

基于本征正交分解的显式四维变分同化方法: 理论与验证

田向军*, 谢正辉

中国科学院大气物理研究所, 北京 100029

* E-mail: tianxj@mail.iap.ac.cn

收稿日期: 2008-03-05; 接受日期: 2008-07-15

国家自然科学基金(批准号: 40705035)、国家高技术研究发展计划(编号: 2009AA12Z129, 2007AA12Z144)、国家重点基础研究发展计划项目(编号: 2009CB421407)和财政部/科技部公益类行业专项(编号: GYHY200706005)资助

摘要 将本征正交分解(POD)方法用于四维空间预报集合提取标准正交基, 该标准正交基在最小二乘意义下是最优的, 与奇异值分解(SVD)技术相比, 它能捕捉到预报集合空间更多的能量, 能够更好地表征四维变量的空间结构以及时间演变特征. 将分析向量依截断的 POD 基展开后, 控制变量会显式地出现在代价函数中, 避免了传统的四维变分方法所必需的伴随模式的运用, 使得同化过程简单. 用土壤湿度预报方程和人造资料进行一系列的数值试验对该基于本征正交分解的显式变分方法与基于 SVD 基的方法以及集合 Kalman 滤波进行比较, 结果表明: POD/SVD 方法从同化精度和同化时效上都要远远优于一般的集合 Kalman 滤波方法; 由于 POD 基在最小二乘意义下的最优性, 基于 POD 分解的同化方法要优于基于 SVD 分解的方法, 尤其在模式存在误差的情况下表现得更为明显.

关键词

资料同化
四维变分
显式方法
本征正交分解

四维变分同化方法^[1,2](four-dimensional variational data assimilation, 简称为 4DVAR)本质上是将资料同化问题转化为一个最优控制问题, 根据同化窗内所有可能的观测优化大气(海洋、陆面)模式的初始场, 使得据此得到的预报与观测之间的距离最小, 其优势明显: (1) 利用完整的模式方程作为动力约束; (2) 可以同时同化多个时刻的观测资料, 但该方法也存在弊端: 4DVAR 给出的代价函数中, 控制变量(初始场)是以隐函数的形式出现的, 为了求得代价函数相对初始场的梯度, 需要给出预报模式的切线性模式及相应的伴随模式; 极小化代价函数也需要反复积分模式方程和伴随方程, 计算效率极低. 与大气和海洋模式相比较, 陆面模式中高度的非线性和间断性质表现得更为突出^[3], 或许正是由于这些原因使得

四维变分同化方法在陆面数据同化中应用极少. 很多工作在改进 4DVAR 方法上取得进展, 这些方法仍然需要利用切线性模式或伴随模式^[4~6], Qiu 等^[7,8]从大气吸引子的角度将奇异值分解(singular value decomposition, 简称 SVD)用于四维变分同化方法, 大大简化了同化过程并取得了不错的结果. 本文将本征正交分解(proper orthogonal decomposition, 简称 POD)^[9~12]技术和集合 Kalman 滤波(ensemble kalman filter, 简称为 EnKF)^[13]思想相结合发展了一种新的基于本征正交分解的显式四维变分方法: 类似于 EnKF 那样利用蒙特卡罗方法产生一个四维的样本集合, 在将 POD 方法应用于四维空间的预报集合提取正交基, 由于 POD 正交基在最小二乘意义下的最优性, 与奇异值分解(SVD)技术相比, 它能捕捉到预报集合

更多的能量, 能更好地表征四维变量的空间结构以及时间演变特征. 该方法无需积分伴随模式, 实现起来简单便捷, 计算量与Qiu等^[7,8]的方法相当. 本文利用土壤湿度预报方程以及人造的观测数据进行一系列的数值试验对该基于 POD 基的显式变分方法与基于SVD基的方法^[7,8]以及集合Kalman滤波进行了比较.

