

全球资料同化中误差协方差三维结构的 准确估计与应用 : 背景误差协方差调整与数值试验分析*

龚建东 赵 刚

国家气象中心,北京,100081

摘 要

利用 NMC 方法针对背景误差协方差的方差、三维相关与特征长度来揭示 T213L31 模式的误差主要特点,并与传统更新矢量方法的计算结果进行了对比与调整。结果表明 NMC 方法结果与更新矢量方法结果在大体特征上基本吻合,但细节上的差异不可忽视,特别是对背景误差方差与特征长度的估计存在显著的差异,其主要原因是 NMC 方法倾向于高估天气尺度波的背景误差,而低估次天气尺度到中尺度波的背景误差。通过对背景误差方差、特征长度的调整,显著改善了背景误差功率谱的分布特点,使得 NMC 方法结果与更新矢量方法结果更为吻合。通过三维变分同化与最优插值中观测与背景误差相对重要性的比较,发现两者结果基本一致,但三维变分同化在 850 hPa 以下的温度场和 300 hPa 以上的风场统计结果都表现出背景误差相对于观测误差偏小的特点。背景误差相对于观测误差偏小有助于保证分析场中质量场与风场平衡,消除了大气底层和高层质量场与风场不匹配现象。在数值试验中,针对不同的背景误差均方差与特征长度的特点,分析了分析增量和预报效果的差异,结果表明,准确的背景误差估计与优化工作改善了预报效果,使得北半球三维变分同化的 120 h 预报效果整体好于现有最优插值。

关键词:背景误差协方差, NMC 方法估计,更新矢量结果调整,数值对比试验。

1 引 言

全球资料同化从最优插值向三维变分升级过程中,首先要解决对现有业务可用资料两者的同化分析与预报技巧相当,这样当卫星 ATOVS 等资料直接同化后预报效果才能进一步改善。在采用国外业务三维变分同化后,由于全球中期预报模式、观测资料的使用量在不同业务中心相差甚远,需要发展基于现有观测资料量与分布特点、适合 T213L31 全球中期预报模式的观测误差与背景误差协方差,而目前中国针对全球范围的误差估计工作并不多见。

利用在同一时刻两个不同时效的预报偏差场来描述背景误差的增长,以这些偏差场作为样本可以估计背景误差协方差^[1](简称 NMC 方法)。这种方

法的优点是可以提供全球同化所需的所有波谱和模式垂直层次背景误差协方差特征,统计出不同变量间的平衡与约束关系。全球模式中一般用同一时刻 48 h 预报减 24 h 预报乘以一个经验系数来缩小该误差幅度去表征 6 h 的背景场误差^[2]。NMC 方法的主要缺陷在于一些经验系数与参数的确定上,如误差方差与相关特征长度等都不太准确,需要在实际使用时进行调整。由于背景误差协方差三维空间结构特征与数值模式分辨率和观测网密切相关,不同的数值模式、不同的分辨率,背景场误差的特征差别较大。更新矢量方法^[3-7](Innovation Vector,简称 IV 方法)在观测系统比较均一、局地均质、各向同性假设条件下,利用贝塞尔函数拟合分析方法可以正确地分离背景场误差与观测误差,并决定出空

* 初稿时间:2005 年 9 月 27 日;修改稿时间:2006 年 7 月 10 日。

资助课题:国家自然科学基金(40305015),国家气象中心自筹资金课题“三维变分同化阶段性突出问题研究”(04ZK2003A-09)。

作者简介:龚建东,男,1971 年生,博士,副研究员。主要从事资料同化与集合预报方法研究。E-mail: gongjd@cma.gov.cn

间相关形状与特征长度。以 NMC 方法估计结果为主、利用 IV 方法进行参数调整是对背景误差协方差静态准确估计的重要途径。此外,现有业务最优插值提供了观测误差和背景误差的特征,观测误差与背景误差的相对大小决定了观测信息与背景信息的相对重要性,而特征长度决定了观测信息的影响范围,需要将这些参数和三维变分同化的参数进行比对来明确差异,这将有助于分析随后同化和预报的表现。

2 背景误差协方差分析

2.1 三维背景误差协方差计算

中国国家气象中心的 T213L31 全球中期数值预报业务模式是静力平衡水平三角形截断的谱模式,采用涡度()、散度()、温度(T)和地面气压对数($\ln p_s$)与比湿场(q)作为大气基本状态变量。对这些状态变量分别计算其在谱空间与格点空间的方差分布特点,就可以较全面地揭示背景误差方差的特征。在文中尽量采用 Ide^[8]的标准符号。在谱空间中 Derber、Courtier 等^[2,9]建议以每个二维总波数为单元来计算背景误差方差,全球平均总方差是所有二维总波数方差的和。当谱空间中每个二维总波数中的每个波都具有相同方差时,则在水平格点空间的协方差具有各向同性特点。任意两点 P, Q 的协方差可表示为

$$PQ = \sum_{n=0}^{n_j} b_n (2n+1) P_n(\cos \theta) = f(\theta) \quad (1)$$

其中 n_j 为三角形截断波数, $2n+1$ 为总波数为 n 的所有二维波的个数(也即经向波 m 的个数), b_n 为总波数为 n 的单个二维波方差,不依赖经向波变化。 θ 为两点间相对于球心的夹角, P_n 为归一化后

的勒让德多项式。显然,格点空间协方差随距离的变化形式可由不同总波数 b_n 的方差唯一确定。通过考察 b_n 的分布特点,就可以计算水平空间相关形式。在格点空间计算方差要简单一些,直接取成纬向平均值。

由二维波的方差计算对应的特征长度,依据 Daley^[10]特征长度的定义表示为

$$L_s^2 = -2 \frac{f(0)}{\nabla^2 f(0)} = 2a^2 \frac{b_n}{b_n n(n+1)} \quad (2)$$

其中 a 为地球半径, ∇^2 为拉普拉斯算子。

利用在同一时刻 T213L31 模式 24 和 48 h 两个不同时效的预报偏差场估计背景误差协方差参数^[11-13]。样本从 1 月 30 日至 4 月 19 日共 80 个。取 24 h 模式比较协调,可以减弱最优插值分析以及 spir-up 震荡的影响,取 48 h 是考虑到预报差异足够大。背景误差因其原始性对季节的敏感性很小。按准地转线性平衡的需求从 $-p$ 混合坐标插值到 σ 坐标,且顶层气压从 10 hPa 外推到 2 hPa 以便考虑随后卫星 ATOVS 资料同化,如涡度误差向量定义为

$$e(n, m, j, s) = 48(n, m, j, s) - 24(n, m, j, s) \quad (3)$$

其中 j 是 σ 坐标的层次指数, s 是个例数或样本数。

2.2 标准差的空间变化

从背景误差标准差的水平分布特征来看,由于现有最优插值中主要用到常规观测资料和卫星反演产品(云导风、测厚资料等),陆地上的资料量要明显多于海洋上,北半球明显多于南半球,图 1a 给出最优插值中由观测资料实时计算得到的 500 hPa 温度背景误差,呈现出敏感的依赖于观测分布的特点。NMC 方法的结果基本与最优插值相符(图 1b),南

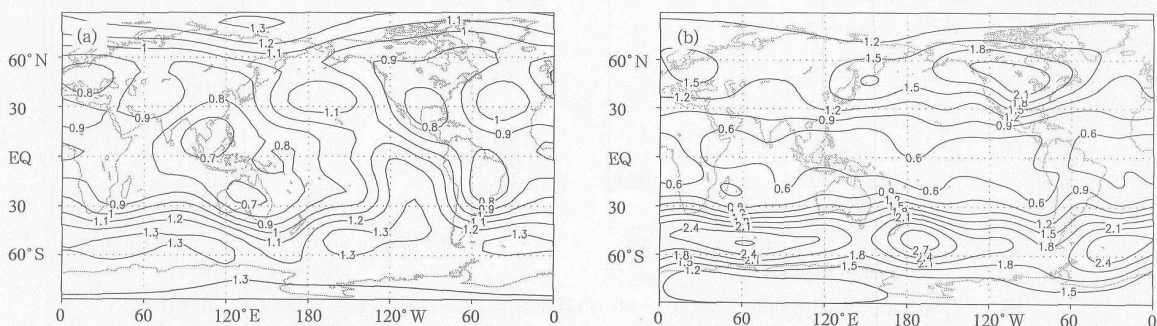


图 1 业务最优插值计算得到的 500 hPa 温度背景误差与 NMC 方法估计值比较(单位: K)

(a. 业务最优插值计算标准差, b. NMC 方法估计结果)

Fig. 1 500 hPa temperature background error standard deviation by OI (a) and by NMC (b) method (units: K)

