中，$C_B(r,p_m,p_n)$为背景误差协方差，$C_O(p_m,p_n)$为观测误差协方差。(3)式表明，在同一观测点($i=j$)，观测增量的协方差等于观测误差协方差与背景误差协方差的和；在不同观测点($i\neq j$)上，由于观测误差之间不相关，观测增量的协方差等于背景误差协方差。

利用以上方法可以得到观测站点间的背景误差协方差样本，为了得到背景误差协方差模型，文中用高斯型相关函数对样本进行拟合。所采用的相关函数为

$$C_B(r)=E_B^2\exp\left(-\frac{r^2}{2L^2}\right), \quad (4)$$

其中，L 为水平特征尺度，r 为距离，E_B^2 为背景误差方差。这种高斯模型的谱密度总是正的，满足构造相关模型的条件。

3　使用的资料和方法

本文所用的无线电探空观测资料为 2003 年 5 月 14 日～2003 年 7 月 14 日，0000 UTC 和 1200 UTC 时刻的资料。预报场资料来源于国家气象中心 T213 模式的 12 小时和 24 小时预报。T213 资料的标准气压层有 17 层，从 1000 hPa 到 10 hPa。研究总区域范围为：(5°N～65°N，55°E～140°E)。研究区域在 0000 UTC 有 238 个探空观测站点，在 1200 UTC 有 221 个探空观测站点。

由 0000 UTC 研究区域范围与站点分布(图 1)可见，接近大陆的海域和较高纬度的无线电探空资料比较稀少，探空资料主要集中在文中选定的矩形区域中(20°N～60°N，80°E～120°E)。在此矩形区域中，0000 UTC 有 127 个探空观测站点，1200 UTC 有 120 个探空观测站点。文中通过计算此区域的每一个点和总区域其他点的观测余差协方差，最终得到背景误差协方差样本。

把搜集到的预报场资料用双线性插值方法插值到观测站点上，然后用观测余差方法，把每一时间点每一层的观测值减去预报值，得到观测增量。由于假定观测值和背景值是无偏的，观测增量还要除

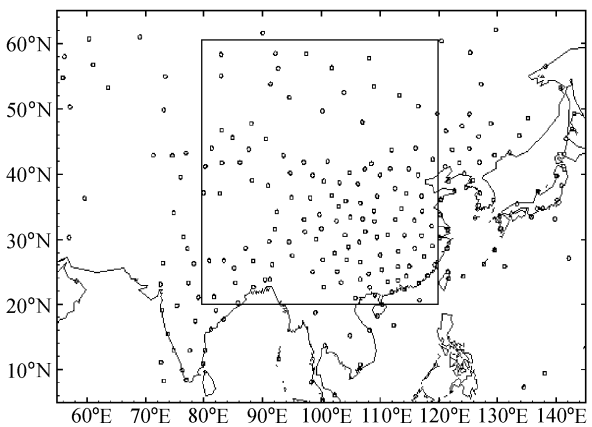

图 1　0000 UTC 研究区域范围与站点分布

Fig. 1　The studied region and station locations at 0000 UTC

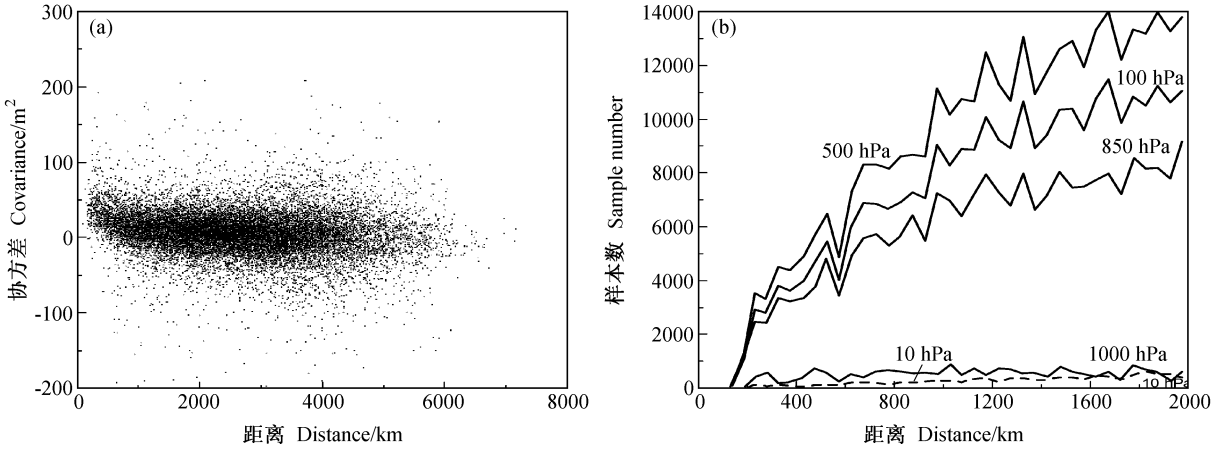

图2 (a) 500 hPa 0000 UTC背景误差协方差样本点的分布；(b) 0000 UTC不同高度层的背景误差协方差样本点数量分布

Fig. 2 (a) Distribution of the sample point of background error covariance at 500 hPa at 0000 UTC; (b) distribution of the sample point number of background error covariance at different levels at 0000 UTC

去偏差，即再减去每一个站点每一层的观测增量的时间平均，得到需要的分析数据，在此分析数据基础上研究背景误差协方差的结构。在进行背景误差协方差统计之前，对探空观测资料进行了极值检测，对观测余差进行了初估场的极值检查和水平一致性检查。

由观测余差方法可获得背景误差协方差样本点对$(d_{i,j}, E_{i,j})$和所有可能的样本点对的数量$N_{i,j,t}$，其中$d_{i,j}$、$E_{i,j}$表示i、j两点间的距离和背景误差协方差。图2a为500 hPa 0000 UTC的位势高度背景误差协方差随距离的分布，从图中可以看出，背景误差协方差样本点在5000 km之内非常稠密集中，很难看出背景误差协方差的分布规律。图2b是以每50 km间隔统计的不同高度层背景误差协方差样本数量随距离的分布，在1000 hPa和10 hPa样本点数量最多时也只有几百；在850 hPa、500 hPa和100 hPa，300 km以内背景误差协方差样本点较少，300 km以外，样本点数量达到2000个以上，并随距离迅速增加。从图2还可以看出500 hPa的背景误差协方差样本数量最多。由于大尺度误差的存在，在两点距离为2000 km以外的背景误差协方差趋近于常数，本文所求的背景误差协方差样本点对之间的距离限制在2000 km之内。

