

# 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\*

龚建东<sup>1</sup> 魏 丽<sup>2</sup> 陶士伟<sup>1</sup> 赵 刚<sup>1</sup> 万 丰<sup>1</sup>

<sup>1</sup> 国家气象中心,北京,100081

<sup>2</sup> 中国科学院寒区旱区环境与工程研究所,兰州,730000

## 摘 要

观测误差与背景误差协方差在四维资料同化和业务资料分析系统中起到决定性作用,它决定着观测信息与背景初猜值信息的相对重要性以及这些信息在空间及不同变量间的扩展方式。由于实际大气的真值并不知道,需要发展不同的技巧来估计观测误差与背景误差协方差,其中在观测空间利用观测与背景初猜值之差来分离观测误差与背景误差协方差的方法估计出的结果较为准确,其估计出的观测误差可直接用于资料分析系统中,背景误差可作为标尺来度量其他方法估计结果的可靠性。文章采用国家气象中心 T213L31 全球中期分析预报系统的 6 h 预报作为背景初猜场及同时段冬夏两个季节的北半球探空,利用贝塞尔函数拟合方法来分离观测误差与背景误差协方差,并比较了东亚区、北美区、欧洲区 3 个探空资料均匀密集区的区域与季节变化结果。结果表明,观测空间拟合方法所要求的水平均质、各向同性在欧洲区和北美区成立程度较好,在东亚区略差,使用时需要斟酌。此外均方差区域间差别较大,在冬季明显大于夏季,温度场偏大 0.2 K,风场偏大 0.9 m/s。温度场在 400 hPa 以下与 150 hPa 以上,背景误差略小于观测误差,而在 200—300 hPa,背景误差略大一些。风场的特点与温度场比较一致。温度与风场背景误差主要集中在前 40 波,并在 20 波左右达到最大,水平相关季节区域差别不大,而温度垂直相关比风场窄,两者相关范围比较大的波数主要集中在前 20 波。此外利用贝塞尔函数拟合方法获得结果的分析表明,在质量场中不同区域季节间温度误差的稳定性要明显好于高度场,涡度散度的稳定性要明显好于流函数和势函数,特别是对于特征长度更为明显。这表明利用贝塞尔函数拟合方法获得的结果对校准在全球资料同化中采用温度、涡度散度作为资料同化的分析变量具有一定的优势。

关键词: 观测误差,背景误差协方差,贝塞尔函数拟合,观测空间分离,准确估计。

## 1 引 言

数值天气预报技术已经成为天气预报的核心支撑,是世界各主要气象中心的重要业务组成部分。中国国家气象中心从 2002 年起建立了基于 T213L31 模式的全球中期业务分析预报系统,水平分辨率约 60 km,预报能力有了大幅度提高,其中北半球夏季和冬季 500 hPa 环流形势可用预报时效分别达到了 5 d 和 6—6.5 d<sup>[1]</sup>。尽管如此,与世界先

进水平仍然有 1—2 d 的差距,其核心原因是业务上采用的最优插值系统以常规资料为主,不能直接同化大量的卫星雷达等遥感遥测资料。从 2002 年起中国国家气象中心着手全球资料同化从最优插值向三维变分的升级工作,在这个过程中,分析资料同化中的误差协方差三维结构来提供客观的依据与标准,对改进资料同化效果非常必要。

资料同化需要综合处理来自两方面的信息,一方面是源自于以前时刻的、由预报模式演变而来的背景

\* 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

初猜值信息,另一方面是当前时刻的观测信息<sup>[2]</sup>。按最优估计理论,这两种信息在分析场中的可靠程度依赖于各自的误差统计特征,以及不同变量间的关联关系<sup>[3]</sup>。由于实际大气状态的真值并不知道,无法直接获得这两种误差大小,但可以在一定的合理假设条件下,利用大量的模式预报结果与观测结果通过统计方法来估计误差协方差的三维结构<sup>[4-7]</sup>。由于误差协方差三维结构涉及到变量的方差、特征长度和变量间的协相关关系,不同的假设侧重点有所突出,其中一种是在观测空间进行,所谓的更新矢量方法(Innovation Vector 方法,简称 IV 方法)<sup>[4-7]</sup>。这种方法的核心思路是先求取观测与背景初猜值间的偏差量,而该偏差量等价于观测误差与背景误差之差,下一步通过合理假设分离观测误差与背景误差。该思路是由 Hollingsworth 和 Loennberg<sup>[6-7]</sup>在 Gandin<sup>[4]</sup>和 Rutherford<sup>[5]</sup>的思路提出,根据北美地区探空观测资料和 ECMWF 预报资料,在水平均质、各向同性假设下分析给出了位势高度、风场以及高度-风之间地转关系的误差结构特征。在对样本进行最小二乘拟合中,采用的协相关函数也从早期的二阶自回归函数(SOAR)发展到利用贝塞尔函数,后者最大的优势在于不仅可以给出方差、特征长度等相关特征,而且可以揭示出不同波谱能量的误差分布特征。这种方法可以获得区域平均的观测误差和背景误差协方差,适用于观测仪器比较统一、观测站密集地区,如北美、东亚、欧洲大陆地区,其最大优势是在所有误差估计方法中较为准确,可以作为其他估计方法的参考值。这种方法也衍生出其他方法,如最大似然估计方法(Maximum-likelihood)<sup>[8]</sup>,适合于移动性观测网(如船舶等)。这种方法的弱点除了只能在观测比较均一密集的地区且主要针对探空外,对不同变量间的协相关约束关系无能为力。早期的质量场与风场间的地转近似在实际应用中显得比较粗糙,需要其他方法匹配来补充。

由于贝塞尔函数是二维傅立叶变换在柱坐标下的表示形式,贝塞尔函数的谱系数同样代表了不同波数的误差能量分布,因此贝塞尔函数的谱系数在正定条件上有比较严格的限制。对于单层变量自协相关,要求每个系数满足半正定条件即可。而对于多(层)变量协相关,要求谱系数的函数表达满足半正定条件。在 Hollingsworth 和 Loennberg 的工作中,正定条件存在漏洞<sup>[9,10]</sup>。Xu 和 Wei<sup>[9,10]</sup>(简称

XW01) 针对 Hollingsworth 和 Loennberg 的工作进行优化,提出了更为合理的正定约束条件和表达计算方式。他们利用美国海军全球模式和北美地区观测对北美地区的误差三维结构特征进行分析,一个主要结论是 1986—2002 年的 16 年,北美区的高度背景误差减小了约 20%,同时观测误差(探空)也有轻微的减小<sup>[9]</sup>。

这个结论给予我们启示,国家气象中心的 T213L31 全球中期分析预报系统误差参数是 20 世纪 80 年代后期确定的,主要参考了 Hollingsworth 和 Loennberg 的结果,同期的模式分辨率是 T42L9,相当于 320 km 的水平分辨率<sup>[11]</sup>。模式预报性能的提高带来背景误差的相应减小,同时水平分辨率的提高也会造成观测误差中代表误差项也相应减小。此外不同国家探空仪器各异,误差特点不同,欧洲、北美的结果不完全适合于中国,因而有必要重新估计误差的三维结构特征,修正现有业务中资料同化系统中的误差参数。同时估计结果也将主要作为基准用于变分同化系统的误差参数校准中。

本文重点在于分析揭示中国国家气象中心全球中期系统的误差三维结构特征,采用 XW01 发展的误差协方差三维估计方法及 T213L31 全球中期分析预报系统的预报结果和探空观测,估计北美、东亚、欧洲大陆等资料密集区的温度、风场误差结构特征。考虑到全球三维变分同化系统采用涡度、散度而不是流函数、势函数作为控制变量,在数学表达上对 XW01 工作进行了扩充。

## 2 IV 方法假定条件特征分析

IV 方法假设的基础条件是观测仪器比较一致,且相关系数在水平方向均匀和各向同性<sup>[4-7]</sup>,只有在这种假设下,样本才能足够的多,结果才是一种平均意义上的且有代表性,而实际上这种假设在不同的区域满足程度不同。观测系统的不一致、地理位置的差异都限制了水平均匀性和各向同性。这就需要分析资料来检验 IV 方法假设的无偏性、均一性和各向同性合理性。此外,虽然背景场特性比较稳定,观测资料却往往受到非气象意义误差(Gross error)干扰,在进行误差协方差三维结构估计工作之前,需要对资料进行预处理。

### 2.1 误差分离原理与基本假定

分别定义大气背景初猜值为  $z^f$ 、真值为  $z^t$  和观

测值为  $z^o$ , 背景误差为  $e^f = z^f - z^t$ ; 观测误差为  $e^o = z^o - z^t$ ; IV 值定义为  $d = z^f - z^o$ 。

