

空气质量数值模式预报中资料同化的初步研究

崔应杰^{1,2} 王自发^{1,2} 朱江³ 付晴艳⁴

1 中国科学院大气物理研究所大气边界层物理与大气化学国家重点实验室, 北京 100029

2 中国科学院大气物理研究所竺可桢-南森国际中心, 北京 100029

3 中国科学院大气物理研究所国际气候与环境科学中心, 北京 100029

4 上海市环境监测中心, 上海 200030

摘要 利用嵌套网格空气质量预报模式, 在对上海市环境监测中心提供的观测数据进行必要的质量控制后, 采用最优插值方法对可吸入颗粒物(PM₁₀)、二氧化氮(NO₂)和二氧化硫(SO₂)进行资料同化。选取2004年8月1~20日做作逐日同化试验的结果表明, 无论是PM₁₀、NO₂还是SO₂, 其同化偏差平均值均在20 $\mu\text{g} \cdot \text{m}^{-3}$ 以下, 比同化前减少了至少50%; 3种污染物的同化偏差小于其未同化偏差的天数均在16天以上。在大气清洁和污染两种情况下, 对PM₁₀分别作10天的同化试验表明, 同化后的均方根误差均小于同化之前。此同化方法能利用观测数据较好地修正空气质量模式预报场, 从而为模式提供与实际更加接近的初始场。

关键词 大气化学 空气质量数值模式预报 资料同化 最优插值 质量控制

文章编号 1006-9585 (2006) 05-0616-11 **中图分类号** X16 **文献标识码** A

A Preliminary Study on Data Assimilation for Numerical Air Quality Model Prediction

CUI Ying-Jie^{1,2}, WANG Zi-Fa^{1,2}, ZHU Jiang³, and FU Qing-Yan⁴

1 *The State Key Laboratory of Atmospheric Boundary Layer Physics and Atmospheric Chemistry, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029*

2 *Nansen-Zhu International Research Center, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029*

3 *International Center for Climate and Environment Sciences, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029*

4 *Shanghai Environmental Monitoring Center, Shanghai 200030*

Abstract Observations of PM₁₀, NO₂ and SO₂ by Shanghai Environmental Monitoring Center with necessary quality control were assimilated into the nested air quality prediction model, using the optimal interpolation approach. The assimilation performed day by day from August 1st to 20th 2004 shows that for all of PM₁₀, NO₂ and SO₂, the mean assimilating departures are less than 20 $\mu\text{g} \cdot \text{m}^{-3}$, decreasing to at least 50 percent of those without assimilation. Also, the bias errors of the three pollutants with assimilation are smaller than those without assimilation for more than 16 days. The assimilation of PM₁₀ for 10 days from both of clean air condition and polluted air condition shows that the root mean square (RMS) errors with assimilation are less than those without assimilation. It is illustrated that this assimilation method can improve a large extent the simulation of urban air quality model using observations so as to afford initial condition more close to the true values.

收稿日期 2005-05-08 收到, 2006-03-30 收到修定稿

资助项目 中国科学院引进国外杰出人才项目“沙尘输送及其气候环境影响”和国家自然科学基金资助项目 40305018

作者简介 崔应杰, 男, 1981年出生, 硕士生, 主要从事大气化学资料同化的研究 E-mail: cuiyingjie2008@hotmail.com

Key words atmospheric chemistry, nested air quality prediction model, data assimilation, optimal interpolation, quality control

1 引言

近年来,城市扩大化和农村乡镇化造成大气污染日益严重,其中对人体健康有危害的大气污染物主要有 NO_2 、 SO_2 、 O_3 和 PM_{10} 等,已影响到国家经济、文化和环境等的可持续发展。因此,深入了解大气环境的状况,开展城市空气质量预报十分必要。城市空气污染预报将同天气预报一样,有利于指导人们安排日常工作和生活,避免或减轻空气污染的损害。

目前国际上对大气污染物浓度预报的方法有两种类型:一种是统计预报,另一种是数值预报。数值预报是以大气动力学理论为基础,建立大气污染物在空气中的稀释扩散数值模型,通过计算机高速计算来模拟和预测大气污染物在空气中的动态分布。空气污染的数值预报,大多数国家还停留在研究阶段,能直接开展日常业务预报的较少。我国北京、上海等城市已经初步进行空气质量数值预报的业务化试验。

对相关大气污染物的监测和数值模式在当前大气化学的研究中均扮演着重要的角色。观测能客观地记录真值,但其时空分布不规则,且观测数量有限。而通过数值模式模拟可以获取规则时空分布的结果,且其时空分布密度较大,但其与真值始终有一定的差距,因此无法代替观测的角色。将模式预报结果与观测信息结合起来的一种客观方法即为资料同化。在数值预报中它主要是为模式提供较优的初始场和再分析资料集。

资料同化在气象上的研究和应用始于 20 世纪 60 年代末^[1],而针对大气化学模式的资料同化方法研究是从 90 年代才开始^[2]。目前国际上开展大气化学资料同化研究的主要有美国^[3]、德国^[4]、荷兰^[5]、法国^[6]和欧洲中期天气预报中心^[7]等国家和组织。Austin^[2]在全球二维化学输送模式中对卫星观测的 O_3 、 HNO_3 、 NO_2 和 H_2O 浓度进行同化, Fisher 等^[7]用四维变分技术通过光化学箱体模式来同化卫星观测资料,利用 Lagrangian 方法将动力过程和化学过程分离,从而减少计算代

价。Levelt 等^[8]在二维全球平流模式中同化 TOVS (TIROS operational vertical sounder) 卫星观测的臭氧总浓度。Lyster 等^[9]采用 Kalman Filter 法根据卫星观测值同化全球大气输送模式计算的 CH_4 浓度。

近年来,区域大气化学模式伴随空气污染和控制问题的研究也逐渐发展起来。Umeda 等^[10]通过对中尺度气象场的同化来分析臭氧浓度对气象场的敏感性发现,臭氧浓度并没有显著的改善。Alberto 等^[11]在光化学模式应用四维资料同化技术来修正模式中亚特兰大的排放总量,对模拟区域人为排放的氮氧化物 (NO_x)、可挥发性有机物 (VOC) 和一氧化碳 (CO) 以及由生物源产生的 VOC 排放量进行评估,在此基础上对排放源进行适当的调整。Remus 等^[12]采用两种 Kalman 滤波技术对欧洲的 135 个固定地面观测点所提供的臭氧观测进行同化并比较其效果。