由于得到的背景误差协方差样本点分布不均匀，为了更好的进行拟合，对背景误差协方差样本按距离每隔50公里分一组，对每一组中的背景误差协方差和站点间的距离求平均，一共可得40组。分组后背景误差协方差样本对为$(d_n、E_n)$，$n=1, 2, \cdots, 40$。然后，采用高斯相关模型来拟合分组后的背景误差协方差样本。拟合时用最小二乘法：

$$\sum_{i=1}^{M} W_i \left[C_{Bi} - C_B(0)\exp\left(-\frac{r_i^2}{2L^2}\right)\right]^2 = \min_{C_B \in H} \sum_{i=1}^{M} W_i \left[C_{Bi} - C_B(0)\exp\left(-\frac{r_i^2}{2L^2}\right)\right]^2, \quad (5)$$

其中，W_i为权重，C_{Bi}与r_i为分组后的背景误差协方差和站点对之间的距离，M为分组后的样本数量，$C_B(0)$为背景误差方差，L是水平特征尺度，H为给定的函数类。考虑到每组中背景误差协方差样本点的数量不一样，距离在0～100 km的样本点几乎没有，200 km之外的样本点数量较多。样本点对的数量越多，资料越可信。参考Franke[13]对权重的选取，文中给出六种权重选择方案，分别为：1、N、N^2、N/R、N/R^2、$1/R$。其中N为每组中的样本点对总数，R为每组的平均距离。通过高斯函数拟合分组后的背景误差协方差样本，可得到某一区域的背景误差方差和水平特征尺度。由于设计的高斯函数较简单，而且1000 km以外的背景误差协方差趋于常数，用高斯函数拟合的背景误差协方差样本的站点间截断距离为1000 km。

4 结果分析

4.1 背景与观测误差方差的分离

用观测余差方法可以得到单层的背景误差协方差样本分布和拟合函数(如图3)。在500 hPa(图

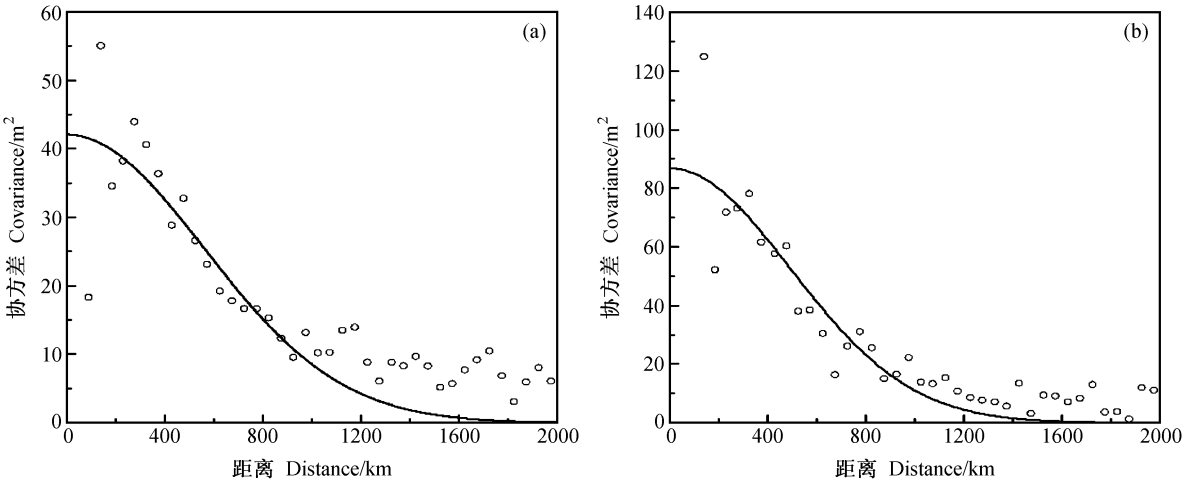

图 3　0000 UTC 背景误差协方差样本点分布和拟合：(a) 500 hPa；(b) 300 hPa

Fig. 3　Distribution and fitting of the sample point of background error covariance at 0000 UTC: (a) 500 hPa; (b) 300 hPa

3a)背景误差协方差随着两点间的距离增大而递减，递减到 1000 km 以后，几乎为一恒量，恒量部分是背景误差的大尺度部分。由于截断距离为 1000 km，在 1000 km 之内高斯拟合曲线对背景误差协方差样本拟合得较好，在 1000 km 之外拟合有一定偏差。在 300 hPa (图 3b)高斯相关函数能很好地拟合背景误差协方差样本，拟合曲线能反映实际背景误差协方差样本的分布特征。在 1000 hPa 和 10 hPa 高斯函数对背景误差协方差样本的拟合略差(图略)，这是由于在最低层和最高层，无线电探空观测资料较少，相应得到的背景误差协方差样本数量也最少，背景误差协方差样本点分布较零散，曲线拟合的效果就不好。

对于不同高度层之间的背景误差协方差，可以用两种不同的方法来得到。最直接的方法与求同一层的背景误差协方差方法一样，另一种方法是用观测余差方法计算出两层间厚度误差协方差，然后根据厚度误差协方差和高度误差协方差的关系，再获得高度误差协方差。如果 D_{lk} 为厚度误差方差，C_{lk} 代表 l 层和 k 层的垂直误差协方差，i 表示水平某一点，可得到

$$D_{lk}=\overline{(z_{li}-z_{ki})(z_{li}-z_{ki})}=C_{ll}-2C_{lk}+C_{kk}. \quad (6)$$

由式(6)，不同高度层间高度预报误差协方差为

$$C_{lk}(r)=\frac{1}{2}\left[C_{ll}\exp\left(-\frac{r^2}{2L_{ll}^2}\right)+C_{kk}\exp\left(-\frac{r^2}{2L_{kk}^2}\right)-D_{lk}\exp\left(-\frac{r^2}{2L_{lk}^2}\right)\right], \quad (7)$$