假定: (1) 背景和观测误差都是无偏的, 即  $e^f = e^o = d = 0$ ;  $\cdot$  表示数学期望。(2) 观测误差与背景误差不相关, 即  $e^o, e^f = 0$ ;  $\cdot$  表示二阶距。(3) 各站之间观测误差不相关, 即当  $i \neq j$  时,  $e_i^o, e_j^o = 0$ ;  $i, j$  为站点位置。(4) 背景误差在水平方向均匀且各向同性,  $e_i^f, e_j^f$  仅为两点间距离  $r_{ij}$  的函数。于是 IV 值的协方差可以写成:

$$C(r_{ij}) = d_i, d_j = e_i^o, e_j^o + e_i^f, e_j^f$$

当  $r_{ij} \rightarrow 0$  时, 上式给出背景误差的协方差; 当  $r_{ij} = 0$  时, 成为背景方差和观测方差之和。所有可能的站点对的  $d_i, d_j$  值则给出背景误差协方差随距离  $r$  变化的函数。由  $r_{ij} = 0$  的点可以拟合出背景误差协方差函数曲线, 外推拟合曲线在  $r = 0$  时的值被看作是背景误差的方差, 同时由 IV 值方差减去背景方差就得到观测误差的方差。

由此可见, 在这些假设下, 根据 IV 值可以分离背景和观测误差的方差, 并估计出背景误差的协方差(或相关函数)。本文仅给出必要的原理与公式说明, 实际上误差估计不是在某一单层上进行分析, 而是计算不同层次与水平距离间的三维协方差进行分析, 更为详细的可以参阅 XW01 工作。针对 T213 模式, 有必要对这些假设的成立程度针对不同地区、不同季节进行分析检验。

## 2.2 资料预处理

IV 方法统计样本采用 2002 年 12 月至 2003 年 2 月和 2003 年 7—9 月冬夏两个季度的 T213L31 全球中期分析预报系统每天 06Z 和 18Z 两个时次的 6 h 预报场插值到标准等压面作为背景初猜场。T213L31 系统已于 2002 年 9 月业务运行, 提供 10 d 中期预报。在此期间中国探空网绝大部分采用 59 型探空仪, L 波段电子探空仪正在换型但数目较少, 有记录的到 2003 年上半年为 7 个, 至 9 月份为 14 个, 为避免干扰对 L 波段电子探空仪温度观测进行剔除处理, 而风场的影响不大予以保留。北美区与欧洲区采用 L 波段电子探空仪。已有分析表明 L 波段电子探空仪在大气高层对温度的观测精度要比 59 型探空仪高。

观测资料采用 00Z 和 12Z 两个时次全球范围内的探空和测风资料。首先对探空资料进行要素基本属性(包括观测经纬度位置、时间、台站高度、地面

气压等有效性)检查、无效观测或要素剔除等必要处理。观测质量控制采用一系列误差检查方法, 对观测资料中存在的非气象意义误差或错误进行检测, 以便在统计分析时排除这些错误的影响, 使统计分析结果具有足够的代表性。质量控制采用综合质量控制(CQC)技术和增量极值检查技术<sup>[12]</sup>。CQC 针对高度及温度, 它的核心思想是首先对要素进行各种有效方法的 CQC 分量检查, 然后对 CQC 分量信息作综合分析来识别出资料中的错误资料及其来源, 后一步称为 CQC 误差决策算法。这里包含的 CQC 分量有 5 个: 静力学检查、极值检查、对预报背景场的一致性检查、空间水平方向的一致性检查与空间垂直方向的一致性检查。而对风场主要采用极值检查, 即将它与事先确定的极值进行对比, 由此来判断非气象意义误差是否存在。

针对北美(25°—65°N, 230°—300°E)、东亚(20°—50°N, 75°—135°E)、欧洲大陆(30°—70°N, 0°—60°E)等资料密集区进行统计和拟合。对经过质量控制的观测点计算在观测位置的背景初猜值与观测的 IV 值, 对 IV 值按时间序列计算不同站点、层次间协方差, 当样本数大于 10 才认为统计有效。按照测站对的距离以 100 km 间隔为节点合并站点对。XW01 工作表明, 采用平均方式合并站点对就可以取得非常好的结果<sup>[9]</sup>。对所有合并好的站点对, 用最小二乘法进行拟合求出功率谱系数, 并由功率谱系数反算出拟合曲线以便分离观测误差与背景误差。

## 2.3 误差无偏性

IV 方法假设观测误差和背景误差都是无偏的, 也即观测误差、背景误差和 IV 值三者的时间平均都应为零。但这种假设依赖于各国的探空仪器类型、质量控制方案水平与地形特征等一系列因素。以冬季 200 hPa 为例, 中国区温度 IV 值平均偏差以正负值成片出现为主, 部分区域正负交错。正值区出现在华北、东北等区域, 占到四分之一象限, 其他大部分区域以负值为主。正值区平均约 0.5 K, 在东北中部较大, 最大可达 1.2 K。负值区在 100°E 以西地区平均约 -1 K, 100°E 以东、长江以南约 -0.5 K, 其中在青藏高原东部、新疆北部、南海地区明显偏大。在北美区温度 IV 值平均偏差以负值为主, 没有正负成片出现现象, 偏差约 -0.5 K, 但南部区域偏差明显大于其他区域, 最大 -1.5 K。在欧洲区温度 IV 值平均偏差以负值为主, 整体基

本均匀,偏差约  $-0.5\text{ K}$ ,但西北区域明显大于其他区域。这种特征在不同层次、不同季节大体相当。这表明观测资料与背景初猜场二者之一或二者都是有偏的,尽管现在无法从 IV 值平均偏差中分离出观测平均偏差和背景初猜场平均偏差,但提醒我们在分析同化系统中有必要考虑这个事实,否则分析同化的结果也将是有偏的。此外温度场偏差特点也表明与地形特点密切相关,特别是青藏高原东部地区。各站点 IV 值平均偏差在进行协方差统计时必须予以扣除,以保证无偏性假设。相对而言,欧洲区、北美区整体均匀性要好于东亚区,扣除平均值偏差对站点间协方差统计结果影响不大,而对中国区要显著一些。

对风矢量也假设观测误差和背景误差都是无偏的,但实际上每个区域都有偏差,一般偏差的风速越小、风向随机性越好则无偏性越好。以冬季  $500\text{ hPa}$  为例,东亚区 IV 值平均偏差风速大,大部分区域在  $0-1\text{ m/s}$ ,而青藏高原东部在  $2-3\text{ m/s}$ ,西南地区在  $2-4\text{ m/s}$ ,印度中北部地区要大得多,部分站达到  $3-5\text{ m/s}$ 。东亚部分区域风向呈现出明显的系统性结构,如在冬季以四川盆地为中心的相当大范围内,风矢量沿逆时针方向呈闭合环流,常年性的印度次大陆和华南地区以南风为主而东北地区以北风为主。这种特征在不同层次、不同季节都会出现,在高层  $200\text{ hPa}$  更为明显,这表明东亚区观测资料与背景场二者之一或二者有偏且无偏性较差。在北美区 IV 值平均偏差风速小且风向比较随机,大约在  $0-1\text{ m/s}$ ,东部沿海略大一些在  $1-2\text{ m/s}$ ,整场基本均匀,表明北美区的风矢量无偏性较好。欧洲区的特点与北美区比较接近。

## 2.4 方差均一性

观测误差与背景误差方差水平均一性是均质假设的一个重要方面,在统计时已经扣除了各站点 IV 值平均偏差。以冬季  $200\text{ hPa}$  为例,中国区 IV 值均方差地域差别不十分明显,均质特点基本成立。 $110\text{ E}$  以东区域好于西部,东部在  $1.6\text{ K}$  左右而西部  $1.8\text{ K}$  左右,方差偏大的主要是青藏高原东部与新疆北部地区。在北美区西部、南部也比较大,在  $2\text{ K}$  左右,而东北部明显偏小,在  $1.5\text{ K}$  左右,从南到北逐步减小。欧洲区除南部略大外整场比较均匀,平均在  $1.5\text{ K}$  左右。东亚与北美的方差在冬季都明显大于夏季,如北美夏季平均仅为  $1\text{ K}$  左右,而欧洲区季节变化不明显。这种特征在不同层次、不同季节都会出现,在高层  $200\text{ hPa}$  更为明显。区域差别的原因可能是在欧洲区、北美区地形平坦,背景初猜场的精度较好,部分原因是 L 波段探空仪器观测资料精度较高。而在东亚区地形很复杂,背景初猜场的精度不如欧洲区。总体而言欧洲区、北美区均质特征略好于东亚区。

$500\text{ hPa}$  风速的均方差在中国大陆东北部为  $4-5\text{ m/s}$ ,其他区域  $5-6\text{ m/s}$ ,印度次大陆偏大一些为  $6-7\text{ m/s}$ 。北美区约  $5\text{ m/s}$  与中国区接近,但西部沿海明显偏大达到  $7\text{ m/s}$ ,欧洲区在  $4-5\text{ m/s}$  与日本接近,南部略大为  $5-6\text{ m/s}$ 。相对于温度,3个区域的风场均方差整体都呈现出更好的均质特征,但总体而言欧洲区、北美区均质特征略好于东亚区。