半球的误差要大于北半球,赤道地区最小,分布与观测有关但有较大差异,从细节上仍然与实际情况有出入。如北半球的误差大值区在北美、东亚、欧洲区,而这些区域都是常规观测,特别是探空资料非常密集的区域,背景误差应当比较小,而 NMC 方法统计的结果显示北美大陆的误差均方差最大,这种现象的原因还不能确定。另外南北半球的极区是观测资料最稀少的位置,背景误差应当比较大,而统计结果要比中纬度小一些。热带地区的误差最小,这和热带地区的大气混沌特性有关,该区域的大气变率比较小,系统维持的时间较短,且观测资料稀少。NMC 方法整体上呈现出中高纬度比最优插值估计大,热带区比最优插值估计小的特点。此外 NMC 估计结果呈现出明显的随纬度变化特点,在实际应用中可以直接取成纬向平均。

在不同垂直层次整体上半球误差大于北半球(图略)。涡度误差高值区出现在 300—400 hPa 的中纬度地区,而在热带上空数值非常小。风场与涡度对应的非常好,但风场大值中心位置略偏高大约在 300 hPa。散度除了在中纬度地区有高值区外,在热带上空 250 hPa 位置出现另一个高值区,这与涡度显著不同。温度高值区主要在中高纬区域,从 850 hPa 一直延伸到对流层顶,但在热带上空

70 hPa 的位置也有一个高值区,这与热带地区对流层顶加热有关。地面气压的特点与温度非常一致。水汽主要在中低层 850 hPa 的位置出现高值区,方差主要集中在 60°S—60°N,在 300 hPa 以上方差基本为零。

与观测空间 IV 方法估计结果进行比对(图 2),对风场而言,由两种方法分别估计出的涡度与标量风这两条廓线之间的比例关系非常一致。可以发现由 NMC 方法估计的风场均方差廓线(取北半球 20°—90°N 平均,基本上是 IV 方法代表的范围)在 400—300 hPa 达到最大值,为 5.5 m/s,比观测空间在 250 hPa 左右的结果明显偏低,但均方差偏大。而温度均方差最大值出现在 200 hPa 左右,为 1.5 K,与观测空间结果非常一致,同样均方差数值偏大。从 150 hPa 往上两种估计方法的均方差结果比较接近,分别约为 2 m/s, 0.9 K。但在 200 hPa 以下的底层, NMC 方法的均方差要明显大于观测空间的结果,且在 500 hPa 左右温度场的差异最大。

从整层倾向的特点来看 NMC 方法的结果从下往上呈现出向低值方向倾斜的特点,而观测空间结果则相反,呈现出向高值方向倾斜的特点。IV 方法结果的可靠性依赖于其无偏、均质、各项同性假设的成立程度以及在代表北半球时考虑到 3 个区域估计

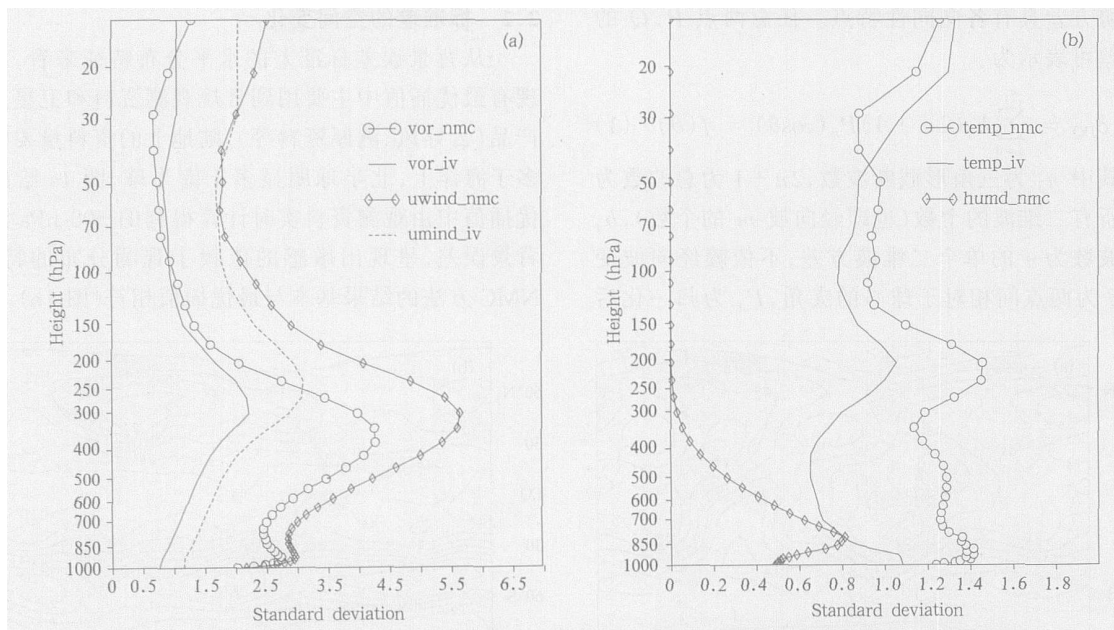


图 2 NMC 方法估计出的北半球(20°—90°N)背景误差均方差垂直平均廓线与 IV 结果比较

(a. 涡度,单位: $10^5 \times \text{s}^{-1}$;风场, b. 比湿,单位: (10^3 g/kg) ;温度)

Fig. 2 Background error standard deviation 20°-90°N mean profiles by the NMC and IV methods for

(a) vorticity (unit: $10^5 \times \text{s}^{-1}$) and scale-wind, (b) specific humidity (unit: 10^3 g/kg) and temperature

结果的差异,这种差异在 200 hPa 以下变率为 11%,而在 200 hPa 以上为 20%左右,但这种变率不足以影响 IV 廓线倾向^[14]。IV 结果与 NMC 结果的均方差偏差在高低层特点不一致,整层上不是简单的倍数缩放关系。这种差异可能和两方面的因素有关,一是平流层以大尺度运动为主,用同一时刻 24 和 48 h 预报时效的偏差来模拟 6 h 背景误差会偏小,而 200 hPa 以下小尺度运动活跃,模拟背景误差会偏大。图 1 已经指出对热带地区由于大气混沌特性变率较小,背景误差会偏小,可见 NMC 方法的可靠性依赖于用同一时刻两个预报时效偏差模拟 6 h 背景误差的成立程度,以上分析表明利用 NMC 方法来估计背景误差协方差时有诸多缺陷,这一点在使用时要谨慎。目前国际上普遍采用对 NMC 方法获得的 24 与 48 h 预报偏差的均方差缩小至 0.9 倍以描述 6 h 的方差^[2]。从以上分析看这是一种折中方法,这种处理方法在 200 hPa 以下的低层使两种方法估计的廓线接近,而在 150 hPa 以上的高层使两者偏差扩大,或者背景误差更偏小。比湿场均方差在低层 850—700 hPa 达到最大值,并随着高度增加迅速减小,在 300 hPa 以上趋于零,这与事实基本相符。

2.3 相关系数的变化特点

背景误差特征长度随高度呈现逐步增长的趋势(图略)。温度特征长度从地面到 500 hPa 逐步增加,从 200 km 到 300 km,再往上特征长度基本维持在 300 km。风场的变化明显呈现两个部分,在西风急流以下层次略大于 200 km,西风急流以上基本维持在 300 km,跨越西风急流层次之间的特征长度变化比较快。与此相反,涡度和散度变化一直比较稳定,在 100 km 左右均匀增加,在西风急流区没有明显的过渡,相对而言涡度比散度的特征长度要略长一些。比湿的特征长度只比涡度略长一些,最长的特征长度是地面气压,达到 470 km。

与观测空间 IV 方法估计结果进行比对,可以发现在垂直方向上对涡度的估计与 NMC 方法非常一致,但温度的特征长度约为 230 km,远小于 NMC 方法的 300 km 左右,整体偏小约 70 km。标量风场间的差别在模式低层相差不大,但随着高度增加差别迅速增大,主要表现在 IV 方法变化不大而 NMC 方法迅速增加,在 200 hPa 以上差别在 100 km 左右。

选取 500 hPa 高度分析相关系数的变化(图

略),可以发现涡度、散度的特征长度都比较短且与零轴相交,然后无限逼近零值。散度的这种变化与 Daley^[10]的结论比较一致,但涡度与现有理论不符,理论上是从正值无限逼近零值。进一步研究发现这种现象在温度、地面气压对数、标量风场、比湿中都一定程度上存在。