$C_{lk}(r)$为两层间的背景误差协方差函数，它可以看作高斯相关模型的一种组合。

从图 4a 可以看出 500 hPa 与 400 hPa 之间的厚度误差协方差样本分布比较集中，和高斯拟合曲线偏差较小。在 1000 km 以内，用厚度方法得到的高度误差协方差近似曲线能比较好地反映两层间的高度误差协方差分布情况，而且得到的预报高度误差方差较大(如图 4b 所示)。厚度方法用 4 个空间点资料进行计算，而高度方法只用两个空间点资料得到高度误差协方差样本。所以，用厚度方法得到的高度误差协方差函数更能准确的反映实际情况。

对同一层的背景误差协方差样本点进行高斯函数拟合后，用观测余差方法可以分离背景误差和观测误差。

图 5 为 0000 UTC 和 1200 UTC 背景误差和观测误差随高度的变化，其中总标准偏差为观测方差和背景方差的和再开方。从图 5 中可以看出，从 1000 hPa 到 100 hPa 背景误差的范围在 5～25 m 之间，而观测误差的范围在 6～40 m 之间。0000 UTC (图 5a)在 1000 hPa 与 500 hPa 之间，背景误差随着垂直高度变化不大，在 500 hPa 以上，背景误差随垂直高度的增加而逐渐变大。观测误差在 850 hPa以下比背景误差略小，在 700 hPa 以上随垂直高度的升高而逐渐变大，且比背景误差大。1200 UTC(图 5b)在 700 hPa 以下背景误差比观测误差大，在 700 hPa 以上两者随垂直高度层的增加而逐渐增大，观测误差随高度变化较剧烈。图 5a 和 b 对比，可以看出位势高度预报误差在 0000 UTC 与 1200 UTC 的大小和变化趋势其本一样，位势高度

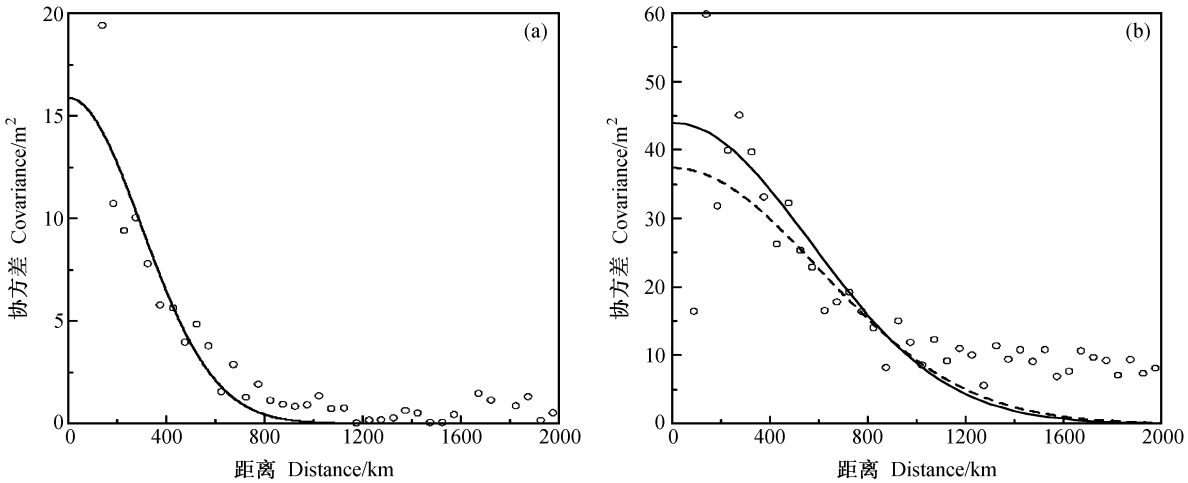

图 4 0000 UTC 500～400 hPa 间的背景误差协方差样本分布及拟合(实线为厚度方法拟合函数；虚线为高度方法拟合函数)：(a)厚度误差协方差；(b)高度误差协方差

Fig. 4 Distribution and fitting of the sample of background error covariance for 500－400 hPa at 0000 UTC: (a) Thickness errors; (b) height errors (solid line is thickness method fitting function and dashed line is height method fitting function)

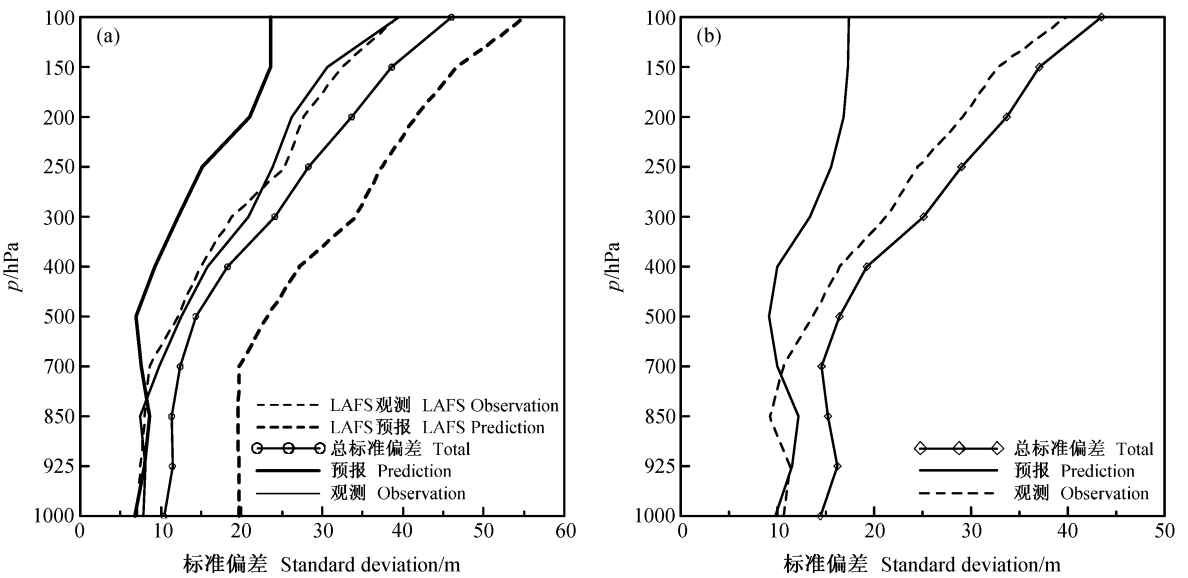

图 5 位势高度预报、观测和总标准偏差随高度的变化：(a) 0000 UTC；(b) 1200 UTC

Fig. 5 Standard deviations of predicted and observed geopotential height, and total errors with the values presently assumed in LAFS: (a) 0000 UTC; (b)1200 UTC