1 本征正交分解(POD)

设 $\bar{U}_i(\bar{x})$, $i=1, 2, \dots, N$ 表示物理过程在 $\bar{x}$ 点的 N 个样本瞬像的集合, 瞬像集合的平均值给出如下:

$$\bar{U} = \frac{1}{N} \sum_{i=1}^N \bar{U}_i(\bar{x}). \quad (1.1)$$

通过集合(1.1)的平均偏差构造如下新集合:

$$V_i = \bar{U}_i - \bar{U}, \quad (1.2)$$

希望能够找到数据序列(1.2)式的一个最优化的简练描述. 该方法的一种描述是用一组基函数的级数展开式. 直观地, 这组基函数应该是(1.2)式集合元素在某种意义下的代表. 这样的一种坐标系统由 Karhunen-Loève 的展开式^[11]提供, 具有一定的最优化性质. 事实上, 这样的基函数 Φ 是所有这些瞬像的代表数和

$$\Phi = \sum_{i=1}^N a_i V_i(\bar{x}), \quad (1.3)$$

其中 a_i 为待定的系数, 由此给出的 Φ 能最大程度的代表集合 $\{V_i(\bar{x})\}_{i=1}^N$: 函数 Φ 使得

$$\frac{1}{N} \sum_{i=1}^N a_i |(V_i, \Phi)|^2, \quad (1.4)$$

最大且满足

$$(\Phi, \Phi) = \|\Phi\|^2 = 1, \quad (1.5)$$

其中 $(\cdot, \cdot)$ 和 $\|\cdot\|$ 分别表示区域 Ω 上的 L^2 内积和 L^2 范数.

由此可知^[11], 基函数是下面积分方程的特征函数

$$\int C(\bar{x}, \bar{x}') d\bar{x}' = \lambda \Phi, \quad (1.6)$$

其中

$$C(\bar{x}, \bar{x}') = \frac{1}{N} \sum_{i=1}^N a_i V_i(\bar{x}) V_i(\bar{x}'), \quad (1.7)$$

将(1.3)式代入(1.7)式得特征值问题

$$\sum_{j=1}^N L_{ij} a_j = \lambda a_i, \quad i=1, 2, \dots, N, \quad (1.8)$$

其中 $L (L_{ij} = (V_i, V_j)/N)$ 为对称非负矩阵. 则问题转化为求解 $N \times N$ 矩阵的特征向量, 其中 N 是瞬像集合的大小. 直接计算^[11]表明: 当(1.3)式的系数 a_i 是对应于 L 的最大特征值的特征向量的分量时, 代价函数

$$\frac{1}{N} \sum_{i=1}^N a_i |(V_i, \Phi)|^2 = (\lambda \Phi, \Phi) = \lambda \quad (1.9)$$

取得最大值.

另外, 我们也可以考虑瞬像集合的一个最优化表示的离散 Karhunen-Loève 的展开式^[11]. 在二维情况下, 每个瞬像 U_i (定义在 $n \times n$ 个节点上组成的集合上) 的样本都能表示为 n^2 维向量 $\bar{u}_i$:

$$\bar{u}_i = (\bar{u}_{i1}, \bar{u}_{i2}, \dots, \bar{u}_{in^2})^T, \quad (1.10)$$

其中 $\bar{u}_{ij} = U_i(x_j, y_j)$ 表示向量 $\bar{u}_i$ 的第 j 个分量. 这里, 总体 $\bar{u}$ 的离散协方差矩阵定义为

$$C_{\bar{u}} = E \{ (\bar{u} - \bar{m}_{\bar{u}})(\bar{u} - \bar{m}_{\bar{u}})^T \}, \quad (1.11)$$

其中

$$\bar{m}_{\bar{u}} = E \{ \bar{u} \} \quad (1.12)$$

是均值向量, E 是期望值. 方程(1.11)和(1.12)可由下面的两式代替:

$$C_{\bar{u}} = \frac{1}{N} \left[\sum_{i=1}^N \bar{u}_i \bar{u}_i^T \right] - \bar{m}_{\bar{u}} \bar{m}_{\bar{u}}^T, \quad (1.13)$$

和

$$\bar{m}_{\bar{u}} = \frac{1}{N} \sum_{i=1}^N \bar{u}_i, \quad (1.14)$$

协方差矩阵的大小具有的维数是 $n^2 \times n^2$, 其中 n^2 是结点的数目(或样本的像素). 对应于 $C_{\bar{u}}$ 的最大特征值的特征向量是主要的 POD 模态. 很显然, 以上同样适用于一维情形.