国内区域及城市空气质量预报模式的研究正处于发展阶段,中国科学院大气物理研究所、中国气象局、南京大学等开发了各种尺度的空气质量模式预报系统,在空气质量模式预报系统中对大气化学物种的同化研究尚未起步。结合上述大气化学模式资料同化及海洋模式资料同化工作的初步成果^[13,14],本文尝试探讨空气质量数值预报中对大气化学成分的同化研究,以期能应用于业务预报。

2 同化方案

常用的资料同化方法主要包括最优插值、三维变分、四维变分、Kalman 滤波和集合 Kalman 滤波等。最优插值是在 20 世纪 50 年代由 Eliassen^[15]提出,由于其计算量小,曾经在气象上得到广泛应用。80 年代 Lewis 和 Derber^[16]提出四维变分伴随同化方法。变分法以数值模式为其动力学限制,通过将预报值与观测值之间的距离(即目标泛函 J)最小化来得到最优的初估场。如果目标泛函 J 定义在三维空间(不包括时间维)上,则对应三维变分问题;若 J 定义在四维空间(包括

时间维), 则为四维变分问题。与最优插值或三维变分相比, 四维变分能利用未来的观测信息, 其缺点是需要给出切线性模式和相应的伴随。理论上的优越性使之成为当前广泛采用的资料同化方法之一。集合 Kalman 滤波 90 年代由 Evensen^[17] 首先提出, 利用了 Kalman 滤波中预报背景场(即模式预报结果)误差的思想, 放弃 Kalman 滤波中的协方差矩阵预报模式, 改为直接利用 Monte-Carlo 方法来多次积分模式, 从而得到背景场的误差协方差矩阵。这也正是它比最优插值优越的地方, 后者假设背景场误差的协方差矩阵是定常的。集合 Kalman 滤波的不足之处是模式误差的更新包含了一些假设, 如假设模式误差符合白噪音分布。相对最优插值而言, 其计算量更大。

由于最优插值相对于其它其他同化方法来说, 其计算量小。从计算复杂度和计算量的角度出发, 本文的同化方案采用最优插值方案。最优插值的一般形式为

$$x_a = x_b + K(y - Hx_b), \quad (1)$$

$$K = BH^T(HBH^T + R)^{-1}, \quad (2)$$

其中, x_a 表示同化后的分析场; x_b 表示背景场, 一般取为上个时刻的模式预报值; y 表示观测场; H 为线性观测算子, 表示从模式空间到观测空间的转换; B 表示背景误差协方差矩阵; R 表示观测误差协方差矩阵, 观测误差包含了观测资料的代表性误差。上“ T ”和“ -1 ”分别表示矩阵的转置和逆。

最优插值在形式上更接近于三维变分。两者的主要区别在于前者假设观测矩阵是线性的, 而后者可以处理观测矩阵是非线性的情况。直接通过(1)和(2)式来计算分析值较为困难, 因为这涉及 B 矩阵的求逆。对于高分辨率模式来说, B 矩阵是一个巨维矩阵, 直接求逆并不现实。但是(1)和(2)式可以通过最小化以下代价函数 J 而得到^[18]:

$$J = \frac{1}{2}(x - x_b)^T B^{-1}(x - x_b) + \frac{1}{2}(y - Hx)^T R^{-1}(y - Hx), \quad (3)$$

令 $X = x - x_b$, $Y = y - Hx_b$, 代入(3)式, 并利用线性观测算子假设, 可得到

$$J = \frac{1}{2}X^T B^{-1}X + \frac{1}{2}(HX - Y)^T R^{-1}(HX - Y), \quad (4)$$

其中, X 表示分析场与背景场的差, 又称分析增量; Y 表示观测场与背景场的差。同化即通过数学手段(如迭代法)寻找一个恰当的分析场使得泛函 J 最小^[19,20]。此法便于使空间上所有的观测资料都可以用于同化分析, 较于其它其他比较复杂的同化方法(如四维变分和集合 Kalman 滤波), 消耗较少的计算机资源。由于假设背景场误差定常, 可能使同化结果易包含虚假的模式误差信息。关于背景场误差的详细讨论见第 5 节。

由(4)可知, B 和 R 在分析过程中起重要作用, 目前尚无精确的公式能直接计算 B 和 R , 只能采用近似的方法估计。

对于背景误差协方差 B 矩阵的估计, 方差的计算采用 Fu 等^[21]提出的公式:

$$(\sigma_e)_j^2 \approx (\sigma_m)_j^2(1 - c_j)^2, \quad (5)$$

其中, $(\sigma_e)_j^2$ 表示第 j 个格点的误差(格点值与该点的真值之差)平方和, 是 B 矩阵的对角元素。 $(\sigma_m)_j^2$ 表示第 j 个格点的统计方差。统计方差中的误差项又称距平项, 是各个样本值(格点值)与统计时间段内平均值之差, 从而使得对背景误差协方差矩阵对角元素的估计只依赖于其统计方差 $(\sigma_m)_j^2$ 和参数 C_j 。此处简化随空间变化的 C_j 为定常的 $(0.5 \leq C_j \leq 1.5)$, $(1 - C_j)^2$ 取值为 0.25。

对于 B 矩阵非对角元素的估计, 采用传统的高斯分布形式:

$$\rho_{i,j}(\Delta x, \Delta y) = (\sigma_e)_i(\sigma_e)_j \exp\left(-\frac{\Delta x^2}{L_x^2} - \frac{\Delta y^2}{L_y^2}\right), \quad (6)$$

其中, i, j 表示格点的序号, Δx 和 Δy 分别表示第 i 个格点与第 j 个格点的水平东西距离和南北距离。此处假设预报误差(背景误差)的水平相关性随水平距离增加呈指数递减。 L_x 和 L_y 分别为预报误差东西和南北方向的水平相关尺度, 表示预报误差的相关系数减少为 $1/e$ 时的水平距离, 取 $L_x = L_y = 20$ km。

3 模式及资料处理

3.1 嵌套网格空气质量预报模式系统

本文所采用的模式为中国科学院大气物理研究所设计的基于中尺度气象预报模式(MM5)驱

动的嵌套网格空气质量预报模式系统^[22]。它主要包括两个部分：MM5 及 NAQM。其中 MM5 是由美国国家大气科学研究中心 NCAR 与宾夕法尼亚州州立大学 (the Pennsylvania State University) 共同开发的第 5 代中尺度非静力模式。NAQM 为空气污染预报系统，其为三维欧拉输送模式，主要包括污染物之排放、平流输送、扩散、气相、液相及非均相反应，干沉降以及湿沉降等物理与化学过程。光化学采用碳键 (CBM-IV) 反应机制。液相过程则采用改进的 RADM2 的液相化学反应机制。垂直坐标采用地形追随坐标。水平结构为多重嵌套网格，采用单向、双向嵌套技术，水平分辨率为 3~81 km，垂直方向不等距分为 18 层。污染物包括 SO₂、NO_x、VOC、O₃、CO、NH₃、和 PM10 等。