## 2.5 协方差各向同性

另一个假设是观测误差与背景误差各向同性。按方位角(以北为正)和站点间距离提取样本进行分析(图1),当协方差随方位角没有系统性的变

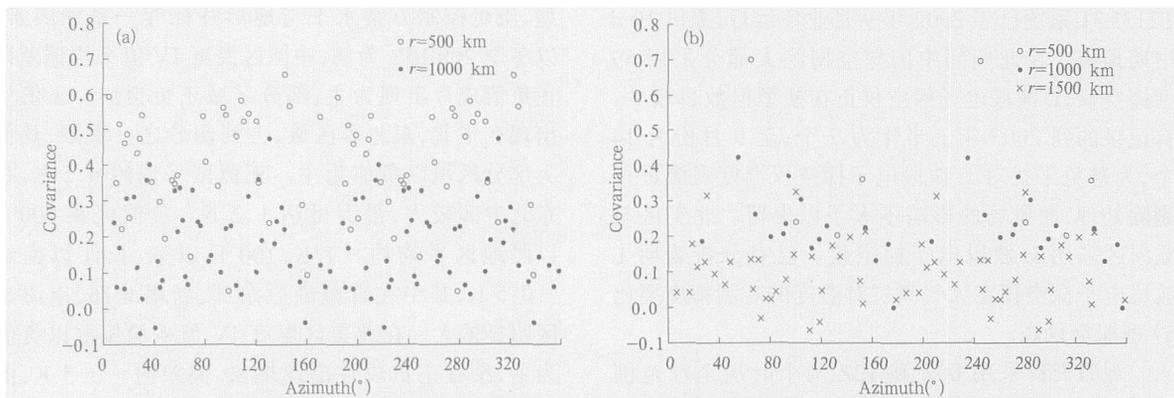


图1 冬季  $500\text{ hPa}$  温度 IV 值水平协方差随方位角和距离的变化特点 (a. 东亚区, b. 北美区)

Fig. 1  $500\text{ hPa}$  temperature IV horizontal covariance distributes with azimuth and distance in winter (a. East Asia, b. North America)

化时,可以认为是各向同性的。3 个区域站点对随方位角没有发现明显的变化趋势,表明各向同性在 3 个区域都近似成立。此外,一个非常明显的差异在于不同区域的站点对数目相差较远,中国区域、欧洲区要比北美区明显多,特别是比较近的站点对,如在 500 km 处中国区域的站点对约有 70 个,而北美区域不到 10 个。在 1000 km 处北美也不如中国区多,而到 1500 km 明显增多。北美区站点对稀少会影响到拟合结果的稳定性。

## 2.6 资料特征对拟合的影响

以上分析了 IV 方法所假定的无偏、均质和各向同性的要求在国家气象中心 T213L31 全球中期

分析预报系统中的区域差异和特征,可见这些假定在严格意义上不成立。相对而言各向同性假设成立最好,均质假设其次,无偏性假设较差。由于在统计时进行了 IV 值平均偏差值的扣除,这显然会影响到站点对间协相关的计算,特别是对东亚区更明显,但统计时不存在无偏性问题,因此主要是无偏扣除、均质假设与样本多寡对拟合的影响。欧洲区与北美区的 IV 平均值偏差的整体性特点与均质特点成立程度要略好于东亚区,且无偏扣除影响要小于东亚区,但东亚区与欧洲区样本数密集有利于拟合的稳定,而北美区样本数稀疏对拟合结果不利(图 2,图 3)。此外风场均质特征要好于温度场。

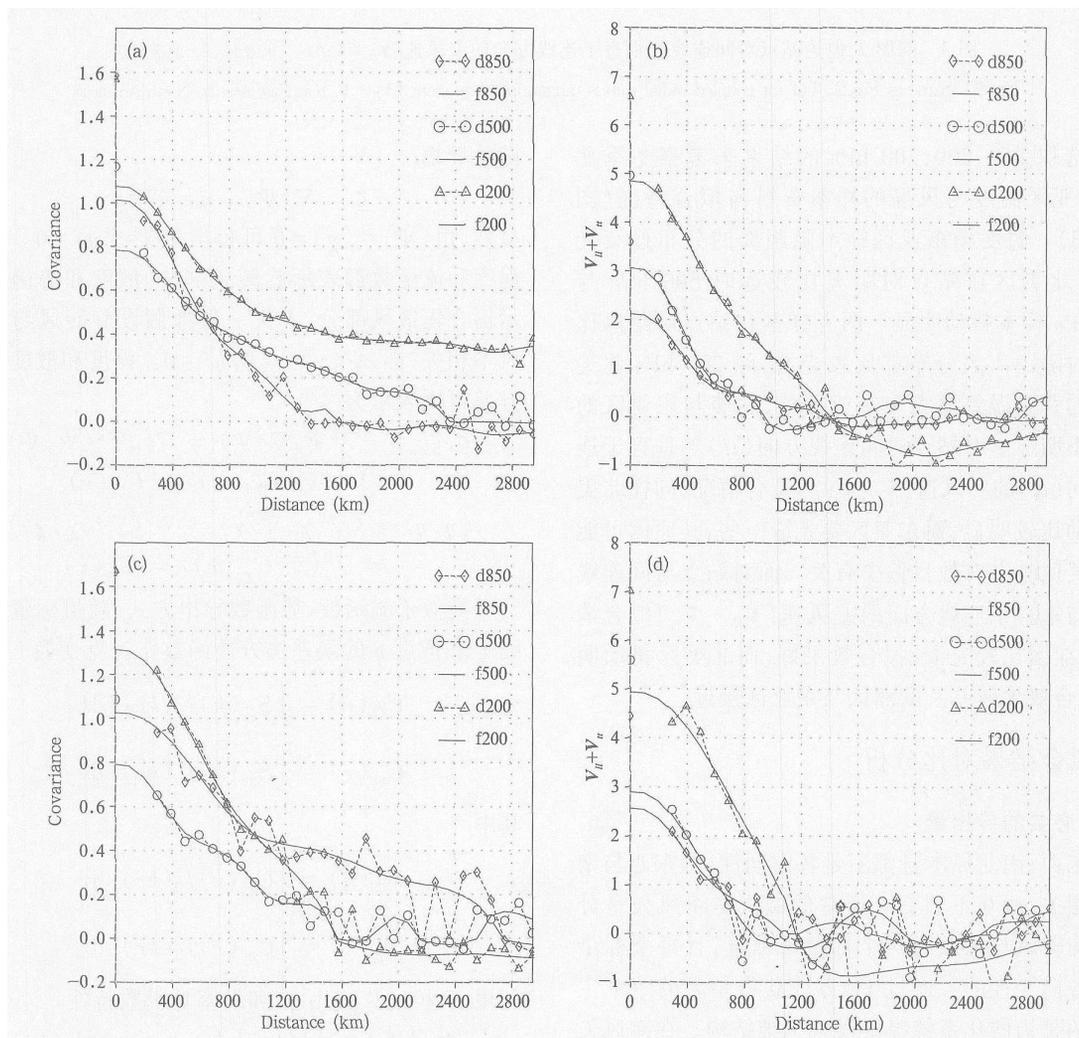


图 2 冬季样本资料对协方差(图标 d850, d500, d200)与拟合的预报误差函数曲线(图标 f850, f500, f200)(a. 东亚温度, c. 北美区温度, b. 东亚区站点对间连线与垂直于连线方向的耦合风速( $V_{||} + V_{\perp}$ )样本拟合, d. 同 b. 但为北美区)

Fig. 2 Binned innovation covariance (symbols for d850, d500 and d200) and fitted forecast error covariance functions (curves for f850, f500 and f200) in winter (a. temperature in East Asia, c. temperature in North America, b. coupled radial plus tangential components ( $V_{||} + V_{\perp}$ ) in East Asia, d. same as b. but for North America)

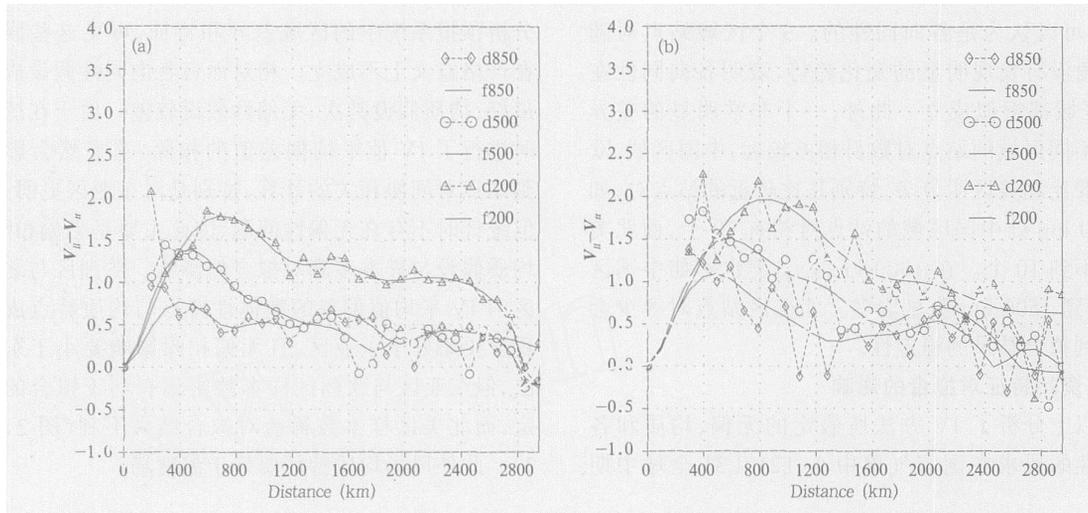


图3 同图2,但为站点对间连线与垂直于连线方向的差风速( $V_{II} - V_{II'}$ ) (a. 东亚区, b. 北美区)