2 基于本征正交分解的显式四维变分方法

4DVAR 是在一个同化时间窗内极小化下面的代价函数:

$$J(\bar{x}) = (\bar{x} - \bar{x}_b)^T B^{-1} (\bar{x} - \bar{x}_b) + \sum_{k=0}^K [\bar{y}_k - H_k(\bar{x}_k)]^T R_k^{-1} [\bar{y}_k - H_k(\bar{x}_k)], \quad (2.1)$$

其中, 向量 $\bar{x}$ 和 $\bar{x}_b$ 分别是模式变量和背景场变量, B 是背景误差协方差矩阵, $\bar{y}_k$ 是 k 时刻的观测, $\bar{x}_k$ 是相应时刻的预报, 它是从初始场 $\bar{x}_0$ 出发积分下面的预报方程得到的:

$$\dot{\bar{x}}_k = M_k(\bar{x}), \quad (2.2)$$

H_k 和 R_k 分别是观测算子和观测误差协方差矩阵. 在上面的代价函数(2.1)中, 控制变量 $\bar{x}$ 是以隐式形式出现的, 需要通过预报方程将它与 $\bar{x}_k$ 联系, 这便造成了计算 J 对 $\bar{x}$ 的梯度的困难, 需要给出预报模式的切线性模式及相应的伴随模式. 本节中我们从 POD 技术与 EnEK 思想相结合的角度出发提出一种新的显式四维变分同化方法, 为叙述方便下面将该方法简称为 POD-E4DVAR. Qiu 等^[7,8]从大气吸引子的角度提出了一种基于 SVD 分解的显式四维变分同化方法(简称为 SVD-E4DVAR), 取得了较好的同化效果. 由于 POD 正交基在最小二乘意义下是最优的^[12], 与奇异值分解(SVD)技术相比, 它能捕捉到预报集合空间更多的能量, 能更好地表征四维变量的空间结构以及时间演变特征, 有望取得更佳的同化效果. 该方法首先选定一个同化时间窗, 然后类似于 EnKF 那样利用蒙特卡罗方法产生一组出预报集合, 在所给定的同化窗内从这组预报集合中的多个时间层上取样, 得到一个四维的样本集合, 再对这一样本集合构成的矩阵作 POD 分解, 将相对于最大特征值的那些特征向量作为基向量, 然后将模式状态变量表示为按照截断的基向量展开的形式, 根据观测最优地确定展开系数, 这时展开系数在代价函数中是显式出现的. 叙述具体做法如下.

设同化时间窗内 $(0, T)$ 内有 S 个时间步, 其中在 K 个时刻 $(t=t_1, \dots, t_k, \dots, t_K)$ 的观测 $\bar{y}(t_k)$, $1 \leq k \leq K$. 对 $\bar{x}_0$ 作扰动后积分模式并在上述 K 或者 S (即有两种扰动方式)个时刻产生 N 个样本的集合 X_n , $1 \leq n \leq N$:

$$\bar{X} = \frac{1}{N} \sum_{n=1}^N X_n, \quad (2.3)$$

以向量 $\delta X_n = X_n - \bar{X}$ ($n=1, 2, \dots, N$) 作为矩阵的列向量组成的一个维数为 $M \times N$ 的矩阵(记为 A), 这里 $M = M_g \times M_v \times K(S)$, M_g , M_v 分别是空间格点数和模式变量数. 下一步求矩阵 $T = (AA^T)_{M \times M}$ 的特征值和特征向量. 因为一般有 $M \gg N$, 因此求矩阵 T 的特征值和特征向量代价将很大. 在这里, 使用时空转化技巧, 转而求矩阵 $T = (A^T A)_{N \times N}$ 的特征值和特征向量. 求下列的特征值问题:

$$TV = \lambda V, \quad (2.4)$$

易证矩阵 T 是对称半正定的, 因此得到一组降序排列的特征值 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N \geq 0$ 以及相应的标准正交基 $V^j = (v_1^j, v_2^j, \dots, v_N^j)$ ($j=1, 2, \dots, N$) 组成的矩阵 $V_{N \times N}$. 由于时空转换, 记矩阵 $F = (A \times V)_{M \times N}$, 此时的矩阵 F 并非标准正交, 但如果 F 的每一列均除以 $\sqrt{\lambda_j}$ ($j=1, 2, \dots, N$), 则矩阵 F 就成为标准正交阵. 由此一组基向量定义为

$$\Phi_{ik} = F_k^i, \quad i=1, 2, \dots, N; k=1, \dots, M, \quad (2.5)$$

其中 F_k^i 为矩阵 F 第 k 行第 i 列的元素, 将四维空间的分析变量 X_a 以由(2.5)式所确定的 POD 基展开可得:

$$X_a = \bar{X} + \sum_{j=1}^p \alpha_j F_j, \quad (2.6)$$

其中 F_j 为矩阵 F 的第 j 个列向量. 且 P 定义为

$$I(p) = \sum_{i=1}^p \lambda_i / \sum_{i=1}^N \lambda_i,$$

选择 p , 使得

$$p = \arg \min \{I(p) : I(p) \geq \gamma\}, \quad (2.7)$$

其中 $0 < \gamma < 1$ 是由降维空间 $D^p = \text{span}\{\Phi_1, \dots, \Phi_p\}$ 所捕获的总能量的百分数. 为了捕获 POD 基的大部分能量, 允许 γ 在 1 附近取值.