2.4 误差协方差的特征

误差的垂直协方差总体上呈现出两部分(图 3),在 100 hPa 以下相关的范围比较宽广,而在 100 hPa 以上相对于低层相关范围比较狭窄。对不同的状态变量垂直相关特点差异比较大。涡度在不同层次都呈现出正相关,且相关的范围比较宽广,而散度要窄得多,并出现明显的负正负结构。这与大气的特点有关,大尺度大气运动整体上以旋转风为主,在垂直方向上呈现出较好的一致性,但在某层辐散/辐合(正/负值)则往往会在其上下层形成辐合/辐散(负/正值)结构。标量风的特点与涡度接近,但相关范围更宽广,这表明在垂直一致性上风场的特点要显著好于涡度场,但在 100 hPa 以上标量风与涡度一致,相关范围都比较窄。温度呈现出明显的负相关,相关系数比较大,达到 -0.5,且范围比较宽广。合理的解释是在对流层温度随高度递减,在对流层顶与同温层交界处达到最小,在同温层温度随高度递增。这样在对流层顶与同温层交界处的温度垂直变率比较大且变率范围比较窄,出现比较陡的温度低谷。不同时刻的预报其对流层顶与同温层交界处的高度有差异,使得两者温度低谷更容易呈现出位相上的差异,这种差异将会造成交界处的上下层呈现出负相关的特点。地面气压对数与温度的相关在 500 hPa 以下和 300 hPa 以上呈现出明显的负相关,最大负相关可以达到 -0.6,但显然 500 hPa 左右层次与地面气压对数的相关并不好,可以认为基本不相关。比湿的相关范围在不同层次都比较窄,而且在不同层次呈现出良好的一致性,这种特点可能是由于水汽在不同高度变化较快,水汽输送在不同层次差异较大(图略)。

对比文献[14]的图 7,可以发现 NMC 方法估计出来的垂直协相关特点与 IV 方法结果基本可比,都是标量风、涡度相关范围最宽,散度最窄且为负正负结构。温度相关范围略窄并都在 150 hPa 出现负相关,但 NMC 结果的负相关幅度明显强于 IV 方法的结果,后者只有 -0.2 左右。NMC 方法结果的相

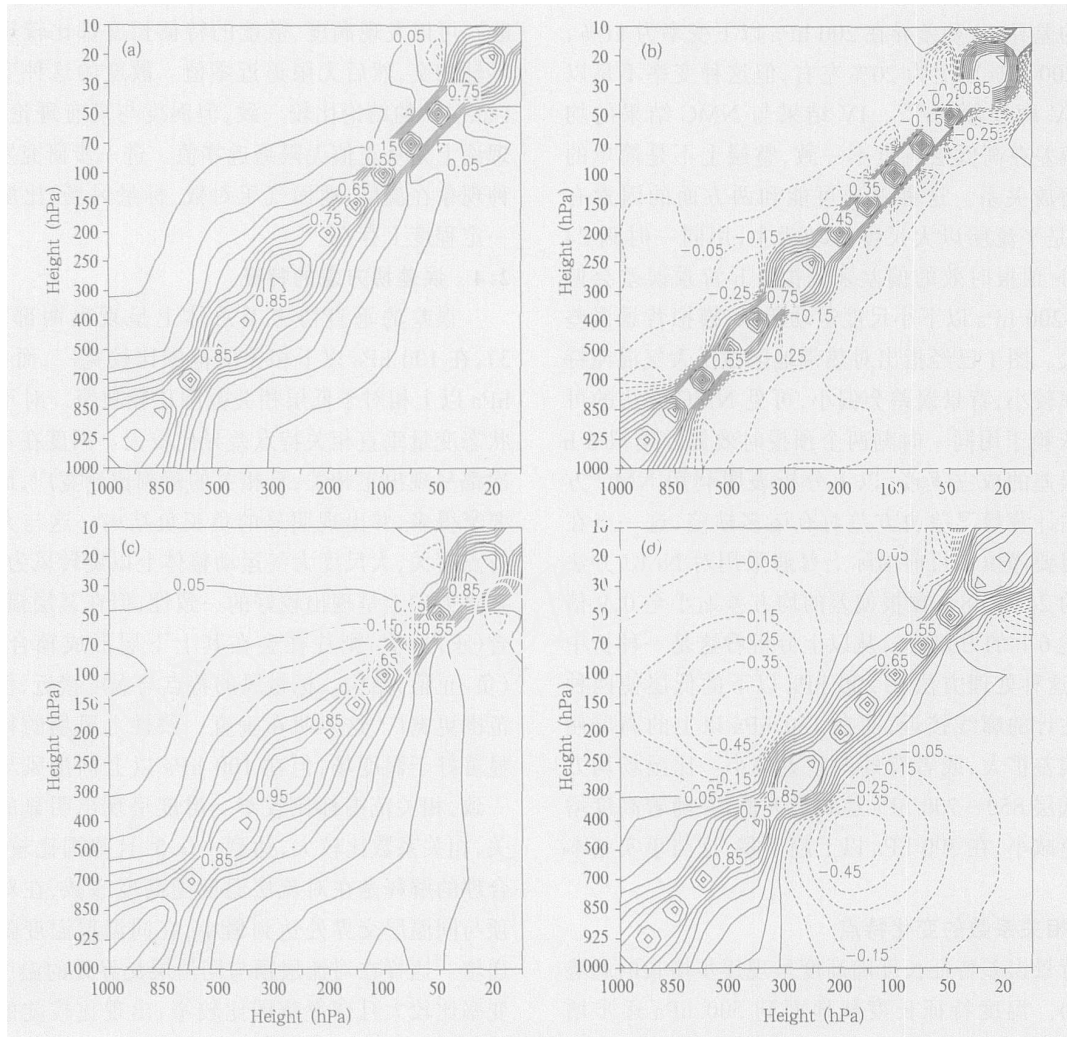


图3 NMC方法估计出的背景误差垂直协相关系数
(a. 涡度, b. 散度, c. 标量风, d. 温度)

Fig. 3 Vertical correlations of background errors by NMC method for (a) vorticity, (b) divergence, (c) scale-wind, and (d) temperature

关范围比 IV 方法结果更为宽广一些是两者间的主要差别。

从不同的二维波数考察垂直协方差的功率谱变化特点(图4),计算500 hPa与其他各层的垂直相关系数。从总体上看随着二维波数变化,垂直相关的特点基本一致,如涡度、比湿、标量风以正相关为主,散度、温度、地面气压对数与温度以负正负结构为主。可以看到0—4波垂直相关比较窄,而5—20波垂直相关最为宽广,10波左右达到最大值,如涡度与1000 hPa相关也能达到0.7。10波以后垂直相关范围逐步缩小并趋于定值,因此仅给出前100波的结果。这表明天气尺度波仍然是误差垂直相关的主体部分。散度的垂直相关随波数没有明显的变

化,都比较窄,且负相关的范围非常宽广,在所有波上都能维持相当大的值,说明辐合辐散在所有波长上都重要。温度的负相关区随波上变化比较明显,大值区主要集中在0—40波,最大负相关达到-0.6。对比文献[14]的图8,可以发现两种方法结果基本可比,都是风场、涡度相关范围比较宽广,散度与温度比较窄且有负相关结构,但NMC结果的相关范围要比IV结果更为宽广一些,且涡度、风场10波左右相关范围异常宽广。

从方差的功率谱特点来看(图5),沿二维波数方向散度的功率谱分布大值区以65波左右为中心且垂直范围较宽(图略),涡度小一些以40波左右为中心,而标量风的大值区垂直范围更窄以10波左右