Fig.3 Same as Fig. 2, but for coupled radial minus tangential component ( $V_{II} - V_{II'}$ ) (a. East Asia, b. North America)

选取 850, 500, 200 hPa 的结果来考察冬季北美、东亚区温度与风场的样本资料与拟合情况(图2,图3)。温度场东亚区样本随距离的分布连续光滑,而北美区在站点对距离比较远时 500 hPa 与 850 hPa 样本抖动明显。两者样本在站点对距离比较近时随距离的分布都比较光滑,除 200 hPa 北美温度的背景误差要大于东亚区外,北美与东亚区的差异不明显。对站点间连线方向( $ll$ )与垂直于连线方向( $tt$ )的和风速( $V_{ll} + V_{tt}$ )拟合情况,同样北美区抖动比较明显,而东亚区要光滑一些,其原因可能与北美区站点对数目偏少有关。而对站点间连线方向与垂直于连线方向的差风速( $V_{ll} - V_{tt}$ )拟合情况,东亚区比较光滑,拟合效果好,而北美区抖动明显,拟合效果较差。欧洲区与东亚区接近。

### 3 拟合结果对比分析

#### 3.1 必要的导出量

东西、南北风本身都不是各向同性,且东西与南北风相关,相互不独立。通常的做法是将风矢量分解为旋转风与辐散风两个标量来估算,且每个标量都是各向同性的。通过 IV 方法获得结果的一个主要目的是为同化系统提供误差三维结构。在对风矢量的分解表达中,除了采用流函数、势函数外,另一种主要表达方式是涡度与散度。本节首先给出涡度与散度的公式表达,可以作为 XW01 工作的补充。

随机标量场涡度 和散度 可以由流函数、势

函数导出

$$= \nabla^2, \quad = \nabla^2 \quad (1)$$

显然,由  $= = 0$  可得出  $= = 0$ 。由于涡度和流函数都表示旋转风部分,散度和势函数表示辐合辐散风部分,不失一般性假设旋转风与辐散风不相关,即  $, = , = 0$ 。涡度和散度的协方差可以表示为

$$, \quad = \nabla^2, \quad \nabla^2 = \nabla^2 \nabla^2, \quad = \nabla^2 \nabla^2 C(r) = C(r) \quad (2a)$$

$$, \quad = \nabla^2, \quad \nabla^2 = \nabla^2 \nabla^2, \quad = \nabla^2 \nabla^2 C(r) = C(r) \quad (2b)$$

类似于流函数、势函数导出方式,随机标量场涡度和散度的误差协方差函数作汉克变换

$$C(r) = S(k) J_0(kr) k dk \quad (3a)$$

$$C(r) = S(k) J_0(kr) k dk \quad (3b)$$

其中

$$S(k) = C(r) J_0(kr) r dr \quad (4a)$$

$$S(k) = C(r) J_0(kr) r dr \quad (4b)$$

利用 XW01(2.8a, b)式和贝塞耳函数特性

$$(d^2/dr^2 + r^{-1}d/dr) J_0(kr) = -k^2 J_0(kr)$$

得到

$$(d^2/dr^2 + r^{-1}d/dr) C(r) = -k^2 S(k) J_0(kr) k dk \quad (5a)$$

$$\begin{aligned} & (d^2/d r^2 + r^{-1}d/dr) C(r) = \\ & - k^2 S(k) J_0(kr) k dk \end{aligned} \quad (5b)$$

对式(5a,5b)再进行一次 Laplace 计算,可得:

$$\begin{aligned} & (d^2/d r^2 + r^{-1}d/dr)^2 C(r) = \\ & k^4 S(k) J_0(kr) k dk \end{aligned} \quad (6a)$$

$$\begin{aligned} & (d^2/d r^2 + r^{-1}d/dr)^2 C(r) = \\ & k^4 S(k) J_0(kr) k dk \end{aligned} \quad (6b)$$

功率谱系数为

$$\begin{aligned} k^4 S(k) &= (d^2/d r^2 + r^{-1}d/dr)^2 \cdot \\ & C(r) J_0(kr) r dr \end{aligned} \quad (7a)$$

$$\begin{aligned} k^4 S(k) &= (d^2/d r^2 + r^{-1}d/dr)^2 \cdot \\ & C(r) J_0(kr) r dr \end{aligned} \quad (7b)$$

由式(3a,3b)和式(6a,6b),以及 XW01(2.11ab)可以导出下述关系

$$S(k) = k^4 S(k) = k^2 S_{II}(k) \quad (8a)$$

$$S(k) = k^4 S(k) = k^2 S_{II}(k) \quad (8b)$$

其中  $l$  与  $t$  分别表示风场在任意两点间连线方向分量与垂直连线方向分量,从而有

$$C(r) = k^2 S_{II}(k) J_0(kr) k dk \quad (9a)$$

$$C(r) = k^2 S_{II}(k) J_0(kr) k dk \quad (9b)$$

对比 XW01 中的表达(2.15ab)推出流函数势函数表达式

$$C(r) = k^{-2} S_{II}(k) J_0(kr) k dk \quad (10a)$$

$$C(r) = k^{-2} S_{II}(k) J_0(kr) k dk \quad (10b)$$

可知两者在表达式中差了波数 4 次方一项。

将上式在有限半径  $r < D = 3000$  km,写成

$$C(r) = k_i^2 S_{II}(k_i) J_0(k_i r) \quad (11a)$$

$$C(r) = k_i^2 S_{II}(k_i) J_0(k_i r) \quad (11b)$$

当  $r=0$  时,利用特征长度的定义(XW01 5.5 式)有

$$\begin{aligned} C(0)/C(0) &= \\ k_i^2 S_{II}(k_i)/k_i^2 S_{II}(k_i) &= O(L^4) \end{aligned} \quad (12a)$$

$$\begin{aligned} C(0)/C(0) &= \\ k_i^2 S_{II}(k_i)/k_i^2 S_{II}(k_i) &= O(L^4) \end{aligned} \quad (12b)$$

由上式可见,流函数、势函数和涡度、散度方差

虽然在表达上非常接近,但不是简单的常数差别关系。其近似关系由 Wu 等<sup>[13]</sup>给出,表示为

$$C(0)/C(0) = L^4/8 \quad (13a)$$

$$C(0)/C(0) = L^4/8 \quad (13b)$$

对计算获得的背景误差协方差  $C(r_{ij}, p_m, p_n)$  与误差协方差功率谱  $S(k_i, p_m, p_n)$ ,对每个距离  $r_{ij}$  和每个波数  $k_i$  都是一个二维协方差矩阵,其中  $p_m$  是某一标准气压层。某一层的垂直协相关系数与垂直功率谱协相关系数  $r_{mn}$  为

$$\begin{aligned} r_{mn}(r_{ij}, p_m, p_n) &= C(r_{ij}, p_m, p_n)/ \\ &^{1/2}(r_{ij}, p_m, p_m)/^{1/2}(r_{ij}, p_n, p_n) \end{aligned} \quad (14a)$$

$$\begin{aligned} r_{mn}(k_i, p_m, p_n) &= S(k_i, p_m, p_n)/ \\ &^{1/2}(k_i, p_m, p_m)/^{1/2}(k_i, p_n, p_n) \end{aligned} \quad (14b)$$

式中  $\sigma$  为方差。

### 3.2 区域与季节变化特征

#### 3.2.1 均方差拟合特点

温度、风场均方差的区域和季节变化都比较明显,存在显著差异,图 4 仅给出东亚区冬季结果。在每个区域不同季节的垂直结构比较一致,但冬季的 IV 均方差都要明显大于夏季,这一特点在整层都比较一致。分别计算 3 个区域在冬夏两季平均值,取 100—850 hPa 的平均,温度场冬季偏大约 0.2 K,风场偏大 0.9 m/s。区域内季节变化可能是由于冬季的天气系统变化强烈而振幅大,长波槽活动非常活跃,而夏季的中小尺度对流活动要强一些,图 6 也表明背景误差功率谱大值区在夏季会向高波数区扩展。而不同区域间的变化比较复杂。因为本文目的是为了全球模式误差协方差校准提供气候基准值,为此计算了 3 个区两个季节共 6 个样本的平均值与均方差,从 IV 均方差平均值看,温度场在 500 hPa 为 1 K,向下增大至 1000 hPa 的 1.7 K,向上在 200 hPa 出现一个温度的凸点,其上的温度误差在 150 hPa 达到最小 1.3 K 后逐步增加。风场随高度增加并在 250 hPa 左右达到最大 6.5 m/s 后减小。将均方差相对于平均值结果作为变率,取 100—850 hPa 平均,温度场变率在 11%,这表明取平均值结果来代替某个区域结果可以接受。但在 100 hPa 以上变率逐步增大(约 20%),这个变率比较大,原因既和区域季节特点有关,也和拟合的结果可信度有关(图 2,图 3),在使用时要谨慎。风场在 200 hPa 以下变率在 11%,200 hPa 以上变率也达到了 17%。