(2.6)式相当于在一个四维空间给出的预报模式的解的近似表达式, 待定的只有展开系数 $\alpha = (\alpha_1, \dots, \alpha_j, \dots, \alpha_p)^T$. 将(2.6)式代入 4DVAR 的代价函数(2.1)中, J 相对于控制变量 a 的梯度很容易计算.

3 POD-E4DVAR 在土壤湿度同化中的验证

利用一个简单的土壤湿度预报模型和人造资料检验该方法(POD-E4DVAR, 其中令仅在 K 个观测时刻进行取样的方式定义为 POD1-E4DVAR, 在整个同化窗口中的每个时刻进行取样的方式为 POD2-E4DVAR)在陆面数据同化中的应用并和 SVD-E4DVAR 以及 EnKF 方法作比较.

3.1 试验设计

采用目前通用的陆面过程模式CLM^[14]中的土壤湿度预报方程:

$$\frac{\partial \theta}{\partial t} = -\frac{\partial q}{\partial z} - E - R_{fm}, \quad (3.1)$$

其中 θ 是体积含水率 (volumetric soil moisture, $\text{m}^3 \cdot \text{m}^{-3}$), q 是水分通量($\text{m} \cdot \text{s}^{-1}$), E 是蒸发, R_{fm} 是源汇项, 为简单起见, E, R_{fm} 都取为 0. 水分通量由 Darcy定律^[15]确定

$$q = -k \frac{\partial(\varphi + z)}{\partial z}, \quad (3.2)$$

其中 k 是水力传导率, φ 是土壤水势, z 竖直向下为正, 坐标原点取在地面. 且有

$$k = k_s \left(\frac{\theta}{\theta_s} \right)^{2b+3}, \quad \varphi = \varphi_s \left(\frac{\theta}{\theta_s} \right)^{-b}, \quad (3.3)$$

k_s 是最大水力传导率, θ_s 是饱和土壤湿度, φ_s 是最大土壤水势, b 是Clapp & Hornberger 常数^[16](表 1). 模型的分层结构与CLM一致, 分为厚度不均的 10 层, 且上边界条件取入渗(infiltration)通量 $q_0(t)$, 下边界条件则采用零通量条件, 时间步长取为 0.5 h.

表 1 土壤水模型中所用的参数

θ_s	k_s	b	φ_s
0.46	$2.07263 \times 10^{-6} \text{ m} \cdot \text{s}^{-1}$	8.634	-3.6779 m

试验中采用两套不同的入渗序列(图 1)表示预报模式有无误差: 预报模型采用“Perfect”(正确)入渗时间序列表示该预报模型无误差; 而预报模型采用“Imperfect”(误差)入渗时间序列表示该预报模型有误差. 另外试验中采用两个不同的初始土壤湿度表示初始场对于模式预报的影响(图 2). 首先利用