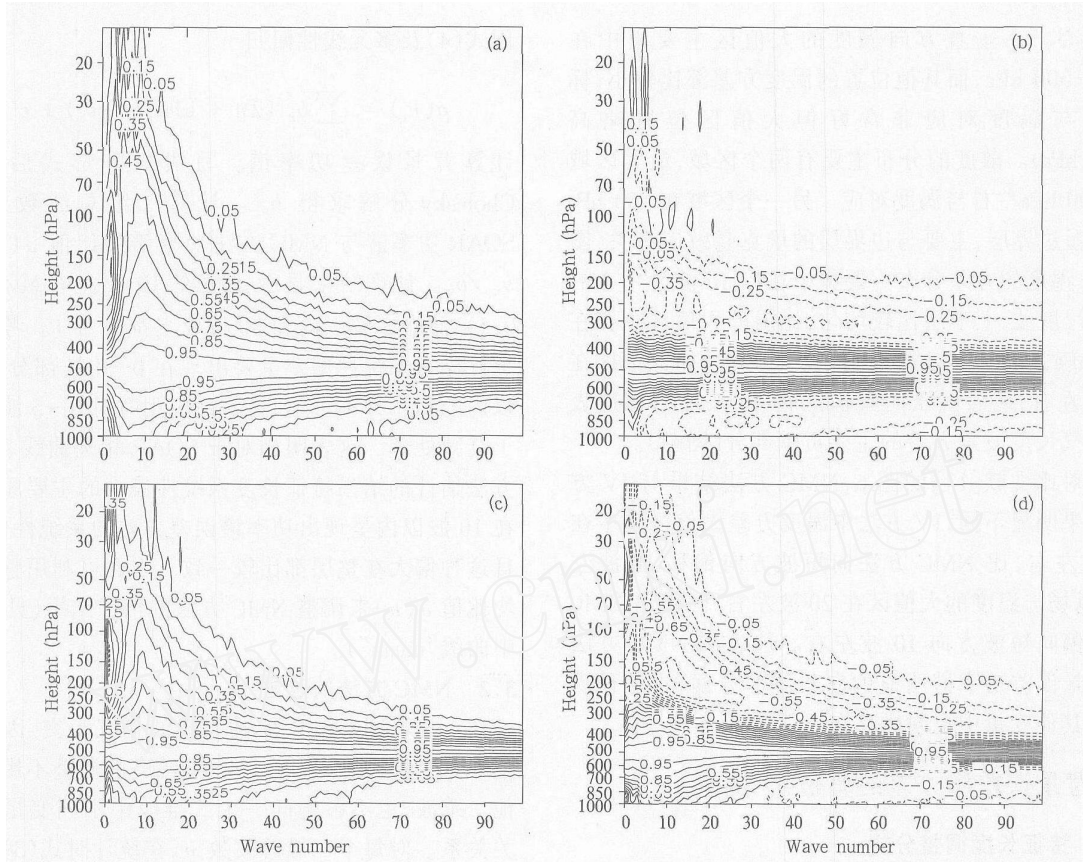


图4 NMC方法估计出的背景误差在500 hPa与其他层垂直相关随二维波数的分布(说明同图3)

Fig.4 500 hPa vertical correlations of background errors by NMC method for (a) vorticity, (b) divergence, (c) scale-wind, and (d) temperature

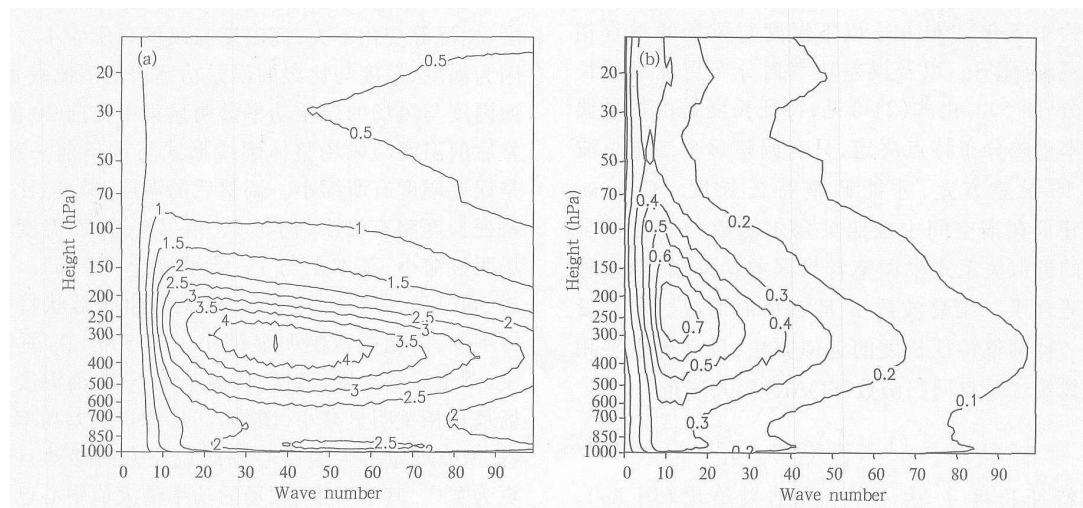


图5 NMC方法估计出的背景误差均方差随二维波数的分布
(a. 涡度, 单位: 10^6 s^{-1} ; b. 风场)

Fig.5 Background error standard deviations by NMC method for (a) vorticity (units: 10^6 s^{-1}), (b) scale-wind

为中心。在垂直方向涡度的大值区主要集中在 400—300 hPa, 而其他位置的涡度方差都比较小, 标量风与涡度对应非常好但大值区位置略高 (250 hPa)。散度的分布主要有两个区域, 最大区域在 300 hPa 左右与涡度对应。另一个区域在 950 hPa 的行星边界层, 主要与边界层的埃克曼效应有关 (图略)。温度的方差分布主要集中在以 10 波为中心的 天气尺度波, 上下层比较均匀, 比湿的大值区主要在 850 hPa, 是水汽的主要通道层次, 方差分布大值区在 30 波左右, 地面气压的大值区在 10 波左右。这表明天气尺度波是方差的主要贡献部分 (图略)。

对比文献 [14] 的图 6, NMC 方法结果与 IV 方法结果明显不同, IV 方法中涡度方差大值区集中在 50 波左右, 比 NMC 方法向短波方向偏移 10 波左右, 风场与温度的大值区在 20 波左右, 同样比 NMC 方法偏向短波方向 10 波左右。这表明 NMC 方法中背景误差功率谱明显的向长波部分偏移, 这种现象在其他变量中也同样存在。

3 背景误差协方差的调整

3.1 特征长度调整分析

NMC 方法估计的背景误差协方差结果与 IV 方法的结果存在显著差异, 表现在 NMC 方法在 200 hPa 以下低层的均方差要明显大于观测空间的结果 (图 2)、在 200 hPa 以上风场与温度的特征长度长出约 1.5 倍, 200 hPa 以下温度与风场的垂直相关更宽阔 (图 3)、以及误差功率谱分布更加靠近长波部分 (图 5)。由式 (2) 可见, 特征长度是由背景误差功率谱的分布特点决定, 只有调整每个二维总波数 n 的误差方差, 才能调整特征长度。Courtier 等 [9] 建议在谱空间中按照贝塞尔函数与勒让德多项式的近似关系式来调整背景误差功率谱, 但这种方法适合于二维总波数 n 足够小的情况。本文设计了一种调整特征长度的近似方法, 设理想水平相关曲线按二阶自回归函数 (SOAR) 形式分布:

$$(r) = (1 + \frac{r}{L}) \exp(-\frac{r}{L}) \quad (4)$$

式中特征长度 L 由 IV 方法结果给出 (图 6a)。NMC 估计的水平相关曲线由式 (1) 除以总方差归一化后计算获得。对 SOAR 相关曲线沿距离方向每 0.2 (约 22 km) 取一个样本点, 共取 400 个点。利

用式 (4) 及多元线性回归

$$(r_s) = \sum_{n=0}^{n_j} b_n (2n+1) P_n(\cos s) + \varepsilon \quad (5)$$

计算背景误差功率谱。写成矩阵形式后可用 Cholesky 分解求得 b_n 。计算二维总波数 n 的 SOAR 功率谱与 NMC 方法功率谱的比值 $\bar{(n)} = b_n / b_n$ 。计算结果显示比值在 10 波左右最小 (约 0.5), 而在 5 波以内和 20 波以上都大于 1。考虑到统计样本系统性偏差主要出现在 0—4 波部分, 5 波以内的结果并不可信, 在实际应用时对 0—5 波进行了衰减处理。这表明相对于 SOAR 相关曲线, NMC 方法估计的结果特征长度系统性偏大的主要原因是在 10 波以内呈现出功率谱误差方差的系统性偏大, 且这种偏大在整层都比较一致。因此可利用整层平均比值 $\bar{(n)}$ 来调整 NMC 方法的估计结果 (见图 9b 中曲线“mean”)。

3.2 NMC 方法的改进

由于不同总波数 n 的二维波相互正交, 因此在整个垂直层次上对某个波方差幅度的调整不影响其他二维波, 也不影响同一个二维波中不同变量的相关关系。对每个二维总波数 n , 在统计时式 (3) 变为

$$e(n, m, j, s) = \bar{(n)} \times ({}^{48}(n, m, j, s) - {}^{24}(n, m, j, s)) \quad (6)$$

重新利用 NMC 方法估计背景误差协方差, 背景误差平均廓线略有变动 (图略), 表现在涡度、散度、比湿等变化不大, 而温度与风场有所缩小。这是因为涡度、散度与比湿的误差功率谱范围比较宽广, 而温度与风场的误差功率谱明显集中在前 20 波, 调整后的温度与风场整体廓线形式与调整前一致, 只是误差幅度有所缩小。调整后的涡度、散度、比湿的特征长度略有缩小, 仍与 IV 结果接近, 而温度与风场明显缩小, 基本上与 IV 结果接近 (图 6a)。从模式 500 hPa 水平相关来看, 相关的曲线形状特征保持不变, 而相关范围明显缩小。对 500 hPa 垂直相关, 调整后的涡度、散度、比湿基本不变, 而温度与风场垂直相关明显减小 (图略)。从背景误差功率谱来看, 所有变量的大值中心都向短波方向移动且范围更为宽广, 其中温度、风场的功率谱大值中心已经从 10 波移到 20 波左右 (图 6b)。不同变量的关系在调整前后没有变化。利用整层比值 $\bar{(n)}$ 调整后的结果与 IV 结果更为接近。