本文同化的区域是模式的第 4 个区域，该区域包括了上海市，其水平分辨率为 3 km。

3.2 观测资料

每日逐时观测资料由中国上海市环境监测中心提供。上海市的 22 个观测站点位置如图 1 所示，其中 20 个站点用于同化试验，其余 2 个用于检验，不参与同化。北边的检验观测点位于上海的苗苗幼儿园附近 (以下简称检验点 A)，南边的检验观测点位于上海静安附近 (以下简称检验点 B)。这些站点每天 24 h 定点定时观测，对于计算每天的平均值只需将 24 h 的观测值作算术平均即可。由于模式预报时段为当天 12 时至次日 11 时，故对观测值也作相应的处理。因为观测站点固定，无须另外再开同化窗口。

3.3 观测资料的质量控制

观测误差大致可分为 3 类。第一类，随机误差，无偏的，并满足 Gauss 分布，可用标准差来刻画；第二类，系统误差，作为有偏误差 (即期望值不为零)，主要由观测系统的误差引起；第三类，重大误差，由观测仪器的故障、数据处理或者通信中的故障引起，虽然其所占比重很小，但如不剔除，将对客观分析和数值预报的精度产生很大的影响^[23]。

质量控制 (QC) 是指在用观测资料进行客观分析或资料同化之前对资料进行检验并删除重大观测误差数据的过程。本文所采取的 QC 方法主

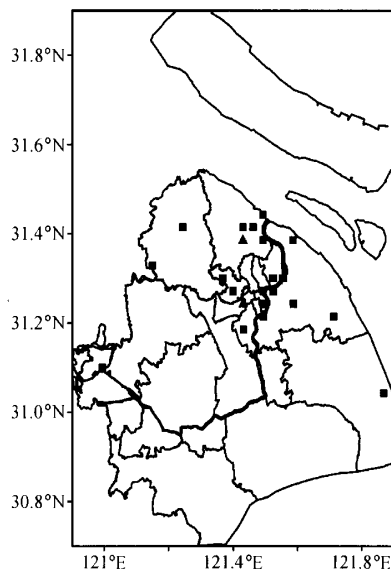


图 1 上海市环境监测站分布 (■ 站点资料用于同化, ▲ 站点资料用于检验)

Fig. 1 The distribution of monitoring stations in Shanghai (■ for assimilation, ▲ for assessment)

要是根据我国城市空气质量日报 API 指数分级标准^[24]，并基于以下原则：

(1) 合理性原则。在保证观测值正常的条件下 (例如污染物浓度不能为负值)，如果一个观测值与气候值或模式预报值相差很大，则可认为观测中有重大误差而删除。本文将 $10 \sim 600 \mu\text{g} \cdot \text{m}^{-3}$ 作为 PM10 观测值的区间， $10 \sim 2\,620 \mu\text{g} \cdot \text{m}^{-3}$ 作为 SO₂ 观测值的区间， $10 \sim 940 \mu\text{g} \cdot \text{m}^{-3}$ 作为 NO₂ 观测值的区间，不在该区间内的观测则认为该观测异常而剔除；由于 API 指数为 100 是区分非污染天气和污染天气的临界值，其对应的 SO₂、NO₂ 和 PM10 的浓度值分别为 150、120、150 $\mu\text{g} \cdot \text{m}^{-3}$ ，故将 150 $\mu\text{g} \cdot \text{m}^{-3}$ 当成观测值与模式预报值之差的上限，若两者之差大于 150 $\mu\text{g} \cdot \text{m}^{-3}$ ，则剔除观测。

(2) 空间连续性原则。如果一个观测值与其周围的观测值相差很大可认为是重大误差而删除。取 150 $\mu\text{g} \cdot \text{m}^{-3}$ 为单个观测与区域内所有观测的平均值之差的上限，超过该值即剔除。由于空气污染物日平均浓度的时间连续性较差，故不用时间连续性来判别重大误差。

4 同化试验

本研究的目的是检验最优插值同化修正大气化学模式初始场的效果,并了解该方法对不同空气质量状况下的模式初始场的修正能力。主要针对目前国内各环保重点城市空气质量预报中的 3 种污染物 PM₁₀、SO₂ 和 NO₂ 作地面层(二维)的同化试验。

试验一:对 2004 年 8 月 1~20 日 PM₁₀、SO₂ 和 NO₂ 浓度进行逐日同化。

理论上可以用分析误差矩阵来检验同化效果,

分析误差矩阵为

$$B_a = (I - KH)B, \tag{7}$$

其中, K 为 (2) 式所定义, B 为背景误差协方差矩阵。但是, (2) 式包含了对 B 矩阵的求逆,因此直接求解 (7) 非常困难。目前常用的方法是计算同化后的分析值与观测值的偏差以及未同化的模式预报值与观测值的偏差,并比较两者的大小。

图 2 和图 3 分别为此期间同化前后 3 种污染物在 2 个检验观测点上的偏差及其平均值。可见在 2 个检验点上,PM₁₀ 的同化偏差小于未同化偏差的天数均为 16 天,其余 4 天同化后偏差大于同化之前;在检验点 A、B 上 PM₁₀ 的未同化