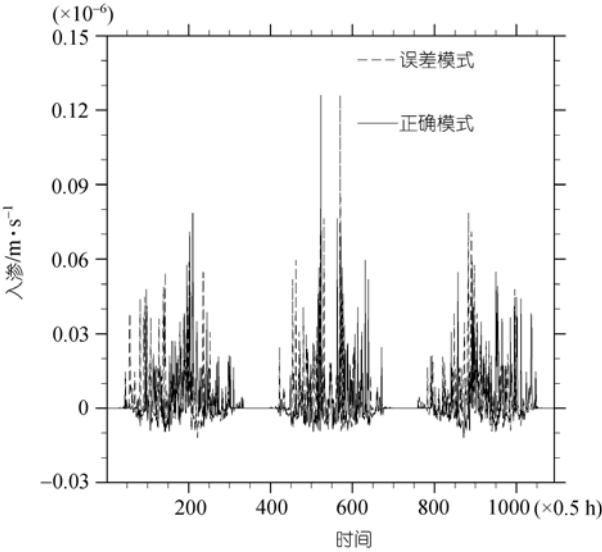


图 1 用于土壤湿度模式预报的入渗(Infiltration, $q_0(t)$)的时间序列

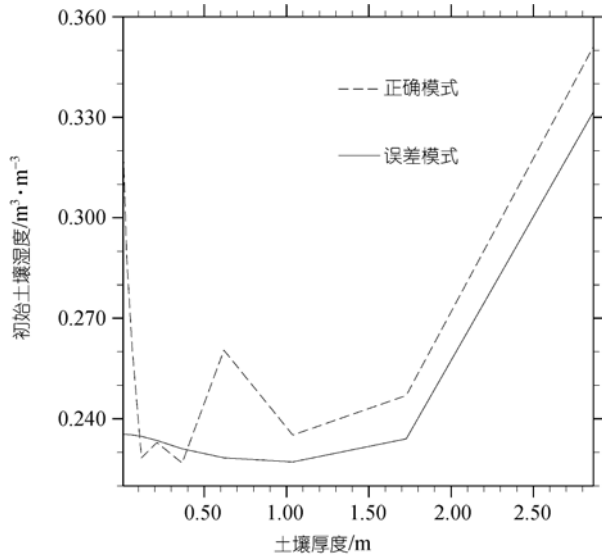


图 2 用于土壤湿度模式预报的初始土壤湿度(θ_0)

“Perfect”的初始土壤湿度以及“Perfect”的入渗时间序列积分 1092 步得到土壤湿度的时间序列作为“真实值”, 再在“真实值”的基础上加上 3% 的误差扰动得到观测的时间序列. 图 3 给出的是分别用(1)正确模式+正确初始值; (2)正确模式+误差初始值, 以及(3)误差模式+误差初始值积分所得到土壤表层的土壤湿度的时间序列, 可以看出 3 个时间序列有着较大的差异. 再以有误差的初始土壤湿度作为初始值, 以正确和误差入渗时间序列分别采用 POD/SVD-E4DVAR 以及

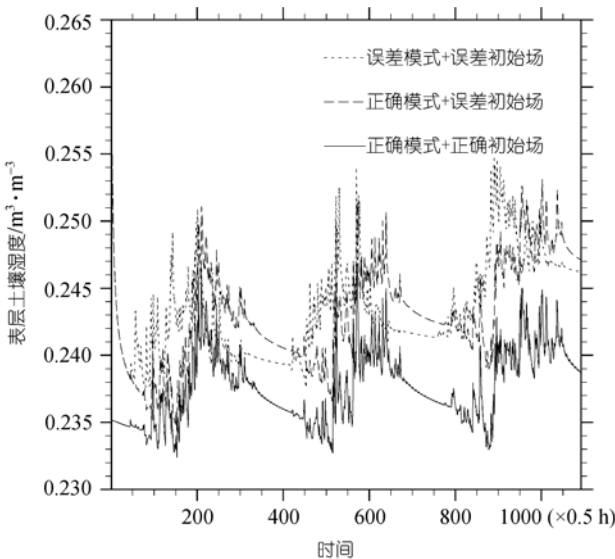


图 3 不同条件积分得到土壤表层土壤湿度($\theta(t)$)的时间序列

EnKF 方法进行两组同化试验(分别表示初始场有误差+模式预报无误差、初始场有误差+模式预报有误差两种情形,见表 2).一个同化窗口共有 7 个积分步(共有 156 个同化窗口),在每一组同化试验中又分别在同化窗口内分为在(1) 1, 2, 3, 4, 5, 6, 7 步; (2) 1, 3, 5, 7 步; (3) 1, 4, 7 步引入观测资料用以表示观测频次的强弱.如果在 4DVAR 中考虑背景场的约束,同化结果会与 B 矩阵的设置有很大关系,而 B 矩阵的确定带有一定的主观性,所以在我们的试验中两种种变分同化方法都不考虑背景场的约束,这样代价函数变得简单.