观测误差也是这样，说明预报误差、观测误差和总误差在一天中基本不变化。

从图 5a 还可以看出，统计得到的位势高度误差(实线)与当前国家气象中心有限区域数值预报业务系统(LAFS：Limited area Analysis and Forecasting System)①所使用数值(虚线)作比较，统计的观测误差与 LAFS 的大小量级相差不大，两者随着垂直高度的增加而逐渐增大的趋势和变化剧烈程度也类似。而统计得到的位势高度预报误差与 LAFS 的预报误差大小相差较大，两者随垂直高度变化的趋势相同。

Franke[13]用二阶自回归函数拟合得到的位势高度预报误差在 100 hPa 以下也比观测误差小，和 NOGAPS 所使用的误差标准相反。许秦[14]用贝塞

① 有限区域数值预报业务系统(LAFS). 国家气象中心编，1994，7～30

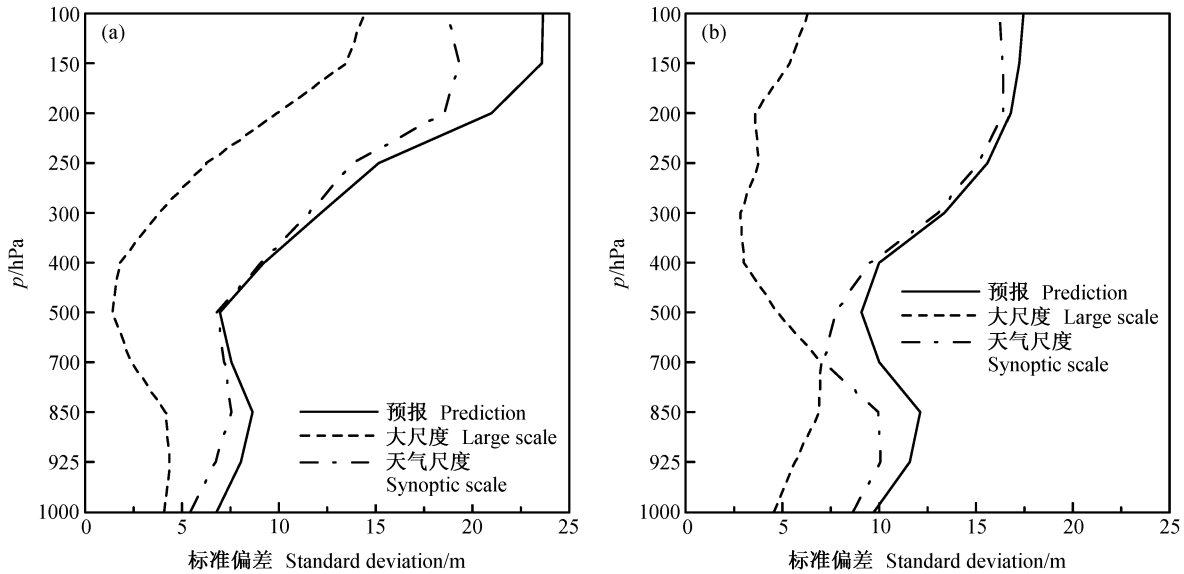

图 6　位势高度天气尺度、大尺度和总预报误差随高度的变化：(a) 0000 UTC；(b) 1200 UTC

Fig. 6　Synoptic scale, large scale and total forecast errors: (a) 0000 UTC; (b) 1200 UTC

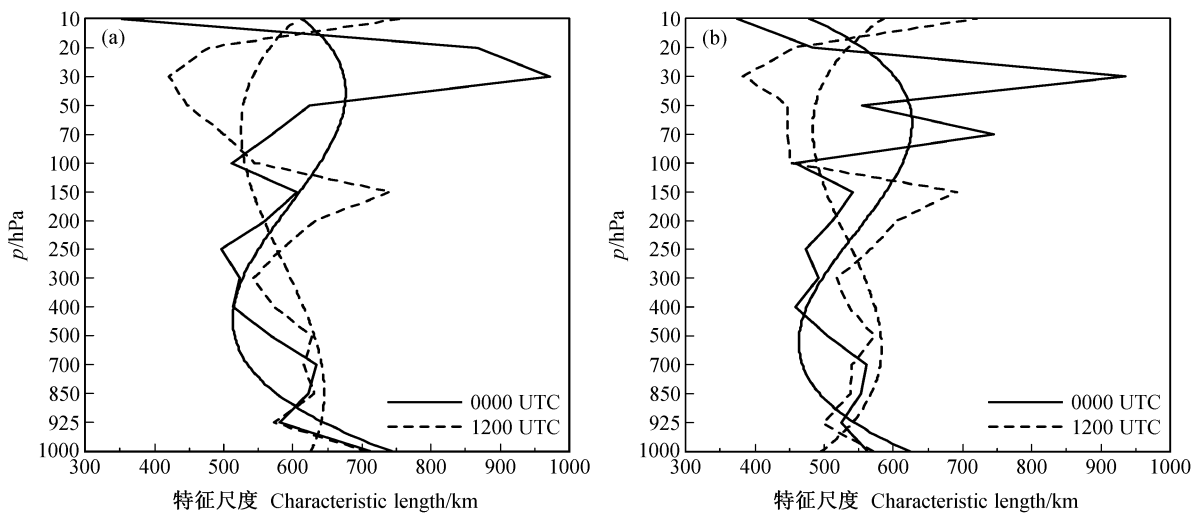

图 7　水平特征尺度的变化及三次方拟合：(a)预报误差的水平特征尺度变化；(b)天气尺度预报误差的水平特征尺度的变化

Fig. 7　Characteristic lengths for the levels and cubic approximations: (a) For total forecast error; (b) for synoptic scale forecast errors

尔函数拟合得到的位势高度预报误差在 100 hPa 以下在 5～15 m 之间，观测误差与预报误差大小差不多。可以发现三种不同拟合函数得到的背景误差大小相差不大，有限区域数值预报业务系统可能对背景误差估计过高。