不同区域和季节分离出来的背景误差与观测误差均方差差异比较大(图略)。从冬夏区域层次平均值来看,对温度场观测误差冬季仅比夏季偏大0.1 K,而背景误差要偏大约0.2 K,对风场观测误差冬季比夏季偏大0.7 m/s,背景误差偏大0.6 m/s。可见温度场IV均方差主要来自背景初猜场,而风场观测与背景初猜场都重要。计算分离出来的背景误差与观测误差比值的平均值,温度场在400 hPa以下与150 hPa以上,背景误差略小于观测误差,而在200—300 hPa,背景误差略大一些。风场的特点与温度场比较一致,但在400 hPa以下的低层,背景误差要明显小于观测误差。比值的变率也在20%—30%,这表明不同区域季节背景误差与观测误差在

IV方差中的比例变化十分显著。观测误差与背景误差的变率特点并不一致,主要表现在其变率要明显小于背景的变率。观测误差的变率在10%左右,与IV均方差的变率接近,而背景场误差的变率在18%左右。温度观测误差平均值在500 hPa左右达到最小约为0.8 K,向下至1000 hPa增大到1.3 K,向上100 hPa增大至1.1 K。风场观测误差在300 hPa左右达到最大约4.8 m/s,向下至1000 hPa减小为2.8 m/s,100 hPa为3.6 m/s。温度与风场观测误差的平均值与业务参数相当。涡度的分布与风场非常相似,但变率比风场略大为19%。散度场明显小于涡度场,但变率达到39%,平均值基本不可用。

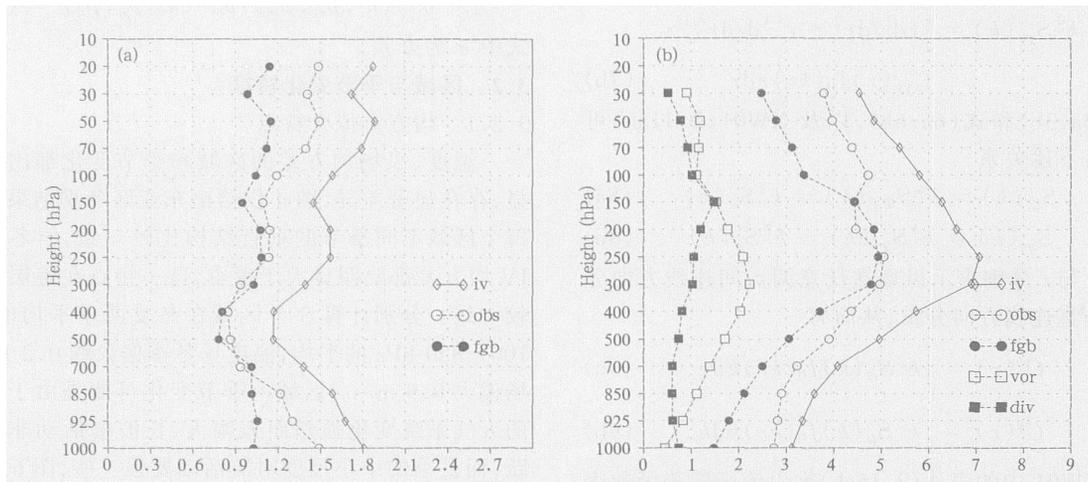


图4 东亚区冬季温度与风场IV均方差拟合结果

(a. 温度, b. 标量风场; 图标中的均方差依次为IV值(iv), 观测(obs)、背景(fgb), 涡度(vor)与散度(div); vor和div单位:  $10^5 \text{ s}^{-1}$ )

Fig. 4 Fitted innovation standard deviations of East Asian temperature and wind in winter

(a. temperature, b. scale wind; symbols for innovation (iv), observation (obs), background(fgb), vorticity (vor) and divergence (div) standard deviation; vor and div units:  $10^5 \text{ s}^{-1}$ )

3.2.2 背景误差水平相关系数的变化特点

对T213L31约60 km分辨率的模式而言,温度、风场、涡度、散度的特征长度依次为230, 200, 140, 140 km左右(图5)。拟合出的特征长度在垂直方向变率很大,出现明显的“之”字型,对明显的错误情况需要进行质量控制予以剔除,质量控制阈值取为变率的2倍均方差。显然东亚区与欧洲区的特征长度估计结果比较稳定,而北美区的结果要差一些。这是因为,特征长度由误差功率谱特点来决定,特别是长波部分有明显的直接影响。北美区在站点

对距离  $r$  比较大时样本点的抖动比较明显(图2,图3),这表明远距离的相关不稳定,这会影响到特征长度。而误差方差的决定主要是站点对距离  $r$  比较小的样本决定,在这些距离上3个区域的样本点都比较稳定一致。

特征长度的季节变化不明显。对质量控制后的特征长度计算季节与区域的平均值与均方差,将均方差相对于平均值结果作为变率,取100—850 hPa平均为参考值,高度场的变率比较大在20%左右,而温度场为12%明显小于高度场。涡度场的变率

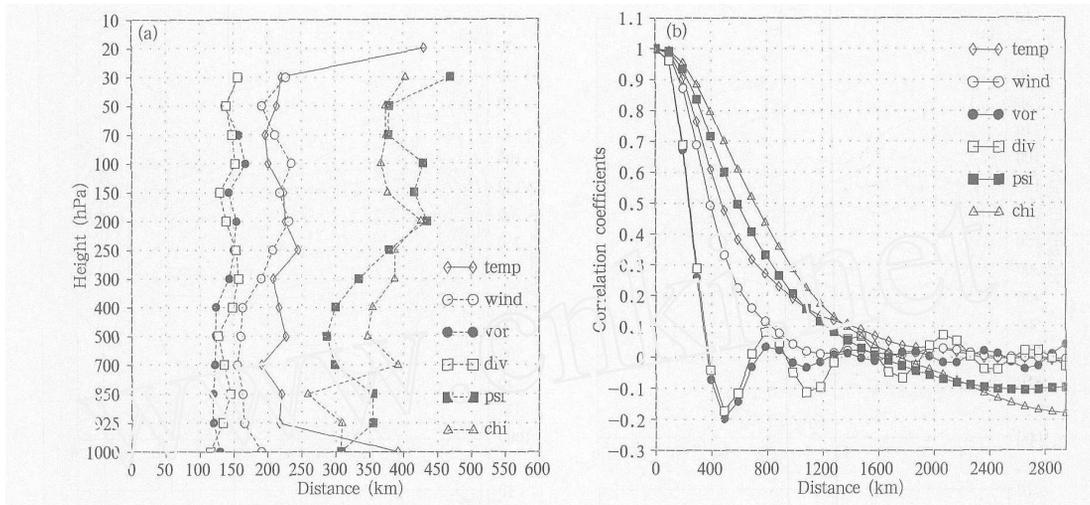


图 5 东亚区冬季背景误差特征长度的变化(单位:km) (a. 特征长度垂直分布, b. 500 hPa 水平相关系数; 图标依次为温度(temp)、风场(wind)、涡度(vor)、散度(div)、流函数(psi)、势函数(chi))

Fig. 5 East Asia background error de-correlation lengths in winter (units: km) (a. vertical distribution, b. 500 hPa horizontal correlation coefficients. Symbols are temperature (temp), scale wind (wind), vorticity (vor), divergence (div), stream function (psi) and potential velocity (chi))

为 8% 比风场 (10%) 略小, 散度场变率比较大达到 17%, 涡度的变率明显小于流函数 (12%) 和势函数 (12%)。这表明特征长度在全球不能变化时, 选择温度、涡度和散度作为分析的控制变量时, 利用 IV 方法对背景误差特征长度的估计要明显好于利用高度场、流函数和势函数。

温度场的特征长度基本上不随高度变化, 而风场在 200 hPa 以下基本为 170 km 左右, 在 200 hPa 以上为 200 km 左右, 有明显的特征长度变长的特点, 这和平流层气流特点有关。涡度也有类似的特点, 在 200 hPa 以下为 130 km 左右, 以上为 145 km, 而散度场随高度的变化不明显, 都在 150 km 左右。从 500 hPa 的水平截面来看 (图 5b), 温度、风场以正相关为主, 而涡度散度出现从负值趋近零的特点, 这和 Daley 等<sup>[21]</sup>的结论一致。从垂直切面来看 (图略), 负值区结构比较深厚, 且 200 hPa 以下整层都比较均一。为了利用 Wu 等<sup>[13]</sup>公式计算涡度散度方差, 同时给出了流函数和势函数的特征长度。

### 3.2.3 背景误差方差功率谱

从贝塞尔函数的波谱 (局地波数) 特征来看, 对误差相关的贡献在 0—7 波谱。以截断范围  $D = 3000$  km 计算, 10 波对应的波长已小于 300 km, 小于现有探空网密度, 因此拟合时截断 10 个波谱已经足够了。根据局地波数与全球波数的关系 (见 XW01 式 (5.2)),  $K_i / (K_i + 1) = (ak_i / D)^2$ ,  $i = 1$  将波

数转换为全球波数, 其中  $a$  是地球半径,  $K_i$  是全球波数。以东亚区为例 (图略), 冬季温度背景误差最大值区域在 250 hPa 左右, 这正是温度凸点的位置, 在 500 hPa 高度背景误差最小, 在 850 hPa 以下的大气低层出现另一个大值区。背景误差功率谱在 20 波左右达到最大, 且夏季比冬季的功率谱大值区明显向高波数区扩展, 而 40 波以后拟合噪音太大不可用。可以发现相同季节不同区域差别较大, 以夏季为例基本上背景误差沿功率谱的分布特点比较接近, 其中北美区以 20 波左右为中心出现大值区, 且大气低层的误差更强, 而欧洲区的大值中心也出现在 20 波左右, 强度要明显强于其他两个区域。