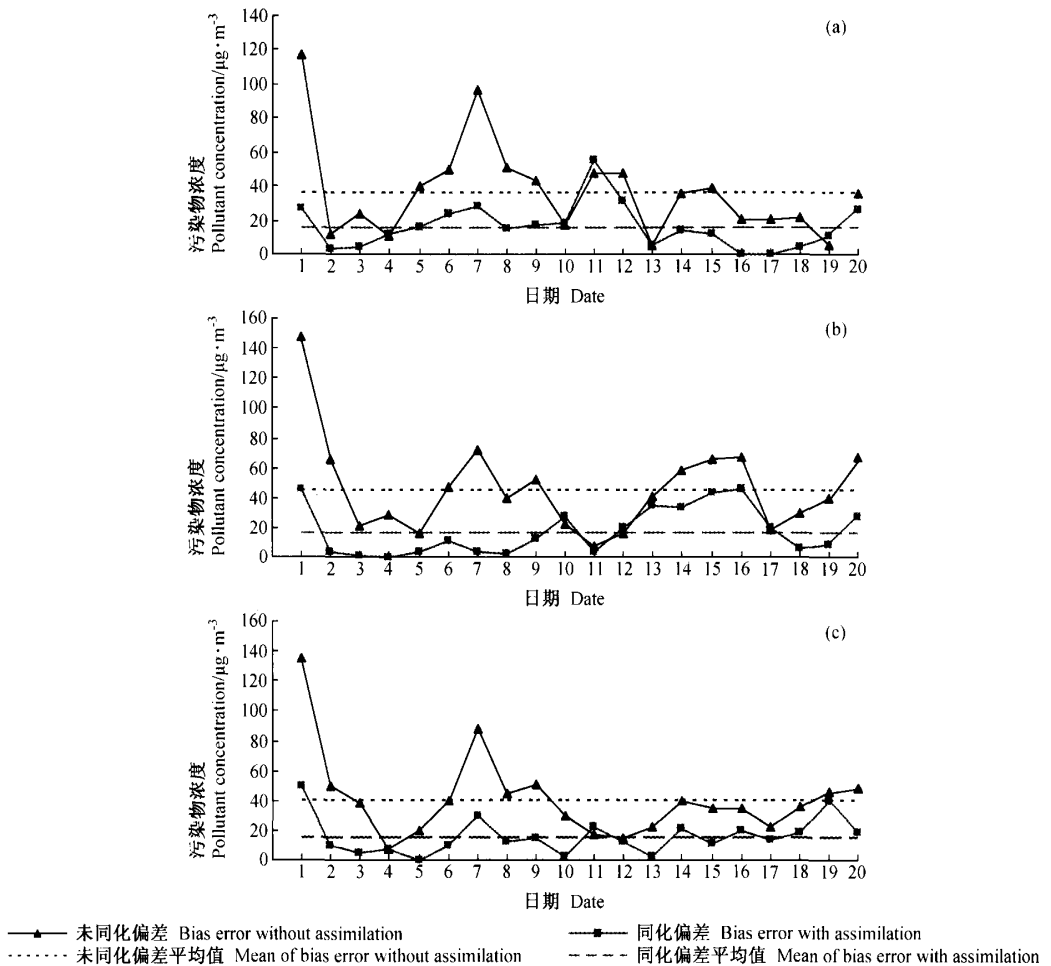


图 2 2004 年 8 月 1~20 日模式预报结果与同化后结果在检验点 A 上的偏差及其平均值: (a) PM₁₀; (b) SO₂; (c) NO₂

Fig. 2 Time series and mean of the bias error without and with assimilation of modeling results at the assessing station A during August 1-20, 2004; (a) PM₁₀; (b) SO₂; (c) NO₂