位势高度预报误差协方差样本在 1000 km 以外逐渐趋近于常数，为大尺度误差部分。从图 6 可以看出天气尺度预报误差占预报误差的主要部分。在 300 hPa 以上，0000 UTC 大尺度预报误差有所增加，而在 1200 UTC 大尺度预报误差随高度变化不大。

4.2　背景场误差的水平特征尺度

对同一层的背景误差协方差样本点进行高斯函数拟合后，还可以得到水平特征尺度。在两维均匀各向同性的条件下，水平特征尺度定义为

$$L^2 = -\left.\frac{2\rho_B}{\nabla^2\rho_B}\right|_{r=0},$$

其中，ρ_B 为背景误差相关系数。水平特征尺度表征背景误差的影响范围，反映背景误差协方差的空间特征。

图 7 为背景误差协方差的水平特征尺度随高度

的变化，较平滑的曲线是对水平特征尺度三次方拟合的结果。从图 7a 可以看出，在对流层，水平特征尺度的大小在 500～750 km 之间，在平流层 0000 UTC 水平特征尺度较大，这与 Bruce[17] 得到的结论相一致。由水平特征尺度三次方拟合曲线可以看出，0000 UTC 的水平特征尺度在对流层底较大，越往高层有所减少，在 500 hPa 左右，减小到最小，然后逐渐增大，在平流层 50 hPa 附近出现最大值。1200 UTC 的水平特征尺度在对流层底部较大，越往高层，逐渐减小，在 50 hPa 附近减到最小。从图 7b 可以看出，在对流层，天气尺度预报误差的水平特征尺度大小在 450～650 km 之间，比总预报误差水平特征尺度小，说明大尺度预报误差影响的范围比中尺度预报误差影响范围大。Franke[13] 用二阶自回归相关函数拟合背景误差协方差样本点所得到的水平特征尺度在 250 km 和 700 km 之间，水平特征尺度随高度变化趋势和本文统计的差不多。

4.3 背景场误差的垂直相关模型

在 GRAPES 3D-Var 中，背景误差垂直相关由经验函数获得，本文参考 Franke[13] 使用的二阶自回归型相关函数近似垂直误差相关，试图用统计得到的数据获得背景误差垂直相关模型。

从以上几节分析可以得到高度层间的背景误差协方差矩阵 $\boldsymbol{C}_{lk}^{\mathrm{B}}$ 和垂直观测误差协方差矩阵 $\boldsymbol{C}_{lk}^{\mathrm{O}}$，因而可以获得垂直相关系数矩阵。图 8 为 0000 UTC 垂直误差相关系数样本点的分布，垂直距离为高度层间的距离取自然对数。可以看出两高度层间距离为零的垂直背景误差相关系数接近于 1，两高度层相距越远，相关系数越小。从图 8b 可以看出，垂直观测误差相关系数分布随着高度层间距离的增大而比较零散，然而可以看出其分布形势类似于垂直背景误差相关系数。

在 GRAPES 3DVar 中垂直误差相关函数为

$$\rho_{lk}=\frac{1.0}{1.0+k_p(\ln p_l-\ln p_k)^2}, \tag{8}$$

其中，k_p 为分析变量数目，$\ln p$ 为气压的自然对数。本文用二阶自回归型相关函数近似垂直误差相关：

$$\rho_{lk}(p)=(1-|a|)(1+|\ln p_l-\ln p_k|)\cdot \mathrm{e}^{-|\ln p_l-\ln p_k|}+|a|, \tag{9}$$

其中，$\rho_{lk}(p)$ 为随气压层间的距离而变化的垂直相关系数，p_l 为第 l 层的气压值，p_k 为第 k 层的气压值，a 为参数。如果按(9)式直接对垂直误差相关系数进行拟合，在高度层间距离远的区域，样本分布比较零散，则拟合的效果就不是很好。采用垂直距离变换函数 $h_l=h(\ln p_l)$，则垂直背景误差相关近似函数为

$$\rho_{lk}(h)=(1-|a|)(1+|h_l-h_k|)\mathrm{e}^{-|h_l-h_k|}+|a|, \tag{10}$$

式中 h 与 a 都是未知的参数，而且 h 是 $\ln p$ 的函数，$h=h(\ln p)$。通过对垂直误差相关样本进行拟合，用最小二乘法可以得到垂直距离变换公式 $h_l=h(\ln p_l)$ 和参数 a。由于距离变换公式中函数 h 的具体形式不知道，用最小二乘法得到 h_l 是标准气压层 p_l 对应的距离变换值。

图 9a 为垂直距离变换后，垂直背景误差相关系数样本分布及其拟合函数曲线，可以看出通过垂直

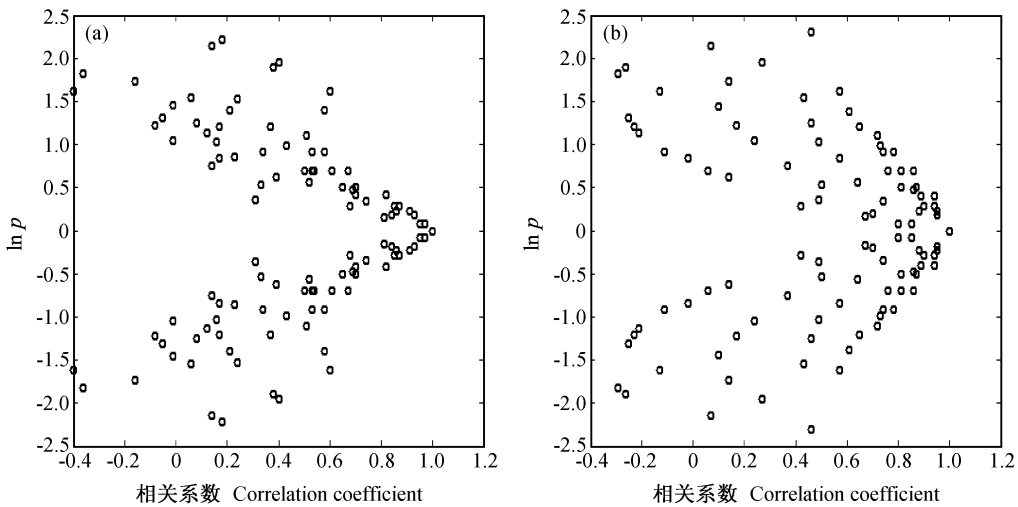

图 8 0000 UTC 垂直误差相关样本点随 $\ln p$ 距离分布：(a) 预报；(b) 观测

Fig. 8 Distribution of vertical error correlation with $\ln p$ distance at 0000 UTC: (a) Prediction; (b) observation