风场分布也有类似特点 (图 6), 冬季方差要明显大于夏季。此外夏季方差大值区要比冬季更向高波数区扩展, 冬季大值区在 20 波左右, 而夏季在 25 波左右, 这和夏季短波活动比冬季强有关。风场大值区的高度在 200 hPa 左右, 这和方差的廓线非常一致。风场的拟合噪音不如温度场那么大。在不同区域 (图略), 方差分布特点比较接近, 但方差差别较大, 以东亚区最大欧洲区最小。特别值得注意的是欧洲区的分布与其他两个区域不同, 高值区不但低 (在 250 hPa) 而且向高波数区扩展的更为明显, 大值区在 35—40 波。从涡度散度的分布来看, 相对于风场大值区更加靠近高波数, 而且涡度比散度略大一些。此外涡度的大值区要比风场低, 在 300—250 hPa, 大值

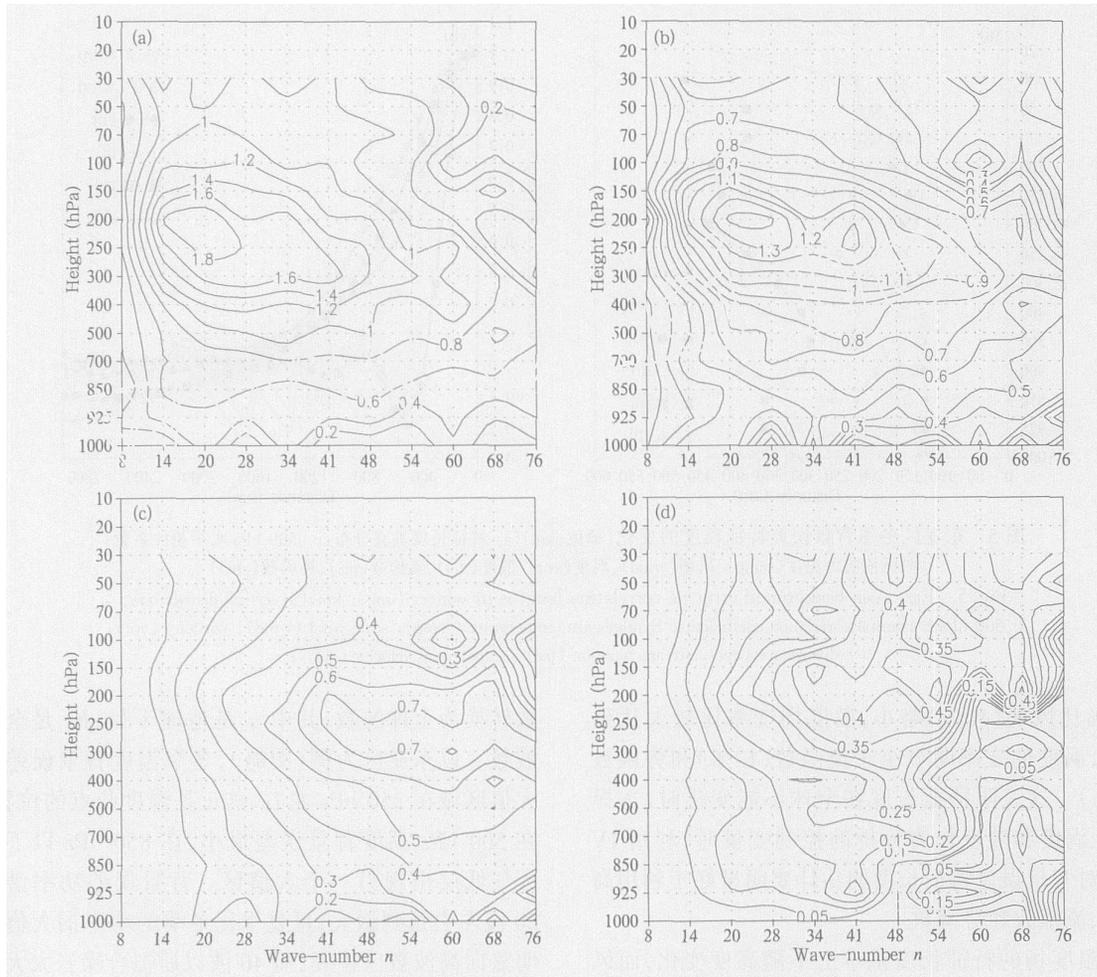


图 6 东亚区夏季风场背景误差功率谱沿二维总波数  $n$  分布特点

(a, b. 标量风, 但 a 为冬季, c. 涡度, d. 散度)

Fig. 6 Energy spectra of East Asia summer wind background errors plotted in section of height and 2-dimensional total wave number  $n$  (a, b. scale-wind, but a. for winter, c. vorticity, d. divergence)

区更偏向短波方向,在 40 波之后;而散度的大值区要比风场高,在 200—150 hPa,这表明旋转风与辐散风的特点并不一致。此外散度场的噪音要明显大于涡度场。

### 3.2.4 背景误差协方差特征

利用式(14a)计算格点空间的背景误差垂直相关系数(图 7),对温度场不同的高度垂直相关的范围都比较窄,但北美区(图略)要明显宽于东亚区,与欧洲区接近。在垂直相关上高度场与温度场差异显著,高度场相关范围要明显大于温度场。值得注意的是在对流层顶位置,在垂直方向出现明显的负相关分布特点,最大负相关的位置存在区域差异,如东亚区、北美区在 150 hPa 的位置,而欧洲区在 200 hPa 的位

置。负相关的存在和对流层顶的存在有关。相对而言,风场的垂直相关范围要比温度场大得多,而且欧洲区要比东亚和北美区更宽广,东亚区最窄。风场的垂直相关以正相关为主,表明在低层的整层一致性比较好。在 200 hPa 以上层次,风场垂直相关要窄得多,表明在平流层不同层次之间的误差相关并不明显。从风场的分量来看,涡度的特点与风场非常一致,欧洲、北美宽而东亚窄。散度场的差别则要大得多,在垂直方向出现明显的负相关,且相关值比较大,说明由于大气的连续性在某层辐合辐散对应其上下层的辐散辐合,但负相关区计算的完整性和连续性在不同区域差别较大。总之从垂直相关特征来看,不同区域的差别非常小,与 XW01 的结果非常

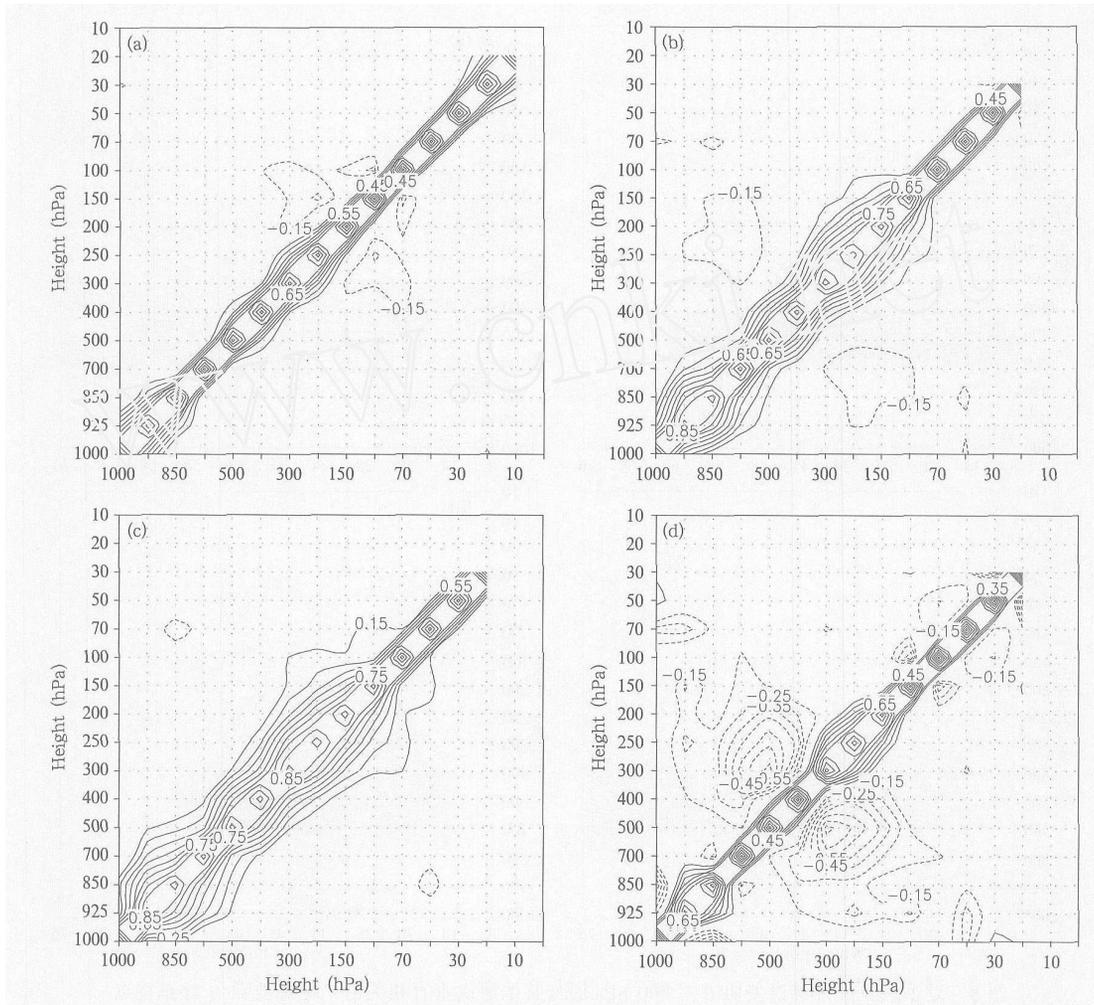


图 7 东亚区夏季背景误差协方差垂直相关系数(a. 温度, b. 标量风, c. 涡度, d. 散度)