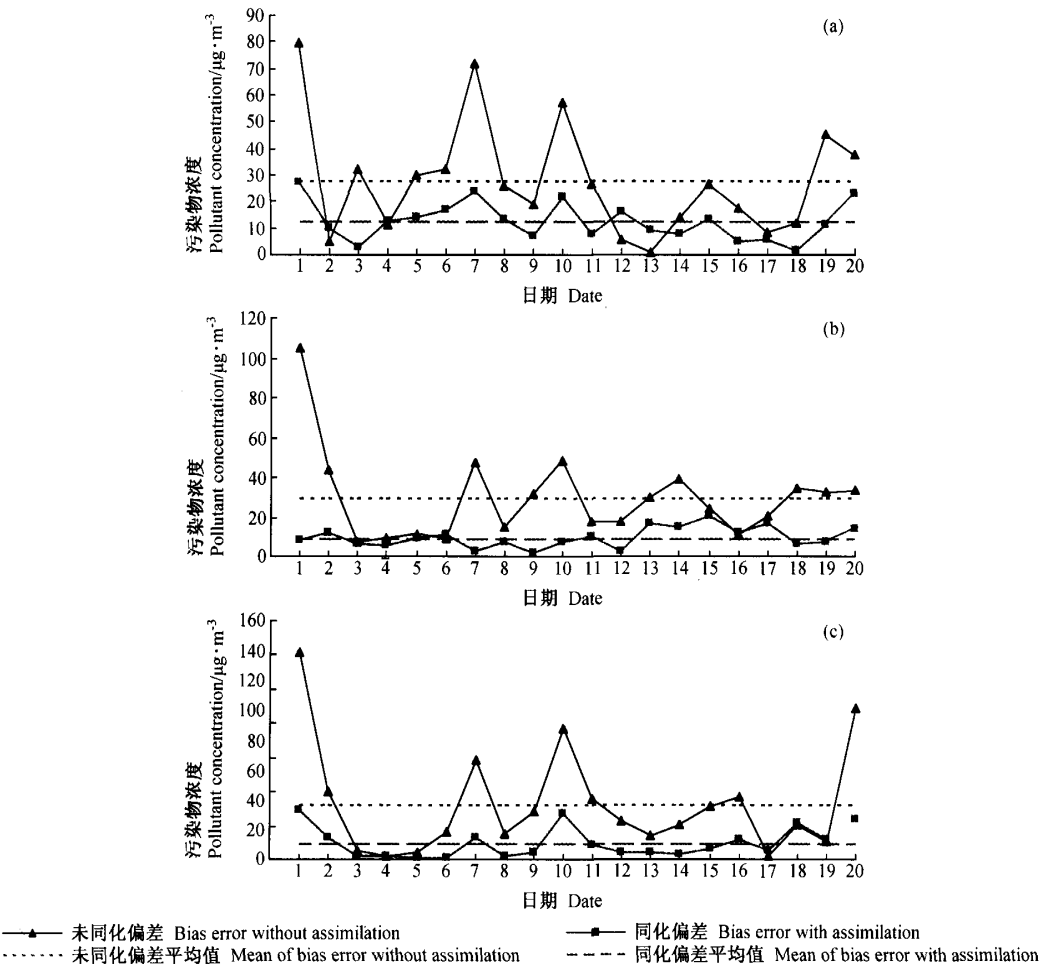


图3 同图2，但为检验点B
Fig. 3 Same as Fig. 2, but for the assessment station B

偏差平均值分别为 $37\ \mu\text{g}\cdot\text{m}^{-3}$ 和 $28\ \mu\text{g}\cdot\text{m}^{-3}$ ，同化偏差平均值分别为 $16\ \mu\text{g}\cdot\text{m}^{-3}$ 和 $12\ \mu\text{g}\cdot\text{m}^{-3}$ ，均不及相应的未同化偏差平均值的 50%。说明同化后 PM10 的模式预报结果能得到较好的改善。在 2 个检验点上 SO₂ 的同化偏差小于未同化偏差的天数均为 18 天，其余 2 天同化偏差略大于未同化偏差；在检验点 A 上，同化前 SO₂ 的偏差平均值为 $46\ \mu\text{g}\cdot\text{m}^{-3}$ ，同化后减少为 $18\ \mu\text{g}\cdot\text{m}^{-3}$ ；在检验点 B 上同化前后 SO₂ 的偏差平均值分别 $30\ \mu\text{g}\cdot\text{m}^{-3}$ 和 $10\ \mu\text{g}\cdot\text{m}^{-3}$ ，后者仅为前者的 1/3。说明对 SO₂ 的同化效果显著。在检验点 A 上，NO₂ 的同化偏差小于未同化偏差的天数为 19 天，同化前其偏差的平均值为 $41\ \mu\text{g}\cdot\text{m}^{-3}$ ，同化后减

少为 $16\ \mu\text{g}\cdot\text{m}^{-3}$ ；在检验点 B 上，NO₂ 的同化偏差小于未同化偏差的天数为 17 天，其余 3 天前者略大于后者，同化前偏差的平均值为 $33\ \mu\text{g}\cdot\text{m}^{-3}$ ，同化后减少为 $10\ \mu\text{g}\cdot\text{m}^{-3}$ ，不及同化前的 1/3。总体而言，8 月 1~20 日，无论是 PM10、NO₂ 还是 SO₂，同化偏差平均值均在 $20\ \mu\text{g}\cdot\text{m}^{-3}$ 以下，比同化前减少了至少 50%；同化后偏差比同化前小的天数在 16 天以上。从而表明，同化后的浓度值与实际观测的差异比同化前明显减少，即同化后的浓度场比同化前更接近于实际的情况。值得注意的是，若模式预报结果较好，则当天的同化偏差比未同化偏差偏大。如 8 月 2 日、12 日和 13 日的模式预报偏差较小，而同化偏差比未同化偏

差略为偏大。可见,对于较好的模式预报结果,同化的作用不显著。其主要原因是:较好的模式预报所对应的背景场误差理应较小,但此处假设背景场误差定常,而这比实际的背景场误差偏大,从而导致同化效果不显著。这也是最优插值方法本身的局限造成的。

试验二:根据环保部门的规定,首要污染物(PM₁₀、SO₂和NO₂中浓度最大的)的API指数以100为界来区分非污染形势和污染形势。API指数在100以下表示空气质量状况良好,100以上则表示空气受到污染(包括轻度污染、中等污染和重污染)。因为首要污染物常常是PM₁₀,所以这里以PM₁₀为同化对象,以 $150\text{ }\mu\text{g}\cdot\text{m}^{-3}$ 为界,分别取小于 $150\text{ }\mu\text{g}\cdot\text{m}^{-3}$ (清洁状况)和大于 $150\text{ }\mu\text{g}\cdot\text{m}^{-3}$ (污染状况)两种形势各10天作同化试验,并比较其效果。由于连续污染10天的情况基本没有,此处10天同化污染形势试验并非连续的,对于清洁形势同化试验则取8月22日到8月~27日和9月25日到9月~28日共10天。

参数选择:同化方案采用共轭梯度法求解,作为一种迭代算法,其优点是能避免对**B**矩阵的直接求逆,在本文中迭代次数取为8;水平相关尺度(*L*)取20 km;经过对观测资料进行初步的质量控制后,同化站点的个数为12~19个,独立检验的观测站点个数为2个;观测误差协方差矩阵假定为对角阵,根据观测仪器的性能,观测误差假定为观测值的10%。

通过计算两种不同空气质量状况下的同化结果与未同化模式预报结果在2个独立检验观测点上的均方根(RMS)误差来比较它们的大小。结果发现,无论污染状况还是清洁状况,同化后的RMS均小于同化前的RMS(图4)。对于污染状况,同化前的RMS为 $40\sim 86\text{ }\mu\text{g}\cdot\text{m}^{-3}$,而同化后的RMS减少为 $10\sim 45\text{ }\mu\text{g}\cdot\text{m}^{-3}$ (图4a)。对于清洁状况,同化前的RMS为 $4\sim 56\text{ }\mu\text{g}\cdot\text{m}^{-3}$,而同化后的RMS减少为 $2\sim 23\text{ }\mu\text{g}\cdot\text{m}^{-3}$ (图4b)。此外,无论是模式预报结果还是同化结果,在污染状况下的RMS均比在清洁状况下的RMS要

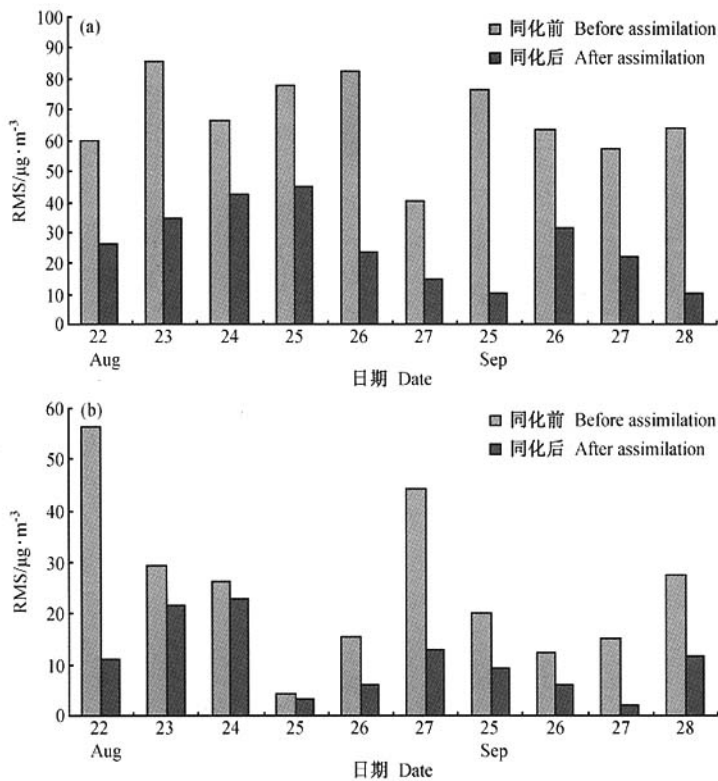


图4 污染状况和清洁状况同化前后的RMS误差比较:(a)污染状况,(b)清洁状况

Fig. 4 The comparison of RMS error between without and with assimilation for (a) polluted days (a) and (b) clean condition