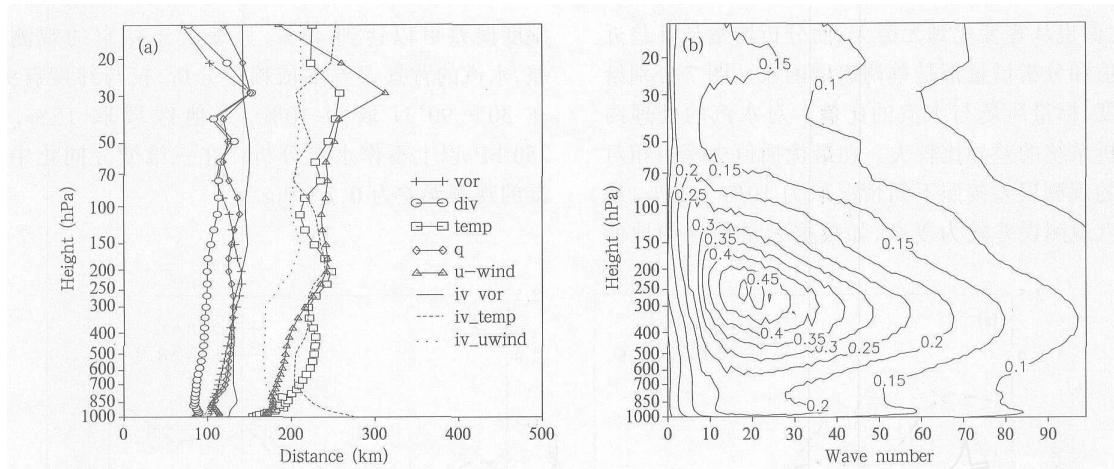


图6 利用整层平均比值调整后的NMC方法统计结果

(a. NMC方法特征长度垂直廓线与IV方法的比较:涡度、散度、温度、比湿、标量风的背景误差特征长度垂直廓线(单位:km); b. 同图5b)

Fig. 6 Tuned background error structures by NMC

(a. background error de-correlation length profile for vorticity, divergence, temperature, specific humidity and scale-wind by NMC and IV method (units: km); b. same as Fig. 5b)

4 背景与观测误差比较

4.1 资料同化系统的差异

全球三维变分同化系统采用美国国家环境预报中心谱统计插值(SSD),在三维变分同化中进行全局分析,采用更加合理的线性平衡关系,三维背景误差协方差的水平和垂直结构耦合在一起不可分离。采用NMC方法来估计背景误差协方差时所用的假设关系非常少,更多利用样本自身所蕴含的统计关系,如旋转风与辐散风间的耦合关系。此外由于分析在谱空间进行,不能完全反映区域变化的差异,如实际上特征长度在热带地区比中高纬地区要长出近一倍,而在三维变分同化中不能反映这一差异,这是该同化系统的主要缺陷。背景误差的方差在格点空间表示,可以随纬度、区域变化以反映不同区域非均质的特点,但只是静态估计结果,而现有最优插值中分析误差与背景误差随着观测信息的多寡与位置的不同而变化,体现逐日演变的背景误差特征。在最优插值中观测资料的质量控制过程与分析过程耦合在一起,资料是否剔除或使用随着分析循环的过程而变,这就保证了更多的资料能够被使用。而三维变分同化中资料的质量控制与分析过程分两步进行,这会丢掉一些在敏感区域非常有用的资料,如锋面、低涡等处的局地资料。这些区域的资料其代表性较

差,主要反映局地特征,与背景场的差异往往比较大,容易在质量控制过程中剔除。这些因素的差异会对分析结果产生一定的影响。

4.2 观测资料种类与观测误差

业务所用观测资料中无线电探空约500个站,飞机报约1600份,小球测风约550份,高低层卫星测风约3000份,地面陆面报告约3600份,船舶、浮标站约1000份,这些资料都经过预处理和质量控制。可见目前业务资料仍然以常规观测资料和卫星反演资料为主,这是目前业务数值预报水平不高的主要原因。随着模式分辨率的提高,观测代表性误差逐步减小,已有的观测误差不能代表目前模式分辨率下的实际情况。利用IV方法分离出无线电探空资料的观测误差,其他观测资料误差通过与探空资料进行比较分析来确定。

4.3 误差均方差的比较

分析增量可以表示为^[14]

$$x_a = BH^T(HBH^T + R)^{-1}(y - H(x_b)) \quad (7)$$

在相同观测增量 $d = y - H(x_b)$ 条件下,分析增量主要取决于背景误差与观测误差的相对大小,为简单起见只分析背景与观测标准差(矩阵对角线的均方根)之比。对单点观测增量 d 为1时,当背景与观测标准差相当时,比值为1而分析增量为0.5,当固定观测误差随着背景误差从0变化到无穷大

时,比值也从零变化到无穷大,而分析增量从 0 趋近 1,比值和分析增量都是单调递增函数。图 7 分别给出温度、标量风场与水汽的比值。对水汽的处理两个分析系统的差异比较大。在最优插值中探空相对湿度观测误差按照不同情况取为 10% - 20%,地面水汽观测误差较为复杂,如根据云量估计的地面

湿度误差可以达到 45%,显著大于探空的观测误差,水汽的背景误差不依赖于分析,仅与纬度有关,在 30°—90°N 取为 10%,其他区域取 15%,在 250 hPa 以上不作水汽分析。在三维变分同化中比湿的观测误差为 0.2 g/kg。

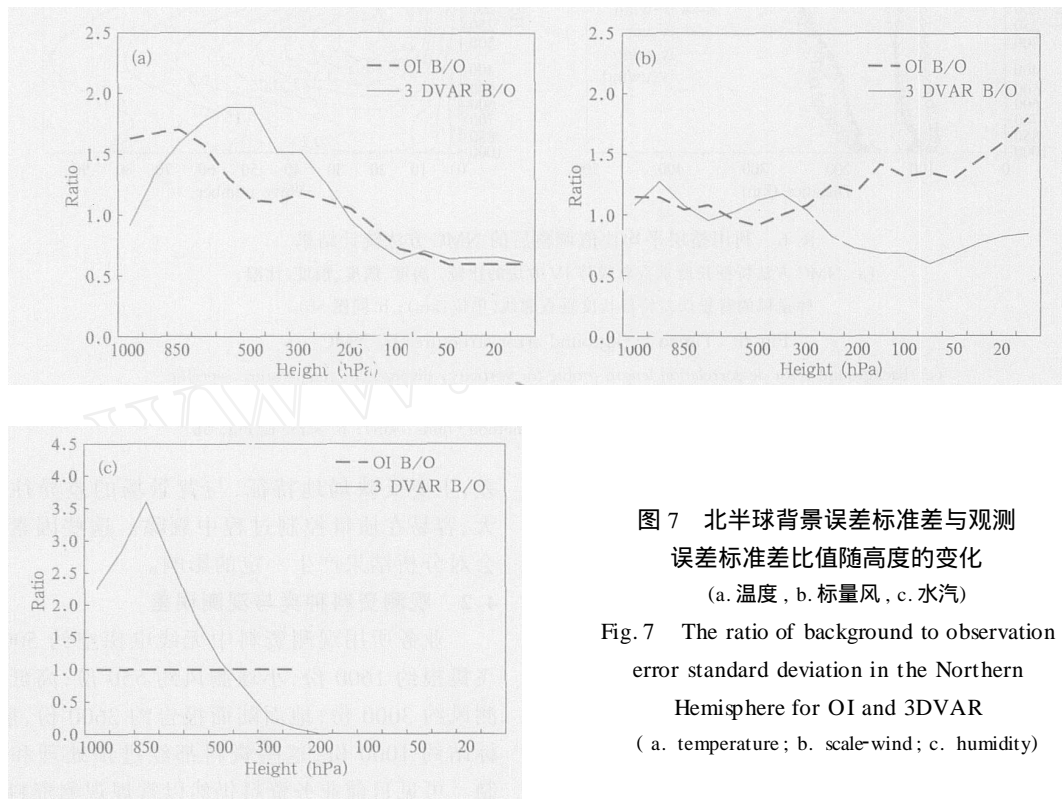


图 7 北半球背景误差标准差与观测误差标准差比值随高度的变化
(a. 温度, b. 标量风, c. 水汽)

Fig. 7 The ratio of background to observation error standard deviation in the Northern Hemisphere for OI and 3DVAR
(a. temperature; b. scale-wind; c. humidity)

对温度最优插值在 600 hPa 以下更信任观测,背景误差明显大于观测误差,而在 500 - 200 hPa 背景与观测误差相当,在 200 hPa 以上考虑更多的背景信息。对风场在 200 hPa 以下的中低层背景与观测误差相当,而在顶层背景误差明显偏大。风场与温度差别最大的是 200 hPa 以上的高层。对水汽两者可比性差,这是因为最优插值只给出探空观测误差的结果,而大量地面观测误差不易给出。