Fig. 7 Vertical correlations of background errors in summer in East Asia

(a. temperature, b. scale-wind, c. vorticity, d. divergence)

类似(XW01, 图 7), 说明在垂直方向上不同来源资料的统计结果差异不大, 其原因很大程度上是由于标准等压面层次固定, 垂直分辨率不变。

利用式(14b)在谱空间计算背景误差垂直相关, 以 500 hPa 为例(图 8), 温度的相关在 20 波以内略宽, 特别是和高层的相关比较宽广, 20 波以后随波数相关的特点变化不大。此外 250 hPa 以上负相关区在所有波数都存在。风场要比温度场相关宽广得多, 同样是在前 20 波相关略宽, 20 波以后变化不大, 但风场在高层没有明显的负相关区, 这和温度场有显著区别。涡度分布与风场比较接近, 但散度的特点与涡度差异非常大, 主要表现在不同的波数都呈现出了非常明显的负相关结构。欧洲和北美区的垂直相关范围要比东亚区宽广得多(图略)。

### 3.3 涡度散度两种表达的差异

由于涡度和散度背景误差方差在后文中的应用非常关键, 本文分别采用近似估算方法(式(13a, 13b))和直接利用贝塞尔函数拟合(式(11a, 11b))方法进行估计, 并计算两种估计结果的比值, 并对比值缩小根号 2 倍(图 9 中略去比值大于 3 的结果)。可见比值在 1 附近, 这表明对涡度场两种方法除有根号 2 倍差别外, 结果相当一致。近似估计方法高估了涡度与散度。而散度场噪音要比涡度场大许多, 这和前文中利用贝塞尔函数拟合方法估计散度误差较大的结论一致。

## 4 结果和讨论

对大气的最优估计可以表达为求解下式<sup>[3, 15]</sup>

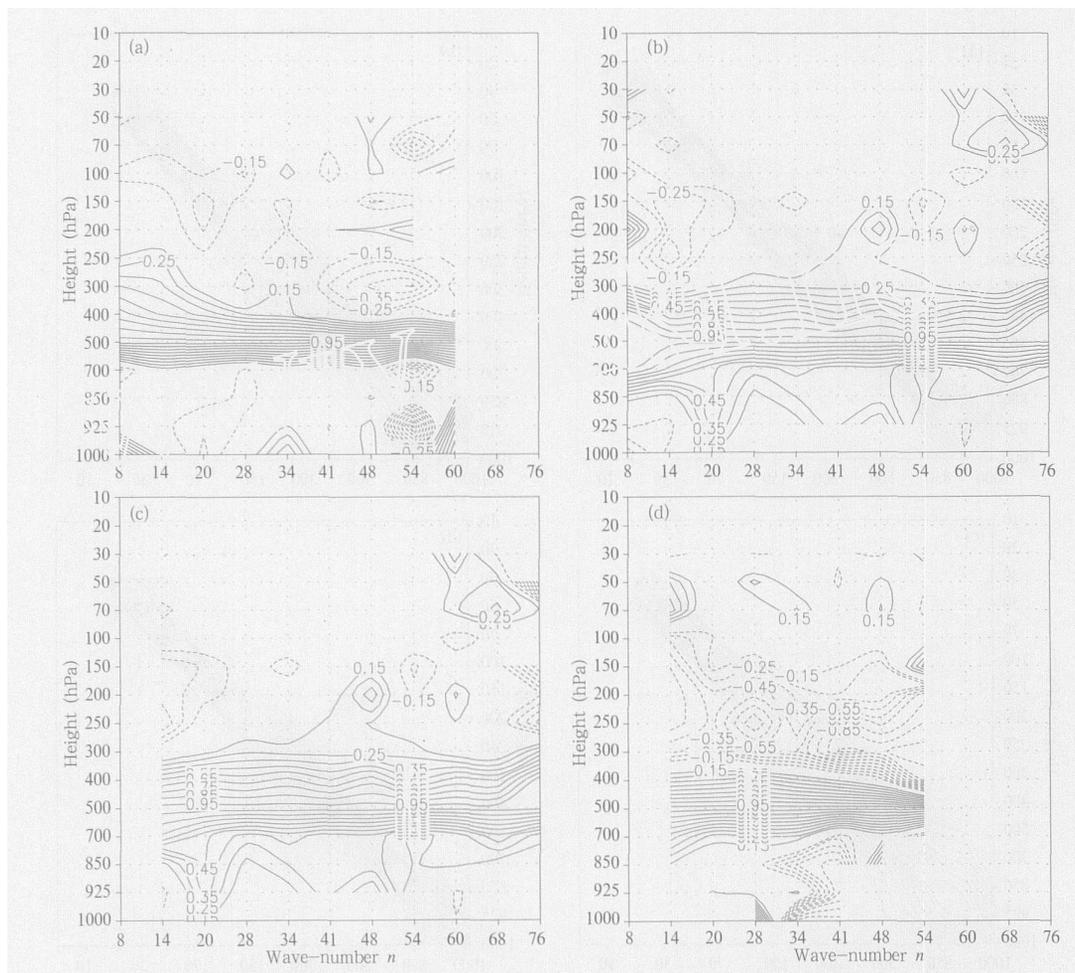


图 8 夏季东亚区背景误差协方差 500 hPa 层与其他层次垂直相关沿二维总波数  $n$  分布特点  
(a. 温度, b. 标量风, c. 涡度, d. 散度)

Fig. 8 Vertical correlations of background error covariances between 500 hPa and other levels plotted in section of height and 2-dimensional total wave number  $n$  in East Asia in summer  
(a. temperature, b. scale-wind, c. vorticity, d. divergence)

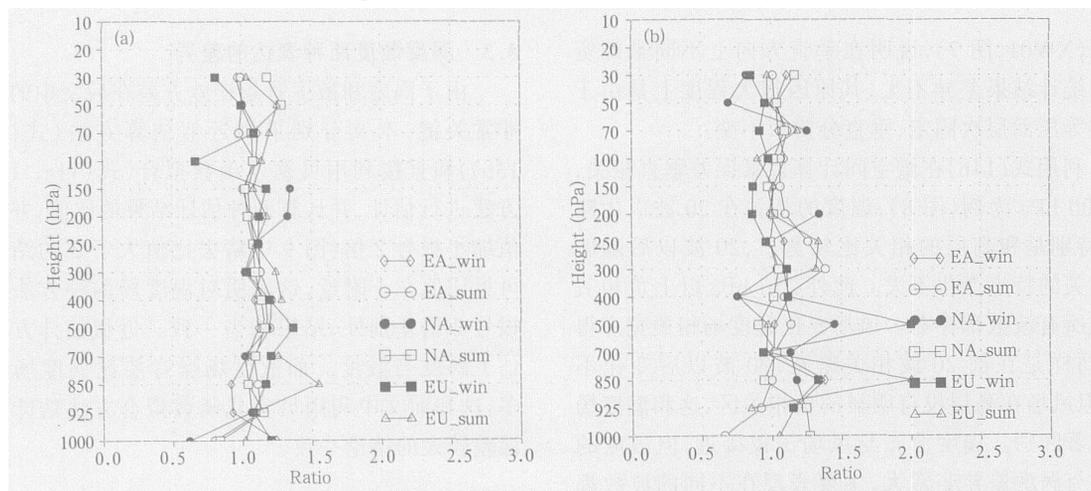


图 9 东亚(EA)、北美(NA)和欧洲区(EU)利用式(11a, 11b)及式(13a, 13b)计算得到的涡度散度之比  
(比值已经缩小 $\sqrt{2}$ : a. 涡度, b. 散度, 图标为冬季(win)与夏季(sum))

Fig. 9 Ratio between values calculated by Equations (11a, 11b) and by to Equations(13a, 13b), and values rescaled by square root of 2 in East Asia (EA), North America (NA) and Europe area (EU) (a. vorticity, b. divergence)

$$J(x_a) = \frac{1}{2} (x_a - x_b)^T B^{-1} (x_a - x_b) + [y - H(x_a)]^T R^{-1} [y - H(x_a)] \quad (14)$$