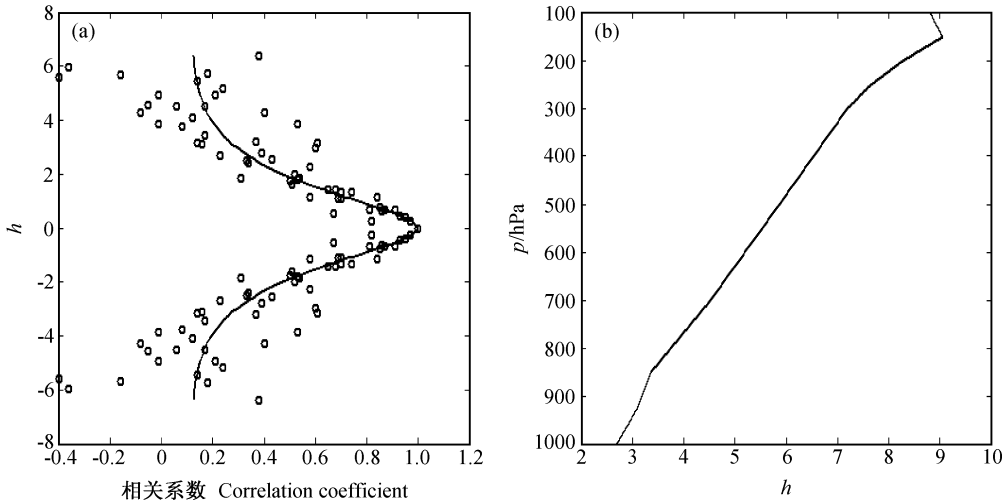

图 9　0000 UTC 垂直距离变换后垂直背景误差相关：(a)误差相关样本分布及拟合；(b)垂直距离变换函数

Fig. 9　Vertical correlation of prediction error in transformed coordinate distance at 0000 UTC: (a) Correlation and fitting; (b) distance transformation

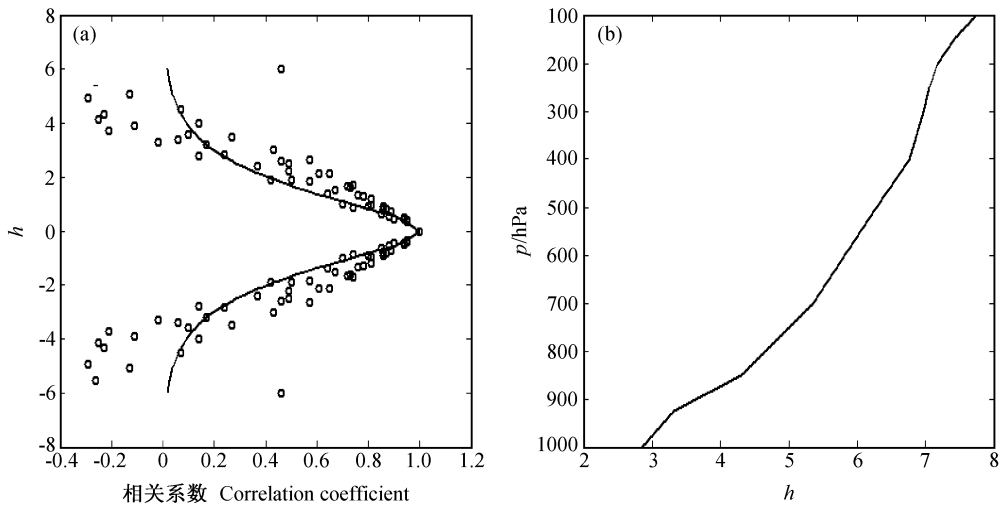

图 10　0000 UTC 垂直距离变换后垂直观测误差相关：(a)误差相关样本分布及拟合；(b)垂直距离变换函数

Fig. 10　Vertical correlation of observation error in transformed coordinate distance at 0000 UTC: (a) Correlation and fitting; (b) distance transformation

距离变换后的样本分布较垂直距离变换前集中，拟合的效果也较好。图 9b 为标准气压层与垂直距离变换函数的对应关系。随着高度的增加，垂直距离变换函数 h 是增加的，但并不是单纯的线性增加。

从图 10a 可以看出，经过垂直距离变换后，0000 UTC 垂直观测误差相关样本分布有了较大的改变，样本分布明显变得集中，垂直相关函数的拟合效果较好。垂直距离变换函数随着高度层的增加而逐渐增大，但与垂直背景误差相关的垂直距离变换函数有所不同。

在 GRAPES 三维变分同化系统中，对背景误差协方差作垂直变换分析时，需要对背景误差协方差矩阵作 EOF 分解。本文取 11×11 背景误差协方差矩阵，作 EOF 分解(如图 11 所示)。图 11a 所示的 11 个特征值中前 4 个特征值最大，其他特征值较小。前 4 个特征值的特征向量如图 11b 所示。

利用 EOF 分解对背景误差协方差进行垂直变换，可以用较少的指标代替原来较多的指标，从而减小计算量。从特征值的大小排列，可以看出主成分的方差贡献，从累积解释方差可以看出主成分的