温度误差在三维变分同化中的特点与最优插值基本相当,但在 500 hPa 左右的中层背景误差比观测误差略大,而在 850 hPa 以下背景误差与观测误差相当。风场在 300 hPa 以下差别不大,但是在高层与温度场特点一致,都是考虑更多的背景信息。这两者在 300 hPa 以上高层的差别决定了它们在高层差异较大。对水汽场在 850 hPa 达到最大,然后

随高度迅速减弱。

5 分析增量特点

在同一时刻利用相同的背景场和相同的观测资料对比最优插值与三维变分同化的分析增量。虽然两者的质量控制方法有较大差异,但统计发现两者的差异比较小,绝大部分观测资料一致。因此分析增量的差异主要反映了背景误差、观测误差与同化系统的差异。由于陆地上的资料明显多于海洋上,北半球的资料明显多于南半球,因此分析增量的对比集中在北半球,并划分成多个小区域,在每个区域上对比分析增量的差异。

5.1 格点空间差异

在格点空间最优插值与三维变分同化的分析增量差异特征在模式中高层并不明显,主要差异表现

在最优插值的分析增量高低值间的梯度比较大,高值系统和低值系统非常清晰,而三维变分同化的分析增量比较平滑,从高值系统向低值系统的过渡比较光滑,梯度比较小。其原因可以从水平误差相关曲线的特点上给出合理解释,最优插值采用高斯型相关曲线,过高斯曲线拐点后相关系数衰减很快,这种特点决定了分析增量高低值系统非常明显。而三维变分同化的水平误差相关曲线是由不同二维总波数的功率谱决定,通过 NMC 方法由样本计算获得,其相关曲线随距离下降远没有高斯型曲线明显,在同样的特征长度情况下,后者的范围要比高斯型宽广得多。另一个原因是最优插值采用局地分析,分析的高低值系统更深一些。

在模式低层和高层的分析增量中,最优插值分析的质量场与风场间的平衡关系没有很好的保证,其原因可能是在高层中特征长度比低层增长了近一倍,在平衡关系不变的前提下平衡半径的范围要大得多。最优插值中特征长度不随高度变化,这就在高层造成风场与质量场间的不匹配。而三维变分同化的特征长度是从实际预报样本中统计得到,能够较好地体现风场与质量场的关系。此外,最优插

值在模式低层和高层都是背景误差大于观测误差(图7),平衡特征不容易保证。在最优插值分析增量中模式顶层易出现风场的反气旋特点和质量场的高值区在位置上并不匹配,存在非常明显的差异。而三维变分同化的分析增量中风场与质量场的匹配关系非常好,呈现出明显的线性平衡关系。这表明最优插值中过多的假设关系并不完全与实际关系相符,通过实际样本统计得到的误差特征与实际情况更为相符。

5.2 背景误差调整对同化分析的影响

背景误差调整前后对风场观测的分析影响较大。以北太平洋区(0° — 60° N, 150° E— 130° W)为例,观测资料主要为导风资料,与最优插值相比分析增量差异显著(图8)。最优插值分析增量的范围在垂直方向比较集中,在二维波数方向其大值区在40波以内且范围较宽,与IV方法估计的风场背景误差功率谱的分布特点比较一致(文献[14]的图6),而基于NMC方法来估计背景误差协方差的三维变分同化分析增量范围在垂直方向非常宽广,但在二维波数方向比较窄,大值范围在10波左右,同样与风场的背景误差功率谱的分布特点一致(图5b)。

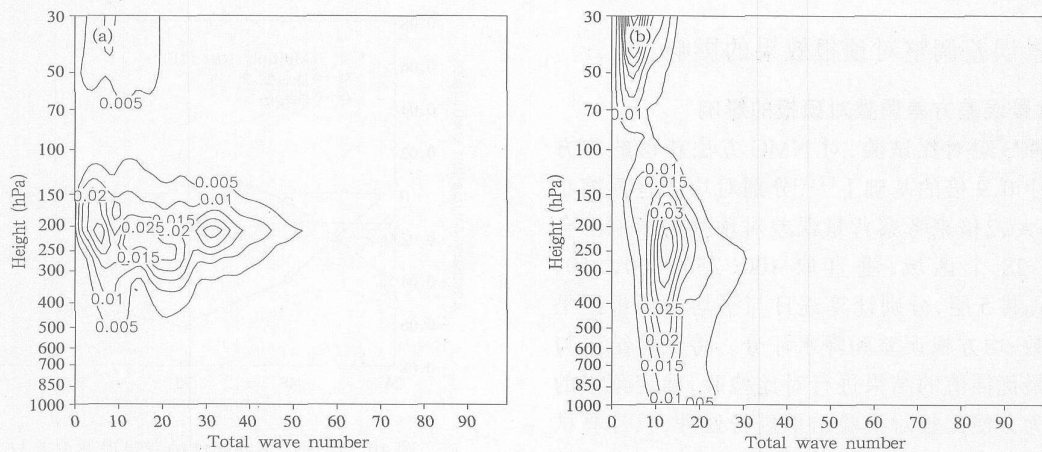


图8 北太平洋区标量风分析增量功率谱随二维波数的分布

(a. 最优插值, b. 三维变分)

Fig. 8 Energy spectra of scale-wind analysis increments in the North Pacific Ocean (a. OI, b. 3DVAR)

在北半球按照模式的垂直层次分成高中低3层,分别计算三维变分与最优插值分析增量功率谱比值(图9a)。尽管在不同层次不同的变量有所差异,但沿二维功率谱方向比值从高值(大于1)向低值(小于1)递减。超长波部分由于噪音过大结果不

可信。这表明在5—20波的天气尺度波部分三维变分同化的分析增量要明显强于最优插值,而在30波以后要明显弱于最优插值,背景误差功率谱大值区明显向长波方向偏移。将925—100 hPa整层平均 $\bar{(\cdot)}$ (n)比值与图9b曲线mean进行比较,可以发现这

两者的特点除了在 40 波以内幅度上的差异外,整体趋势非常一致。对背景误差调整后,三维变分分析增量功率谱与最优插值的比值递减特点显著改善。

不同区域分析增量特点有差别,主要原因在于背景误差和观测误差的非均质特点,如最优插值在北美区的观测误差要比中国区小一倍,比欧洲区小 $\sqrt{2}$ 倍。

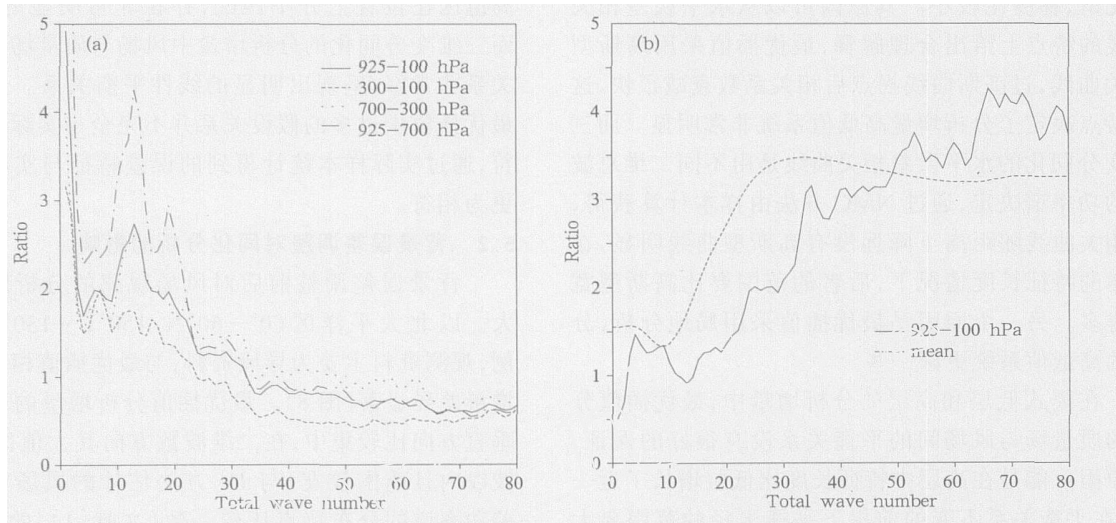


图 9 北半球分析增量功率谱比值随二维波数的分布

(a. 不同层次三维变分分析增量与最优插值的比值, b. 925—100 hPa 整层平均比值的倒数,同时给出比值 $\bar{r}(n)$)

Fig. 9 Ratio of the energy spectra of analysis increments in the Northern Hemisphere

(a. Ratios of analysis increments of 3DVAR to OI in 4 layers; b. the inverse of 925—100 hPa averaged ratio, and $\bar{r}(n)$ curve “mean”)