3.2 结果分析

定义如下的相对误差来反映分析场的精度:

$$E_n(\theta) = \frac{\sum_{i=1}^S (\theta_{n,i}^a - \theta_{n,i}^t)^2}{\sum_{i=1}^S (\theta_{n,i}^f - \theta_{n,i}^t)^2}, \tag{3.4}$$

表 2 对试验的说明

试验编号	初始场有无误差	观测频次	模式有无误差
1	有	7 次/窗	无
2	有	4 次/窗	无
3	有	3 次/窗	无
4	有	7 次/窗	有
5	有	4 次/窗	有
6	有	3 次/窗	有

其中 $E_n(\theta)$ 表示第 n 个同化窗口内土壤湿度分析场的相对误差, S 为第 n 个同化窗口内模式积分步数, $\theta_{n,i}^a, \theta_{n,i}^t$ 分别表示该同化窗内的分析值和真实值,而 $\theta_{n,i}^f$ 则表示模式直接模拟的预报值(对于试验 1~3, $\theta_{n,i}^f$ 为图 3 中正确模式+误差初始场,即采用“Perfect”入渗序列+“Imperfect”初始场的预报序列;而对于试验 4~6, $\theta_{n,i}^f$ 为图 3 中正确模式+正确初始场,即采用“Imperfect”入渗序列+“Imperfect”初始场的预报序列).

图 4 表现的是试验 1~3 的结果,不难看出 EnKF 由于无法在同时同化多个时刻的观测资料,其分析场的相对误差明显偏大,最大相对误差要超过 0.06(图中并没有画出),而 POD/SVD-E4DVAR 方法的分析场相对误差都要小于 EnKF,绝大多数同化窗内的相对误差都要小于 0.005.由于此时所用的预报模型是没有误差的,模式的预报场的误差只来自于初始值偏差的时间演变,两个变分方法(POD/SVD)同化的分析场的相对误差非常接近,但是还是能够看出 POD 方法由于能够捕捉到预报集合更多的能量,其同化精度要略高一些,相对误差也比 SVD 方法略微偏小.

图 5 给出的是试验 4~6 的分析场相对误差的对比结果.很显然,这个时候由于采用预报方式是采取模型有误差+初始场有误差,模式预报的误差不仅仅来自于初始场的时间演变,还在很大程度上取决于模式本身的不确定性.这种情况下,3 种方法的分析场的相对误差都有所加大,其中 EnKF 的相对误差的最大值超过 0.15(图中没有给出),与试验 1~3 相对比,POD-E4DVAR 方法的相对误差变化不大,显示出该方法在模式有无误差的情况下都能表现良好.而 SVD-E4DVAR 方法则稍微逊色一些,尤其在第 15~40 个同化窗内,其相对误差偏大,甚至与 EnKF 方法的相对误差相差无几,如:第 20 个同化窗附近,图 5(a), (b)中 SVD-E4DVAR 的相对误差超过 0.09, EnKF 的相对误差为 0.12,而 POD-E4DVAR 的相对误差仅为 0.04~0.05;在第 30~40 同化窗内,图 5(b)中 SVD-E4DVAR 的相对误差甚至超过了 0.15;在图 5(c)中 SVD-E4DVAR 的相对误差在第 20 个窗口附近也超过了 0.08, SVD-E4DVAR 表现得与 EnKF 方法几乎一致,而 POD-E4DVAR 方法的相对误差则都能控制 0.06 以