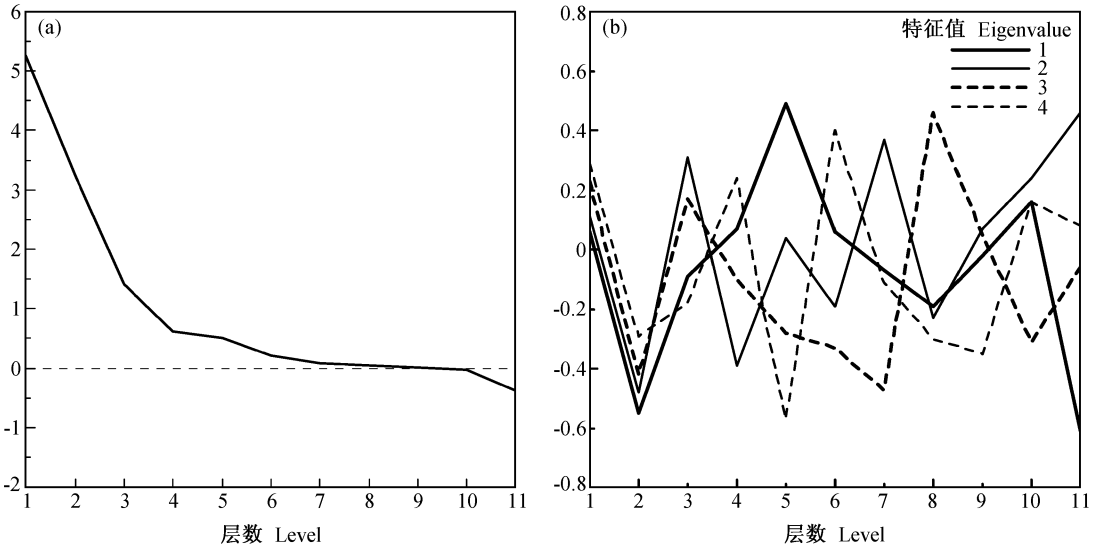

图 11 1200 UTC 背景误差协方差的特征值（a）和各特征值所对应的特征向量（b）

Fig. 11 (a) Eigenvalue of background error covariance at 1200 UTC; (b) each eigenvector with eigenvalue at 1200 UTC

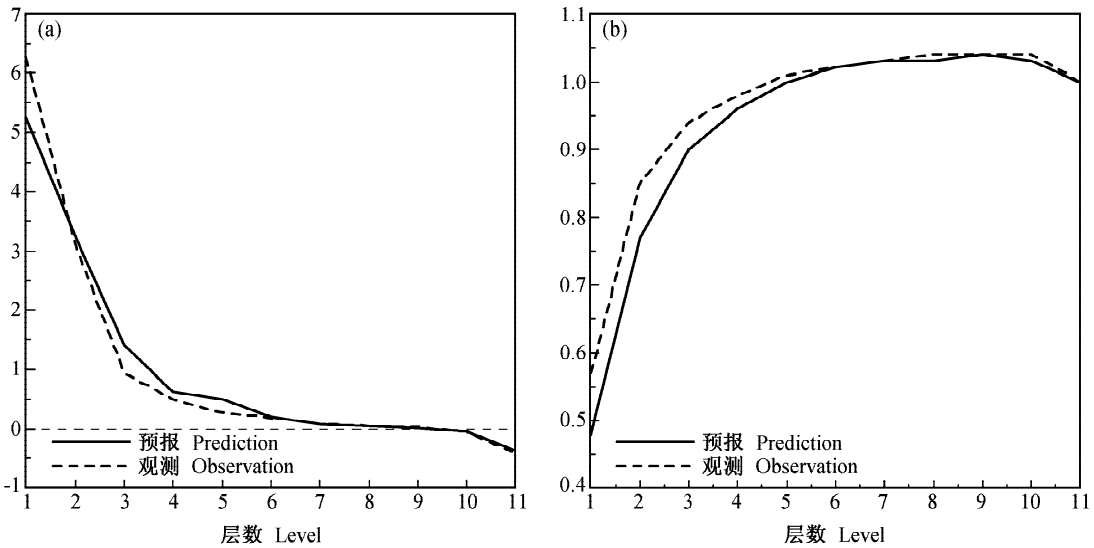

图 12 1200 UTC 背景与观测误差协方差的特征值（a）和累积解释方差（b）

Fig. 12 (a) Eigenvalue of covariance of prediction and observation errors at 1200 UTC; (b) cumulative variance at 1200 UTC

贡献。若共有 n 个主分量，取前 m 个主分量，则其累积解释方差的为

$$G(m)=\frac{\sum_{i=1}^{m}\lambda_i}{\sum_{i=1}^{n}\lambda_i},$$

λ_i 为第 i 个特征值。从图 12a 1200 UTC 的特征值可以看出，背景误差协方差和观测误差协方差前 6 个特征值较大，第 6 个特征值以后特征值就趋向零。从图 12b 可以看出 $m>4$ 时，其累积解释方差的贡献率已超过 90%。当 $m=6$ 时，累积解释方差的贡献率已接近于 100%。

4.4 三维背景误差协方差相关模型

从前面的分析可以得到水平和垂直方向上的背景误差相关近似函数。在前几节的基础上，综合考虑水平和垂直方向上的背景误差，可以得到三维背景误差协方差模型。

在垂直方向，采用二阶自回归型垂直误差相关模型，由(10)式，各层间的垂直背景误差协方差近似函数为

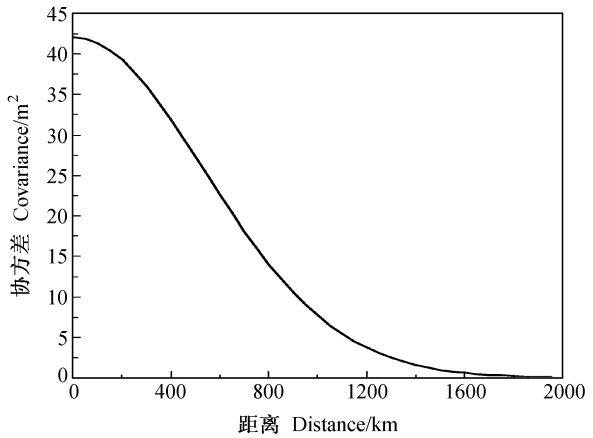

图 13　0000 UTC 500～400 hPa 背景误差协方差近似函数

Fig. 13　Background error covariance approximation at 500 -400 hPa at 0000 UTC

$$\sigma_l[(1-|a|)(1+|h_l-h_k|)\mathrm{e}^{-|h_l-h_k|}+|a|]\sigma_k, \tag{11}$$

其中，σ_l、σ_k 分别为第 l 层和第 k 层的背景误差均方差。再由厚度误差协方差与高度误差协方差的关系，得到厚度误差协方差函数：

$$D_{lk}=\sigma_l^2+\sigma_k^2-2\sigma_l\sigma_k[(1-|a|)(1+|h_l-h_k|)\mathrm{e}^{-|h_l-h_k|}+|a|]. \tag{12}$$