6 背景误差调整对预报效果的影响

6.1 背景误差方差调整对预报的影响

选择 3 组对比试验,对 NMC 方法获得的均方差在缩小 0.9 倍的基础上^[2],分别对均方差再缩小 2 倍,放大 $\sqrt{2}$ 倍来考察背景误差对预报的影响。全球分成 28 个区域,垂直取 100, 250, 500, 850, 1000 hPa 共 5 层,分别计算逐日与平均的预报距平相关系数、均方根误差和降水评分。将试验结果与同时段最优插值的结果进行对比检验,确定预报的差异。为方便比较对检验结果进行处理,只计算试验与业务检验评分的差值,当差值为正时表明三维变分效果占优,反之则不及(图 10)。

试验发现,背景误差的调整对随后预报非常敏感,对 24 到 120 h 预报,当背景误差从偏小调整到偏大,相应 24 h 预报评分从高变到低,然而他们的趋势相反。以中国区 500 hPa 为例,对背景误差偏小情况,在开始 24 h 距平相关系数高,与最优插值效果相当,随后逐渐比最优插值效果差。与此相反对背景误差偏大情况,在开始 24 h 距平相关系数略

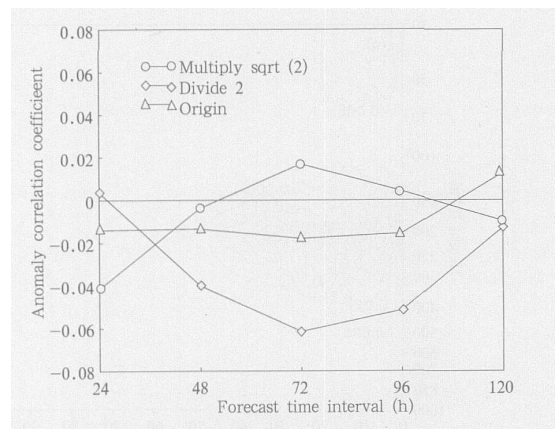


图 10 中国区不同幅度的背景误差方差与相应 500 hPa 预报距平相关系数 (三维变分减业务最优插值)

Fig. 10 China region 500 hPa anomaly correlation coefficients in crement between 3DVAR and OI according to various background error variances

低,不如最优插值效果,然而到 48 h 与最优插值效果持平,再往后超过最优插值效果。对于恰当的背景误差,从 24 h 开始至 120 h 效果始终与最优插值

效果的差距比较稳定。在不同层次和不同区域,虽然没有严格遵守这个规律,但基本上都表现出开始24 h与随后预报的“翘翘板”特点。这表明背景误差偏小虽然会带来24 h预报的改善,但会对随后的预报不利,而背景误差偏大会影响到24 h的预报,对随后的预报会改善。最恰当的情况是背景误差与实际误差一致。

6.2 背景误差特征长度调整对预报的影响

对背景误差功率谱分布特点的调整来改变特征长度会影响到分析增量和随后的预报效果。分析表明在北半球这种调整对以云导风资料为主的区域最

敏感,如北太平洋、北大西洋地区(25° — 65° N, 15° — 60° W)。本文选择两组对比试验,分别是直接利用NMC方法获得的特征长度(图略)和调整后的特征长度(图6a),对250 hPa夏季85 d的平均结果,太平洋地区特征长度调整前三维变分的预报效果除24 h预报时效外明显劣于最优插值,预报技巧相差接近1 d,而调整后的预报效果与最优插值相当。对大西洋区域调整前预报效果与最优插值相当,调整后整体都有改善(图略)。两个区域结果都表明特征长度调整后都较调整前有明显的改善,说明特征长度决定着观测信息在空间的扩散范围,恰

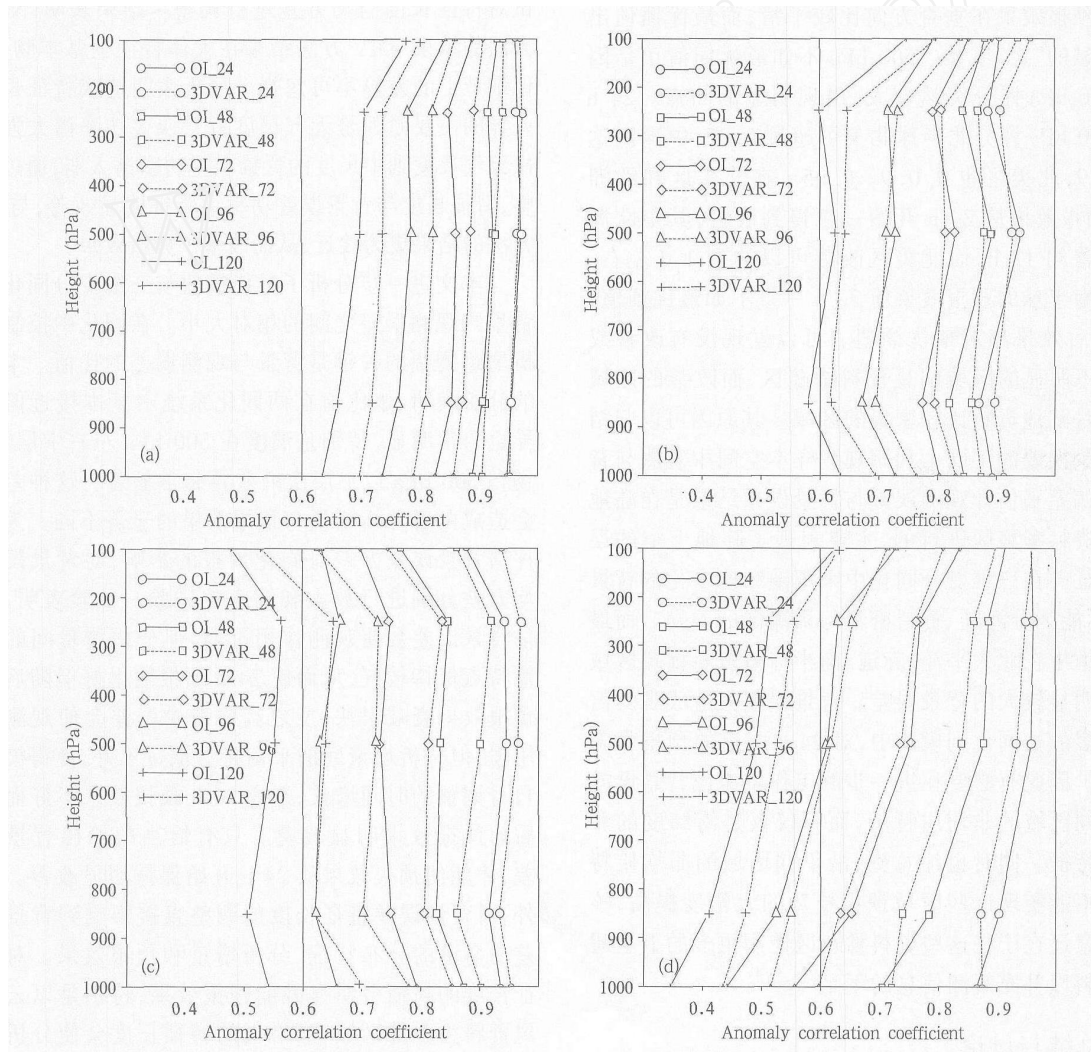


图11 不同分区的100,250,500,850,1000 hPa的120 h预报距平相关系数垂直廓线
(a. 北半球区, b. 中国区, c. 北美区, d. 南半球区;图中实线为最优插值预报评分,虚线为三维变分同化评分)

Fig. 11 120-hour forecast anomaly correlation coefficient profile at 100, 250, 500,

850, 1000 hPa for various regions

(a. North hemisphere; b. China; c. North America; d. South hemisphere. Solid line for OI, and dash line for 3DVAR)