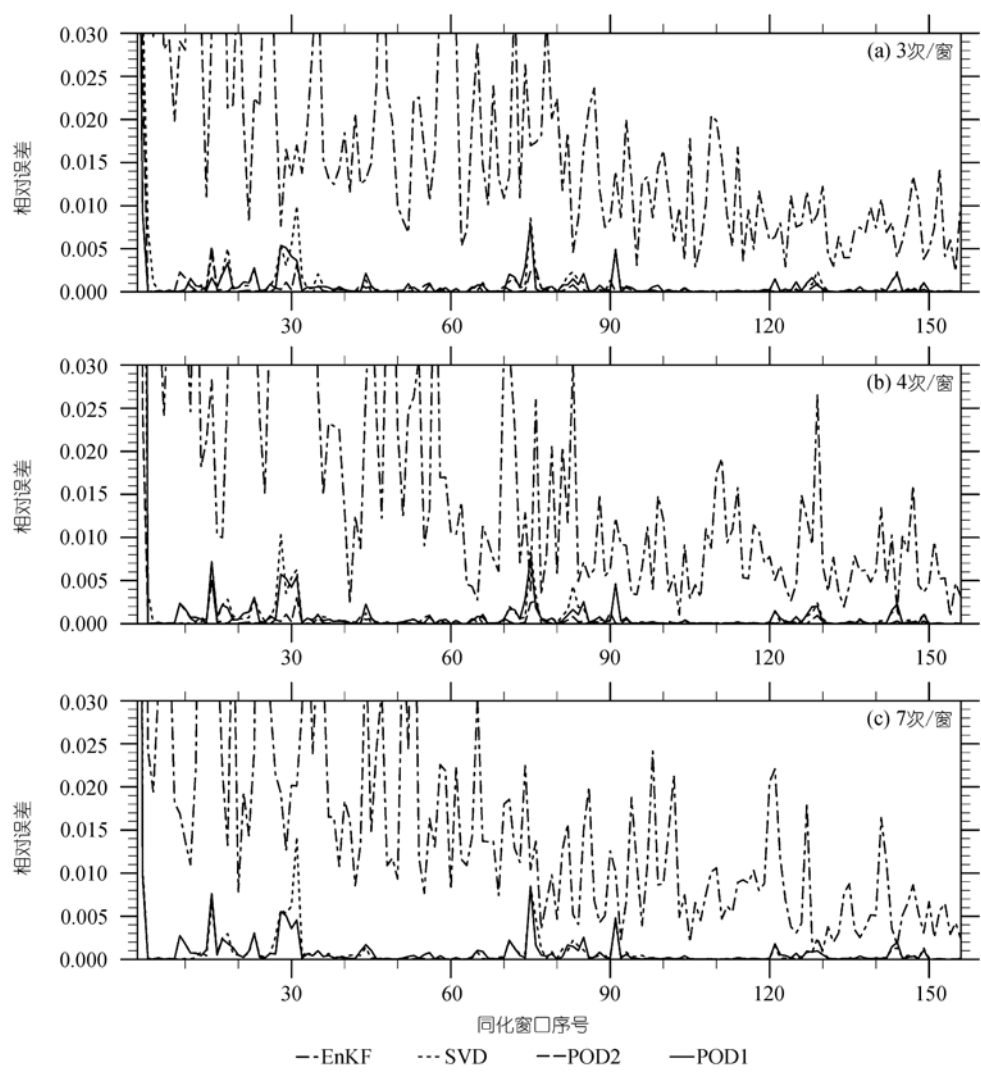


图4 POD/SVD-E4DVAR 和 EnKF 方法所同化的土壤湿度分析场的相对误差序列
预报误差仅来自初始场的误差

下. 由图 5 还可以看出, 观测频次的不同对 SVD-E4DVAR 方法影响较大, 观测频次为 7 次/窗(图 5(c))的 SVD-E4DVAR 方法所同化的分析场的相对误差要比观测频次为 4 次/窗内(图 5(b)), 3 次/窗(图 5(a))小一些, 尤其在第 15~40 同化窗内; 而观测频次对 POD-E4DVAR 方法的影响要相对小一些: 从图 5(a)~(c)上可以看出 POD-E4DVAR 方法所同化的分析场的相对误差变化不大.

就计算时间效率而言, 就这个个例而言(以上 3 种方法所采用样本个数均为 60, 且 POD/SVD 方法特征根的截断误差取为 5%, 也就是(2.7)式中的 γ 取为 0.95)所用时间的比例约为 1:1:30 (POD-E4DVAR:

SVD-E4DVAR:EnKF), 由此不难看出这种 POD-E4DVAR 方法应用的巨大潜力.

对 POD-E4DVAR 方法的两种取样方式而言, 从图 4 和 5 可以看出, POD1-E4DVAR 与 POD2-E4DVAR 方法的同化效果几乎一致, 由于在整个同化窗口内的每个时刻都进行取样能够更充分地反应模式预报的时空变化信息, 因此 POD2-E4DVAR 方法的同化效果要略好一点. 从实际应用上看, 随着控制空间维数的急剧增加, POD2-E4DVAR 方法的计算量会随之变大(当然若在每个积分都有观测信息的时候, 两种取样方式是一致的, 见图 4(c)和 5(c)), 从这个角度出发, POD1-E4DVAR 方法更为有效.