大。这是因为污染状况的浓度场量值较大，从而导致误差容易增大，而清洁状况的浓度场量值较小，误差不易增大的缘故。

图 5 是清洁状况和污染状况下未同化的 PM10 模式预报结果与同化结果的对比。可以看出，清洁状况下（8 月 22 日）在同化区域的中部、西南

部和东南部的观测的浓度值较低，为 $20 \sim 85 \mu\text{g} \cdot \text{m}^{-3}$ ，模式预报结果在中部是高值区（图 5a）；经过同化后，在观测较为集中的中部地区变为低值区，西南部的低值观测（ $20 \mu\text{g} \cdot \text{m}^{-3}$ ）也导致该地区出现一个低值区（图 5b）。这说明对于空气清洁状况，低的观测值能有效地降低模式

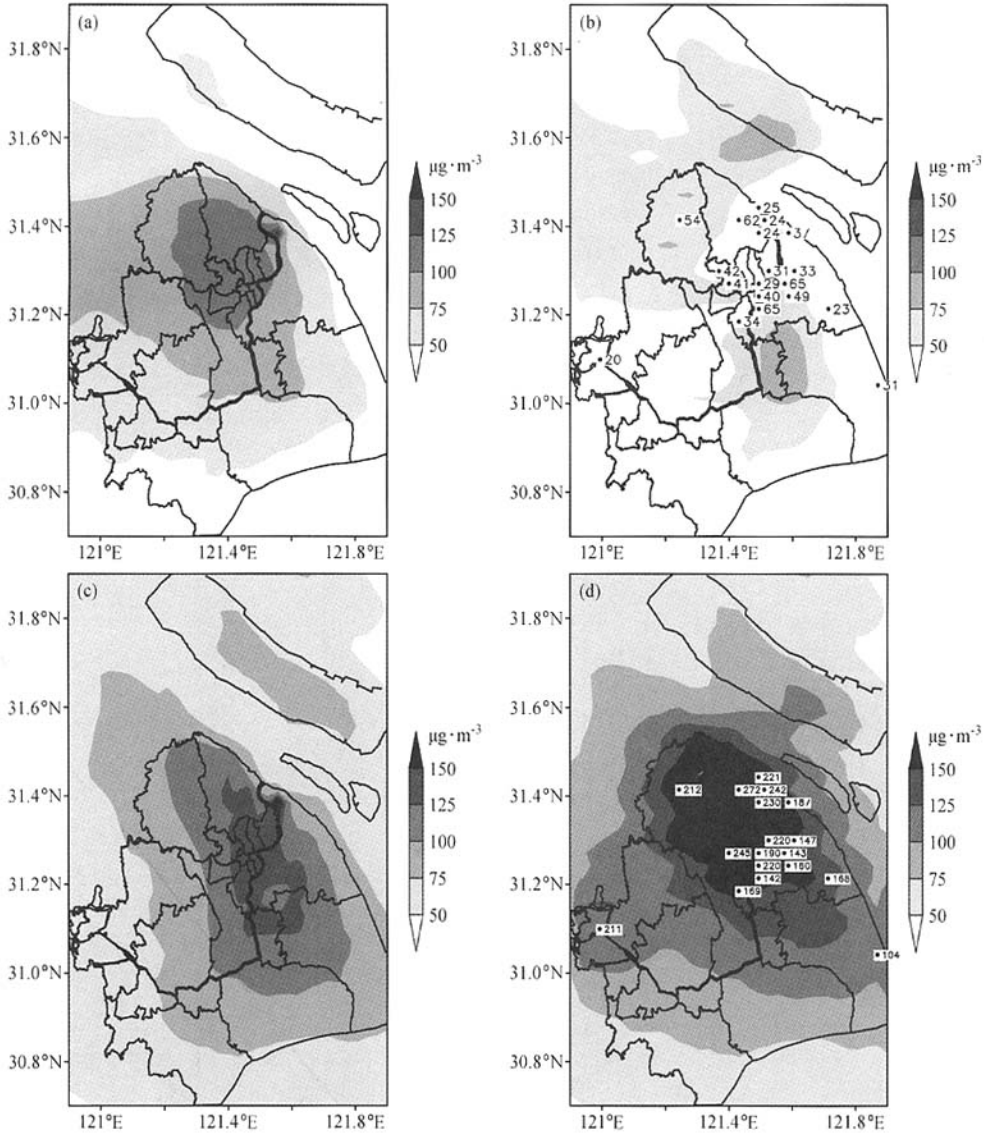


图 5 清洁状况和污染状况未同化的模式预报 PM10 浓度场与同化后的 PM10 浓度场：(a) 8 月 22 日清洁状况未同化浓度场；(b) 8 月 22 日清洁状况同化后浓度场；(c) 12 月 17 日污染状况未同化浓度场；(d) 12 月 17 日污染状况同化后浓度场
Fig. 5 The concentration fields without and with assimilation for clean condition and polluted days: (a) without assimilation for clean air condition on Aug 22; (b) with assimilation for clean air condition on Aug 22; (c) without assimilation for polluted air condition on Dec 17; (d) with assimilation for polluted air condition on Dec 17

预报场的强度。对于空气污染状况(12月17日)未同化模式预报结果在中部地区是高值区(图5c),但观测显示该地区浓度场的实际强度更大,在有观测的地区及其邻近区域,同化场的强度均比未同化场要大(图5d)。总之,在有观测的地区,同化结果比未经同化的模式预报结果更接近于实际观测值。

5 背景误差协方差的讨论

大气化学模式误差不仅依赖于模式本身,还依赖于气象场的强迫误差,而后者很难描述。本文假设背景误差协方差是模式模拟(预报)协方差的一部分,其中一个好处就是避免了对模式误差的直接估计。(6)式中可调参数 L_x 和 L_y 表示模式场误差的水平相关尺度,而大气化学污染物浓度受气象场的影响,故模式场的水平相关性与风向和气压等气象因子有关,应该是空间可变的。为简化计算量,本文假定水平相关尺度定常,且