当的估计至关重要,而利用观测样本空间估计的背景误差特征长度来调整 NMC 方法估计出的特征长度,能有效改善三维变分同化的分析预报效果。

6.3 分区预报的影响

从总体评分来看,北半球 24 h 的预报三维变分与最优插值的差异不大,高低层评分都比较一致,从 48 h 开始三维变分同化效果在不同的检验层上都开始超出,至 96 h 预报评分超出 12 h 左右,差异非常显著。对中国区在 500 hPa 的预报效果都要比最优插值略差一些,但其他层次从 24 h 开始就明显好于最优插值。而且从评分上也注意到,三维变分同化的预报效果在垂直方向比较平滑,而最优插值出现明显的“之”字型。500 hPa 不如最优插值正是因为 500 hPa 评分突然变大,出现明显的凸点。24 h 的 500 hPa 评分北半球北美可超过 0.95,中国区达到 0.9,北美区也在 0.9—0.95。在北美区和欧洲区,预报效果从 24 h 开始一直得到改善,其中欧洲区改善约 12 h,而北美区改善可以达到 1 d 左右。然而南半球的预报效果前 72 h 一直不如最优插值,72 h 后效果超过最优插值。可以发现没有改善或效果不明显的区域都是资料稀疏区,而改善的区域主要是陆地资料比较密集的区域。其原因可以归结到背景误差的不同。利用观测样本空间中观测与背景的背景误差来估计观测误差与背景误差结果是在陆地探空资料密集区进行的,主要表现了陆地上空的误差特征。在三维变分同化中采用随纬度变化的背景场,不能区分海陆、资料密集与稀疏区的差异。而最优插值中在北太平洋、赤道、南半球等资料稀疏区域呈现明显较大的背景误差。背景误差大造成观测信息更多的被同化到模式中,对 24 h 之后的预报带来改善。因此有必要在进一步的工作中考虑背景误差沿不同区域的非均质特点,而不仅仅是沿纬度的非均质特征。同时也注意到,沿不同区域的非均质特点并不能实现这些区域预报技巧的大幅度提高,核心因素还在于在这些资料稀疏区使用更多的卫星遥感资料以补充观测信息的不足^[15,16]。

7 总结与讨论

NMC 方法能克服 IV 方法所要求的主要针对常规探空资料、观测系统均一、区域范围不大,只能描述 10 波以上 100 波以下的波谱特征的不足,对全球谱模式而言这种方法的优点是可以提供全球同化

所需的所有波谱和模式垂直层次背景误差协方差特征,还可以统计出不同变量间的平衡与约束关系。但 NMC 方法的缺陷非常明显,本文的分析表明其估计出的结果表现在 200 hPa 以下底层的均方差要明显大于观测空间的结果(图 2)、在 200 hPa 以上风场与温度的特征长度长出约 1.5 倍,200 hPa 以下温度与风场的垂直相关更宽阔(图 3)、以及误差功率谱分布更加靠近长波部分(图 5),在使用时要谨慎。NMC 方法的均方差与特征长度等信息同样在 IV 结果中也能获得,使用 IV 方法的结果可以提供很好的参考信息来调整 NMC 方法的参数,本文尝试对特征长度与均方差进行调整。结果表明,NMC 方法的结果与 IV 方法结果在大体特征上基本吻合,但细节上的差异不可忽视。分析表明造成特征长度差异的主要原因是天气尺度背景误差功率谱太强而次天气尺度到中尺度的背景误差功率谱太弱,通过调整,特征长度与背景误差功率谱特征明显改善,与 IV 方法的结果更为接近,从而改善了分析效果。

本文进一步分析了最优插值与三维变分同化中背景与观测误差之间的相对大小。在同化中控制分析增量增幅的关键是背景与观测误差的比值。本文的分析表明,该比值在两同化系统中基本接近但差异也非常明显,特别是温度在 500 hPa 左右中层、风场在 300 hPa 以上层次的差异非常显著。这种差异会造成两者分析增量与预报效果的显著不同。为了评估背景误差方差对同化预报的影响,对背景误差均方差分别进行放大和缩小的试验。试验表明,低估背景误差会使观测作用过弱,但分析增量间的平衡特点保持较好,从而使 24 h 预报效果好但随后的预报效果衰减很快,反之高估背景误差会使观测作用强,但分析增量间的平衡特点保持不好,使得模式自身调整的时间增长,造成 24 h 预报效果不好而随后的预报效果明显改善。只有恰当的给出背景误差,才能使预报效果从 24 h 开始保持明显改善。此外,对背景误差特征长度的调整直接影响到背景误差功率谱的分布特征、分析增量和预报效果。对特征长度的高估会明显减弱预报效果,特别是以云导风资料为主的区域,而恰当的特征长度会使分析增量的分布比较合理。

参考文献

- [1] Parrish D F, Derber J C. The national meteorological center's

- spectral statistical-interpolation analysis system. *Mon Wea Rev*, 1992, 120: 1747-1763
- [2] Derber J, Bouttier F A. Reformulation of the background error covariance in the ECMWF global data assimilation system. *Tellus*, 1999, 51A: 195-221
- [3] Hollingsworth A, Lonnberg P. The statistical structure of short-range forecast errors as determined from radiosonde data, Part I: the wind field. *Tellus*, 1986, 38A: 111-136
- [4] Lonnberg P, Hollingsworth A. The statistical structure of short-range forecast errors as determined from radiosonde data, Part II: the covariance of height and wind error. *Tellus*, 1986, 38A: 137-161
- [5] Xu Q, Wei L, Van Tuyl A, et al. Estimation of three-dimensional error covariances, Part I: analysis of height innovation vectors. *Mon Wea Rev*, 2001, 129: 2126-2135
- [6] Xu Q, Wei L. Estimation of three-dimensional error covariances, Part II: Analysis of wind innovation vectors. *Mon Wea Rev*, 2001, 129: 2939-2954
- [7] Xu Q, Wei L. Estimation of three-dimensional error covariances, Part III: height-wind forecast error correlation and related geostrophy. *Mon Wea Rev*, 2002, 130:1052-1062
- [8] Ide K P, Courtier P, Ghil M, et al. Unified notation for data assimilation: operational, sequential and variational. *J Meteor Soc Japan*, 1997, 75B:181-189
- [9] Courtier P, Andersson E, Heckley W, et al. The ECMWF implementation of three-dimensional variational assimilation (3DVAR). I: Formulation. *Quart J Roy Met Soc*, 1998, 124: 1783-1807
- [10] Daley R. *Atmospheric Data Analysis*. Cambridge University Press, 1991. 457pp
- [11] Rabier F, McNally A, Anderson E, et al. The ECMWF implementation of three-dimensional variational assimilation (3DVAR). II: Structure functions. *Quart J Roy Met Soc*, 1998, 124: 1809-1829
- [12] Andersson E, Haseler J, Undén P, et al. The ECMWF implementation of three dimensional variational assimilation. III: Experimental results. *Quart J Roy Met Soc*, 1998, 124: 1830-1860
- [13] Ingleby N B. The statistical structure of forecast errors and its representation in the met office global 3DVAR. *Quart J Roy Met Soc*, 2001, 127: 209-231
- [14] 龚建东,魏丽,陶士伟等. 全球资料同化中误差协方差三维结构的准确估计与应用:观测空间协方差的准确估计. *气象学报*, 2006, 64(6):667-681
- Gong Jiandong, Wei Li, Tao Shiwei, et al. Accurate estimation and application of 3-d error covariance structures in global data assimilation, Part I: Accurate estimation of error covariance in observation space. *Acta Meteor Sinica (in Chinese)*, 2006, 64(6): 667-681
- [15] Derber J C, Wu W-S. The use of TOVS cloud-cleared radiances in the NCEP SSI analysis system. *Mon Wea Rev*, 1998, 126: 2287-2299
- [16] McNally A P, Derber J C, Wu W-S, et al. The use of TOVS level 1B radiances in the NCEP SSI analysis system. *Quart J Roy Met Soc*, 2000, 126:689-724

ACCURATE ESTIMATION AND APPLICATION OF 3-D ERROR COVARIANCE STRUCTURES IN GLOBAL DATA ASSIMILATION

Part : Background Error Covariance Structure Adjustments and Numerical Experiments

Gong Jiandong Zhao Gang

National Meteorological Centre, Beijing 100081

Abstract

The so-called NMC-method is used to analyze the three major characteristics of the T213L31 model outputs: error variance, three-dimensional correlation and de-correlation length. Results from the NMC-method are compared with those from the traditional innovation vector method, and then its parameters are tuned based on the comparison. It is shown that results from both methods are roughly similar, but some differences still exist in the background error variance and de-correlation length estimations. It seems that the NMC-method intends to over-estimate background errors for synoptic scale, but under-estimate them for sub-synoptic scale and meso-scale. After the tuning of background error variance and de-correlation length, results from the NMC-method are closer to those of the innovation vector method, and the analysis results are improved. Comparing the ratios of the background to observation errors for the currently operational optimal interpolation system (OI) with the three-dimensional variational data assimilation (3DVAR), it is shown that the two system results are roughly similar, but the background errors for temperatures below 850 hPa and for winds above 300 hPa in the 3D-VAR assimilation are relatively smaller than their observation errors. This is helpful to make sure the mass and wind fields in balance within the analysis, to reduce the unbalance part of the two fields near surface and upper level in the OI system. This paper shows that the accurate estimations and tunings of background error covariance improved the forecast skill, and in the north hemisphere, the forecast for 120 hours made from the analysis from the three dimensional variational data assimilation system is better than that using the currently operational OI system.

Key words: Background error covariance, NMC-method estimation, Tuning with innovation vector, Numerical contrast experiments.