最后，得到高斯型和二阶自回归型组合的三维空间的背景误差协方差的结构函数：

$$C_{lk}(d,\delta h)=\frac{1}{2}\{\sigma_l^2\mathrm{e}^{-d^2/2L_{ll}^2}+\sigma_k^2\mathrm{e}^{-d^2/2L_{kk}^2}-[\sigma_l^2+\sigma_k^2-2\sigma_l\sigma_k((1-|a|)(1+\delta h)\mathrm{e}^{-\delta h}+|a|)]\mathrm{e}^{-d^2/2L_{lk}^2}\}, \tag{13}$$

可以看出空间的背景误差协方差是高度距离 δh 和水平距离 d 的函数。

从 500 hPa 与 400 hPa 背景误差协方差近似函数曲线的分布(图 13)可以看出，与图 4b 比较，近似函数能反映出背景误差协方差样本的分布状况。

5　小结

本文主要利用 T213 预报资料和无线电探空观测资料，用观测余差方法研究我国区域范围内的背景误差协方差的结构特征，通过以上的研究，主要得到以下结论：

(1) 在观测资料较多的情况下，用观测余差方法统计背景误差协方差是可行的。在均匀各向同性的假设下，单层背景误差协方差的水平分布近似高斯型分布，用高斯型相关函数拟合效果较好。

(2) 对于多层间的背景误差协方差的研究，由于厚度方法比直接方法使用的信息多，用厚度方法得到的高度误差协方差函数更能准确地反映实际情况。

(3) 在对流层，预报误差的水平特征尺度的大小在 500～750 km 之间，天气尺度预报误差的水平特征尺度在 450～650 km 之间。说明大尺度预报误差影响的范围较广。在平流层 0000 UTC 的水平特征尺度较大。

(4) 背景误差和观测误差随着高度逐渐增加，在对流层低层观测误差小于背景误差，在对流层中高层，观测误差大于背景误差。统计得到的预报误差包括天气尺度误差和大尺度误差，其中天气尺度误差占预报误差的主要部分。

(5) 垂直背景误差协方差可以利用垂直距离变换用二阶自回归函数来近似。对垂直背景误差协方差进行 EOF 分解，可以减少其在三维变分同化系统的计算量。三维背景误差协方差的结构可以用高斯型和二阶自回归型组合函数来近似。

(6) 在背景误差协方差的统计过程中，拟合函数权重的选择对拟合结果有一定的影响。由于观测站点分布不均匀和测站资料数量的不同，权重主要和一定区域范围内的样本数量和样本间的距离有关。文中权重的设计和权重的选择具有一定的主观性，需要进一步研究。

本文在均匀、各向同性的假设基础上得到背景误差协方差的结构，只有在一定条件下，背景误差协方差接近各向同性，而实际的背景误差协方差是随天气形势而变化的，非各向同性的[18～20]。因而，如何得到随天气形势变化、更合理的背景误差协方差还需要进一步研究。

参考文献 (References)

[1] Bergthorsson P, Doos B. Numerical weather map analysis. *Tellus*, 1955, **7**: 329～340

[2] Cressman G P. An operational objective analysis system. *Mon. Wea. Rev.*, 1959, **87**: 367～374

[3] Sasaki Y. An objective analysis for determining initial conditions for the primitive equations. Tech. Rep., 1960, (Ref. 60-16T)

[4] Barnes S L. A technique for maximizing details in numerical weather map analysis. *J. Appl. Meteor.*, 1964, **3**: 396～409

[5] Barnes S L. Oklahoma thunderstorms on 29－30 April 1970. Part I: Morphology of a tornadic storm. *Mon. Wea. Rev.*, 1978, **106**: 673～684

[6] Roger Daley. *Atmospheric Data Analysis*. New York: Cambridge University Press, 1991. 105～120

[7] Gustafsson N, Berre L, Hornquist S, et al. Three-dimensional variational data assimilation for a limited area model. *Tellus*, 2001, **53A**: 425～446

[8] 邱崇践. 变分同化中使用背景场时尺度匹配的数值研究. 大气科学, 2001, **25**(1): 103～110
Qiu Chongjian. Matching scales of the retrieved variables with background scale in variational assimilation: a numerical study. *Chinese Journal of Atmospheric Sciences* (in Chinese), 2001, **25**(1): 103～110

[9] Drozdov O, Shepelevskii A. The theory of interpolation in a stochastic field of meteorological elements and its application to meteorological elements and its application to meteorological map and network rationalization problems. Trudy Niu Gugms Series 1, 1946, No. 13

[10] Rutherford I D. Data assimilation by statistical interpolation of forecast error fields. *J. Atmos. Sci.*, 1972, **29**: 809～815

[11] Hollingsworth A, Lönnberg P. The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field. *Tellus*, 1986, **38A**: 111～136

[12] Lönnberg P, Hollingsworth A. The statistical structure of short-range forecast errors as determined from radiosonde data. Part II: The covariance of height and wind errors. *Tellus*, 1986, **38A**: 137～161

[13] Franke R. Three-dimensional covariance functions for NOGAPS Data. *Mon. Wea. Rev.*, 1999, **127**: 2293～2308

[14] Xu Qin, Li Wei, Andrew Van Tuyl, et al. Estimation of three-dimensional error covariances. Part I: Analysis of height innovation vectors. *Mon. Wea. Rev*, 2001, **129**: 2126～2135

[15] Parrish D F, Derber J C. The national meteorological center's spectral statistical-interpolation analysis system. *Mon. Wea. Rev.*, 1992, **120**: 1747～1763

[16] Barker D M, Huang W, Guo Y-R, et al. A three-dimensional variational data assimilation system for MM5: Implementation and initial results. *Mon. Wea. Rev.*, 2004, **132** (4): 897～914

[17] Ingleby N B. The statistical structure of forecast errors and its representation in The Met. Office Global 3-D Variational Data Assimilation Scheme. *Quart. J. Roy. Meteor. Soc.*, 2001, **127**: 209～231

[18] Seaman R S. A systematic description of the spatial variability of geopotential and temperature in the Australian region. *Aust. Meteor. Mag.*, 1982, **30**: 133～141

[19] Benjamin S G. An isentropic meso α-scale analysis system and its sensitivity to aircraft and surface observations. *Mon. Wea. Rev.*, 1989, **117**: 1586～1603

[20] RiishØjgaard L P. A direct way of specifying flow-dependent background error correlations for meteorological analysis systems. *Tellus*, 1998, **51A**: 42～57