$L_x=L_y$,最优插值法的这个缺点,从计算效率的角度看,却是比其它其他同化方法优越的地方。

为了估计 L_x 和 L_y 对同化结果的影响,分别取水平相关尺度为20、40、60、80和100 km,对清洁状况(8月22日)和污染状况(12月17日)的PM₁₀模式预报场做同化试验。可以看出,8月22日模式场的RMS误差为 $56\ \mu\text{g}\cdot\text{m}^{-3}$,相关尺度为20 km和40 km时分析场的RMS误差相差不大,均为 $11\ \mu\text{g}\cdot\text{m}^{-3}$,相关尺度为60、80和100 km时RMS误差略为偏大,分别为13、14和 $15\ \mu\text{g}\cdot\text{m}^{-3}$,均较未同化的模式预报误差小(图6a)。污染状况(12月17日)模式场的RMS误差为 $57\ \mu\text{g}\cdot\text{m}^{-3}$,不同相关尺度的分析场RMS误差均小于模式场的RMS误差,但它们彼此差别不大,相关尺度取20~80 km的分析场RMS误差介于 $22\sim 23\ \mu\text{g}\cdot\text{m}^{-3}$ (图6b)。其原因可能是因为用来检验的观测点的位置与用来同化的观测点位置相距不远,无论取哪个相关尺度,检验站点均在同化影响范围之内。

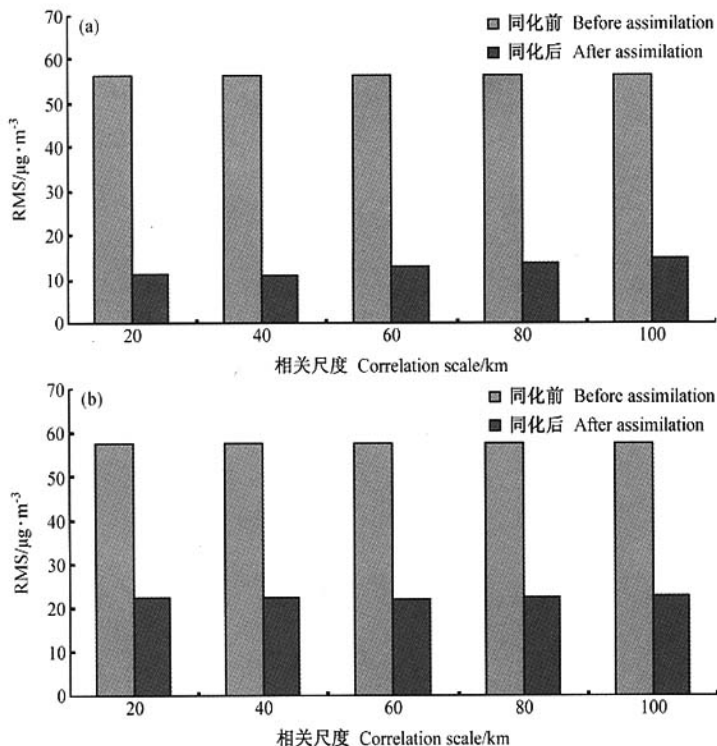


图6 取不同相关尺度时同化前后 RMS 误差比较: (a) 8月22日; (b) 12月17日

Fig. 6 The comparison of RMS error without and with assimilation for different correlation scale: (a) Aug 22; (b) Dec 17

不过,相关尺度的增大,也意味着观测对同化区域的影响范围也增大。由于同化区域大概为 $100\text{ km} \times 130\text{ km}$,若相关尺度增加到 100 km ,意味着同化区域内的任何一个观测都能影响全区域,有可能导致整个同化区域的浓度场趋于一致,而这与事实不符(图 7)。因此,在观测分布不均匀的情况下,为了简化对相关尺度的估计,这里认为取 20 km 是最佳的,因为这样既使得同化偏差较小,又确保同化的影响范围在观测的附近。

6 结论

(1) 观测资料的质量对资料同化非常重要。在进行资料同化之前,应先进行观测资料的质量控制。通过合理性原则和连续性原则将由观测仪器的故障、数据处理或者通信中的故障引起的观测误差剔除,能够在一定程度上保证观测资料的可靠性。

(2) 利用最优插值方法对 PM_{10} 、 SO_2 和 NO_2 进行连续 20 天的同化试验结果表明,无论是 PM_{10} 、 SO_2 还是 NO_2 ,同化偏差的平均值均小于 $20\text{ }\mu\text{g} \cdot \text{m}^{-3}$,比未同化偏差平均值减少了至少 50%。同化偏差小于未同化偏差天数都在 16 天以

上,说明该方法能较好地修正模式预报场。另外,其余的 1 至 4 天模式预报结果较好,同化偏差略大于未同化偏差。这表明,对于较好的模式预报结果,同化的作用不显著。

(3) 取不同空气质量状况下的 PM_{10} 浓度作 10 天的同化试验。结果表明,无论是污染状况还是清洁状况,同化后的 RMS 误差均小于同化之前,说明同化后的浓度场均比同化前更接近实际浓度场。值得注意的是,无论是未同化模式预报结果还是同化结果,在污染状况下的 RMS 均大于在清洁状况下的 RMS。这可能是因为污染状况的浓度场量值较大,从而导致误差容易增大,而清洁状况的浓度场量值较小,误差不易增大的缘故。

(4) 本文假设背景误差协方差是模式模拟(预报)协方差的一部分,可避免对背景误差的直接估计。模式水平相关尺度假设为定常,从而简化计算。理论上对相关尺度选取时应考虑空间和时间的变化,Zhou 等^[25]提出了一种根据模式结果估计随空间变化的相关尺度的方法,并且在海温资料同化中有效改进了同化结果。另外一个解决此问题的办法是采用集合 Kalman 滤波的方法来估计背景协方差矩阵,这是我们今后的工作中改进的方向。