式中  $x_a$  为最优估计值或分析场,为列向量,  $x_b$  为背景初猜场,  $B$  为背景误差协方差矩阵,上标  $T$  表示矩阵转置,  $y$  为观测值,  $R$  为观测误差,一般只考虑方差项。  $H$  是从模式空间到观测空间的观测映射算子,一般为非线性表达式。利用最优控制理论按迭代方法求解上式就是一般意义上的三维变分同化方法。当  $H$  是线性算子时,上式可以表述为最优无偏估计的形式<sup>[3]</sup>

$$x_a = x_b + B H^T (B H^T + R)^{-1} [y - H(x_b)] \quad (15)$$

对上式中的背景误差协方差矩阵  $B$  进行近似处理,写成  $B = D^{T/2} C D^{1/2}$  形式,其中矩阵  $C$  为多变量协相关系数矩阵,  $D$  为变量的方差,考虑  $D$  的时间演变,并利用局部资料求解分析场,就是一般意义上的多变量最优插值方法。

可见,最优插值与三维变分在观测算子  $H$  为线性算子时基本等价,在常规观测资料相同时决定分析场差异的主要因素是背景误差协方差和观测误差的不同,这也表明基于最优插值的误差统计结果同样适用于三维变分。作为发展国家气象中心全球三维变分同化系统的第一步,本文采用 XW01 发展的误差协方差三维估计方法,利用 T213L31 全球中期分析预报系统的预报结果和探空观测,来估计北美、东亚、欧洲大陆等资料密集区的温度、风场误差结构特征,除修正现有业务中资料同化系统中的误差参数外,估计结果也将主要作为标准值来校准全球三维变分同化系统的误差参数。同时,国内中尺度资料同化系统多采用 T213L31 全球中期分析预报系统 6 h 的预报场作为冷启动的初猜场,东亚区的背景误差协方差与观测误差有很好的指导意义。

相对于区域分析预报系统,对全球三维变分同化系统涉及到的下垫面比较复杂,资料分布均匀性不同,涵盖赤道与中高纬、陆地与海洋以及季节的变化,也即观测误差与背景误差非均质特点。需要分析不同区域的季节变化特点。结果表明,观测空间拟合方法所要求的水平均质、各向同性在北美区和欧洲区成立程度较好,在东亚区略差。此外方差在冬季略大于夏季,区域差别较大。温度与风场背景误差主要集中在前 40 波,并在 20 波左右达到最大,

水平相关季节区域差别不大,而垂直相关比较窄,且垂直相关范围比较宽广的波数主要集中在前 20 波。

此外利用贝塞尔函数拟合方法获得结果的分析表明,在质量场中不同区域季节间温度误差的稳定性要明显好于高度场,涡度散度的稳定性要明显好于流函数和势函数,特别是对于特征长度更为明显。这表明利用贝塞尔函数拟合方法获得的结果对校准在全球资料同化中采用温度、涡度散度作为资料同化的分析变量具有一定的优势。多角度分析表明,虽然不同区域、季节有差别,但对这 3 个纬度比较接近的区域,其结果仍然有较大的可比性,利用区域、季节的平均结果可以获得很好的近似。

但 IV 方法所要求的观测系统均一、均质与各向同性假设只能在北美、欧洲和东亚等几个区域基本成立,每个区域范围都在 3000—4000 km 以内。它主要给出中高纬 20°—70°N 范围内的大陆资料均匀密集区平均观测误差与背景误差协方差,不能给出大洋、赤道、南半球、极区等资料稀疏区的结果,也不能给出方差变化的非均质特征。3000—4000 km 只能描述 10 波以上的波谱特征,此外 IV 方法依据的主要是常规探空资料,资料间隔在 300—400 km 以上,只能给出 100 波以内的波谱特征。对于全球模式 T213L31 而言,总波数 214 波,有一半的波不能被 IV 方法所描述,而在背景误差协方差矩阵中需要全部波谱的信息。此外在垂直方向上 17 层标准等压面间距过宽,垂直相关不高,IV 方法不能揭示出太多的垂直协相关信息,对于密得多的模式层次,IV 方法提供的指示意义有限。IV 方法特别是对  $B$  矩阵多变量之间协方差的估计能力缺乏,因此需要将 IV 方法与其他估计方法结合起来一起应用。

致谢:本文原始程序由美国强风暴实验室许秦博士提供,在此谨向他表示感谢!

## 参考文献

- [1] 王建捷,陈起英,姚明明等. 我国中期数值预报又迈上一个新台阶——T213L31 数值预报系统投入业务运行. 北京:国家气象中心年报,2002. 7-8  
Wang Jianjie, Chen Qiyang, Yao Mingming, et al. CMA Medium Range Numerical Prediction Takes a New Step The Operational Running of the T213L31 Numerical Prediction System. Beijing: National Meteorological Center Newsletter. 2002. 7-8

- [2] Daley R. Atmospheric Data Analysis. Cambridge University Press, 1991. 457 pp
- [3] Stephen E C. An introduction to estimation theory. J Met Soc Japan, 1997, 75B: 257-288
- [4] Gandin L S. Objective Analysis of Meteorological Fields (in Russian). Israel Program for Scientific Translation, 1965. 242pp
- [5] Rutherford I D. Data assimilation by statistical interpolation of forecast error fields. J Atmos Sci, 1972, 29: 809-815
- [6] 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
- [7] 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
- [8] Dee D P, Da Silva A M. Maximum-likelihood estimation of forecast and observation error covariance parameters, Part I: methodology, Mon Wea Rev, 1999, 127: 1822-1834
- [9] 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
- [10] Xu Q, Wei L. Estimation of three-dimensional error covariances, Part II: analysis of wind innovation vectors. Mon Wea Rev, 2001, 129: 2939-2954
- [11] 国家气象中心编译. 资料同化和中期数值预报. 北京: 气象出版社, 1991. 436pp
- National Meteorological Center. Data Assimilation and Medium Range Numerical Prediction. Beijing: China Meteorological Press, 1991. 436pp
- [12] 陈德辉等. 下一代数值天气预报业务系统的开发及应用论文集. 国家气象中心“九五”科技攻关项目, 2000. 380pp
- Chen Dehui, et al. Next Generation Numerical Weather Operational Prediction System Development and Application. National Meteorological Center “95” Scientific Key Project, 2000. 380pp
- [13] Wu W-S, Purser J, Parrish D. Three-dimensional variational analysis with spatially inhomogeneous covariances, Mon Wea Rev, 2002, 130: 2905-2916
- [14] Rabier F, McNally A, Anderson E, et al. The ECMWF implementation of three-dimensional variational assimilation (3D-VAR). II: Structure functions. Quart J Roy Met Soc, 1998, 124: 1809-1829
- [15] Lorenc A C. Analysis methods for numerical weather prediction. Quart J Roy Met Soc, 1986, 112: 1177-1194

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

### Part : Accurate Estimation of Error Covariance in Observation Space

Gong Jiandong<sup>1</sup> Wei Li<sup>2</sup> Tao Shiwei<sup>1</sup> Zhao Gang<sup>1</sup> Wan Feng<sup>1</sup>

<sup>1</sup> National Meteorological Centre, Beijing 100081

<sup>2</sup> Cold and Arid Regions Environmental and Engineering Research Institute,  
Chinese Academy of Sciences, Lanzhou 730000

#### Abstract

Observation and background error covariances play important roles in the four-dimensional data assimilation and operational data analysis system, and they determine the relative importance of observation and background information, and the spread of these information in the grid space and among various control variables. Because the true atmospheric state is unknown, it is necessary to develop some techniques to estimate observation and background error covariances. Among these techniques, the method which uses innovation vector (observation minus background in the observation space) to partition observation and background error covariances is more accurate than other methods; and its estimated observation errors can be directly used in the data analysis system, and its estimated background error covariance can be used as a benchmark to tune and verify the results from other techniques, such as so called NMC method. Using National Meteorological Centre global medium-range analysis and forecast system T213L31's 6 hour forecasts as a background, and winter/summer two seasonal radiosonde observations in the Northern Hemisphere, the Bessel fitting function is used to partition observation and background error covariances. The analysis and comparison are performed during winter and summer periods, and among the data from three dense radiosonde observations regions of East Asia, North America, and Europe. The results show that the requirements for Bessel fitting method, such as horizontal homogeneity and isotropy, are roughly satisfied in all three regions, relatively better in North America and Europe, but slightly worse in East Asia. The error variances change greatly with different regions and seasons. For example, for temperature and scale-wind in winter those values are about 0.2 K and 0.9 m/s larger than those in summer, respectively. Below 400 hPa and above 150 hPa the temperature background error variances are smaller than the observation error variances, but within 200 - 300 hPa, the background error is slightly larger. The characteristics of winds are similar with temperatures. In the spectral space, the major temperature and wind background errors are distributed within wave-number 40, with the maximum in about wave-number 20. The horizontal de-correlation length is temporally, spatially stable for all variables, but the vertical de-correlation length of temperature is relative narrower than that of wind, and their main vertical correlations are located within wave-number 20. The analysis results also show that the background error covariance structure for temperature is clearly more stable than that of height, regardless of different seasons or in different regions; and for wind field the covariance structures for vorticity and divergence are clearly more stable than those of streamfunction and velocity potential function, especially for horizontal de-correlation length. Those results are useful for tuning and verification of the global data assimilation which adopts temperature, vorticity and divergence as analysis variables.

**Key words:** Observation error, Background error covariance, Bessel function fitting, Error separation, Accurate estimation.