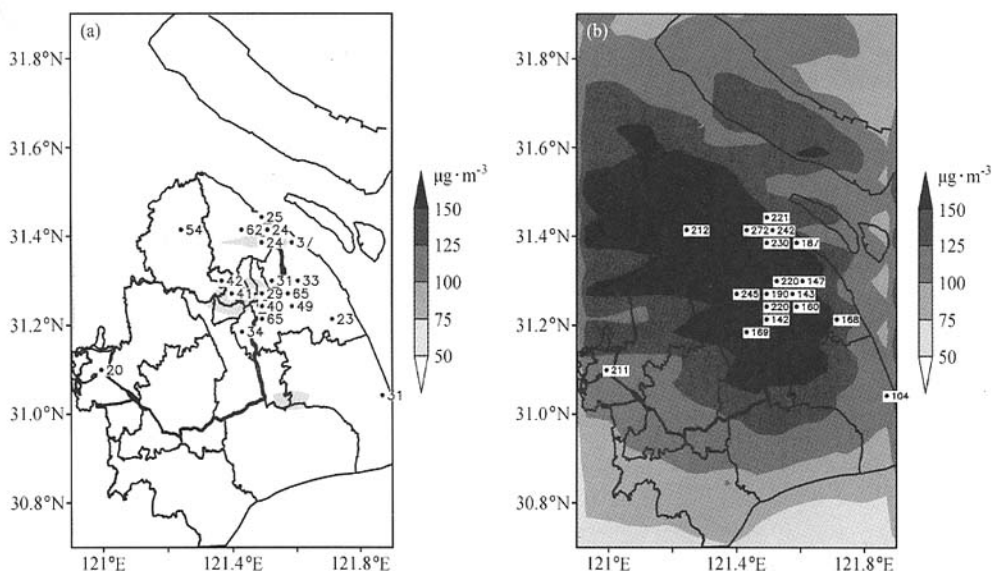


图 7 水平相关尺度取 100 km 的 PM_{10} 同化场: (a) 8 月 22 日; (b) 12 月 17 日

Fig. 7 The assimilated fields of PM_{10} with correlation scale equal to 100 km : (a) Aug 22; (b) Dec 17

参考文献 (References)

- [1] Charney J, Halem M, Jastrow R. Use of incomplete historical data to infer the present state of the atmosphere. *J. Atmos. Sci.*, 1969, **26**: 1160~1163
- [2] Austin J. Toward the four-dimensional assimilation of stratospheric chemical constituents. *J. Geophys. Res.*, 1992, **97**: 2569~2588
- [3] Khattatov B, Lyjak L, Gille J. On Applications of Photochemical Models to the Design of Measurement Strategies. *Geophys. Res. Lett.*, 2001, **28** (12): 2377~2380
- [4] Elbern H, Schmidt H, Ebel A. Variational data assimilation for tropospheric chemistry modeling. *J. Geophys. Res.*, 1997, **102** (D13): 15967~15985
- [5] Eskes H J, PETERS A J M, Levelt P F, et al. Variational assimilation of total ozone satellite data in a 2D lat-lon tracer-transport model. *J. Atmos. Sci.*, 1999, **56**: 3560~3572
- [6] Clerbaux C, Hadji-Lazaro J, Hauglustaine D, et al. Assimilation of carbon monoxide measured from satellite in a three-dimensional chemistry-transport model. *J. Geophys. Res.*, 2001, 106 (D14): 15385~15394
- [7] Fisher M, Lary D J. Lagrangian four dimensional variational data assimilation of chemical species. *Quart. J. Roy. Meteorol. Soc.*, 1995, **121**: 1681~1704
- [8] Levelt P F, Allaart M A F, Kelder H M. On the assimilation of total-ozone satellite data. *Ann. Geophys.*, 1996, **14**: 1111~1118
- [9] Lyster P M, Cohn S E, Menard R, Chang L P, et al. An implementation of a two-dimensional filter for atmospheric chemical constituent assimilation on massively parallel computers. *Mon. Wea. Rev.*, 1997, **125**: 1674~1686
- [10] Umeda T, Martien P T. Evaluation of a Data Assimilation Technique for a Mesoscale Meteorological Model Used for Air Quality Modeling. *J. Appl. Meteor.*, 1999, **41**: 12~29
- [11] Alberto M D, Armistead G R. Estimation of emission adjustments from the application of four-dimensional data assimilation to photochemical air quality modeling. *Atmos. Environ.*, 2001, **35**: 2879~2894
- [12] Remus G H, Guus J M V, Arnold H. Data assimilation of ground-level ozone in Europe with a Kalman filter and chemistry transport model. *J. Geophys. Res.*, 2004, **109**, D10302, doi: 10.1029/2003JD004283
- [13] 朱江, 徐启春, 王赐震, 等. 海温数值预报资料同化试验. *海洋学报*, 1995, **17** (6): 9~20
Zhu Jiang, Xu Qichun, Wang Cizhen, et al. The assimilation experiment of the sea temperature numerical forecast. *Acta Meteorologica Sinica* (in Chinese), 1995, **17** (6): 9~20
- [14] 朱江, 王辉, 周广庆. 海表温度的自适应变分同化. *科学通报*, 2002, **47** (19): 1517~1521
Zhu Jiang, Wang Hui, Zhou Guangqing. The self-adaptive variational assimilation for sea surface temperature. *Chinese Science Bulletin* (in Chinese), 2002, **47** (19): 1517~1521
- [15] Eliassen A. Provisional report on calculation of spatial covariance and autocorrelation of the pressure field. Report no. 5, Videnskaps-Akademiets Institutt for Vær-Og Klimaforskning, Oslo, Norway, 1954. In: *Dynamic Meteorology: Data Assimilation Methods*. Ghil M, Kallen E, Ed. Applied Mathematical Sciences. Springer Verlag, 1981, V36, 319~330
- [16] Lewis, J M, and Derber J C. The use of adjoint equations to solve a variational adjustment problem with advective constraints. *Tellus*, 1985, **37A**: 309~322
- [17] Evensen G. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, 1994, **99** (C5): 10143~10162
- [18] Lorenc A C. Analysis methods for numerical weather prediction. *Quart. J. Roy. Meteor. Soc.*, 1986, **112**: 1177~1194
- [19] Parrish D, Derber D. The National Meteorological Center's Spectral Statistical-Interpolation Analysis System. *Mon. Wea. Rev.*, 1992, **120**: 1747~1763
- [20] Deber J, Rosati A. A global oceanic Data Assimilation System. *J. Phys. Oceanogr.*, 1989, **19**: 1333~1347
- [21] Fu Weiwei, Zhou Guangqing, Wang Huijun. Ocean Data Assimilation with Background Error Covariance Derived from OGCM Outputs. *Adv. Atmos. Sci.*, 2004, **21** (2): 181~192
- [22] Wang Zifa, Maeda T, Hayashi M, et al. A nested air quality prediction modeling system for urban and regional scales; application for high-ozone episode in Taiwan. *Water, Air and Soil Pollution*, 2001, **130**: 391~396
- [23] 朱江. 观测资料的四维质量控制: 变分法. *气象学报*, 1995, **53** (4): 480~487
Zhu Jiang. The four-dimensional quality control for observations; variational method. *Acta Meteorologica Sinica* (in Chinese), 1995, **53** (4): 480~487
- [24] <http://www.sepa.gov.cn/quality/background.php>
- [25] Zhou Guangqing, Fu Weiwei, Zhu Jiang, et al. The impact of location-dependent correlation scales in ocean data assimilation. *Geophys. Res. Lett.*, 2004, **31**, L21306, doi: 10.1029/2004GL020579