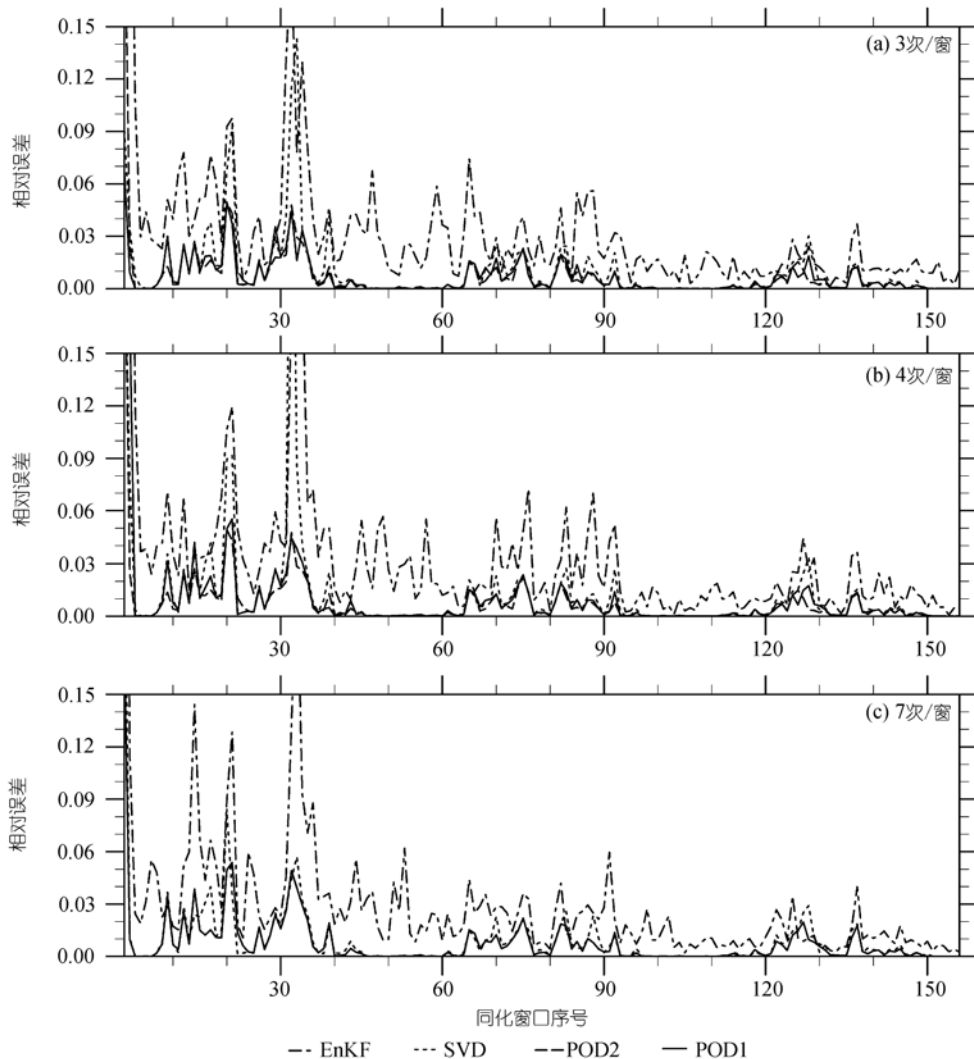


图 5 POD/SVD-E4DVAR 和 EnKF 方法所同化的土壤湿度分析场的相对误差序列

预报误差来自初始场的误差和预报模型的不确定性

4 讨论与小结

本文所发展的一个基于标准正交分解的显式四维变分同化方法与 Qiu 基于 SVD 分解的方法相似, 都能使控制变量显式地出现在代价函数中, 避免了伴随模式的运用, 使得同化过程变得简单. 用土壤湿度预报方程和人造资料进行的数值试验表明: 无论那种情形, 由于 POD/SVD-E4DVAR 变分方法能够在同一个时刻同化多个时间的观测数据, 它们的同化结果都要远远优于 EnKF 方法, 它们的分析场的相对误差都要远远小于 EnKF 方法. 在只有初始值存在误差而模式没有的情况下, POD/SVD-E4DVAR 方法的

同化效果极为相似, 或许是由于 POD 基能捕捉到样本空间更多的能量, POD-E4DVAR 的同化结果略好于 SVD-E4DVAR, 相对误差也略小于 SVD-E4DVAR; 而在初始场和模式都存在误差的情况下, POD-E4DVAR 方法的优势表现地较为明显, 其相对误差与模式不存在误差的情况下变化不大, 而 SVD-E4DVAR 方法则在某些同化窗内的同化误差变大, 甚至与 EnKF 方法的相对误差几乎一样. 另外, 观测频次 POD-E4DVAR 方法的同化效果影响较小, 而对 SVD-E4DVAR 方法的影响较大. 就计算时间效率而言, 数值试验表明 POD/SVD-E4DVAR 方法所用的计算时间大体相当, 都远远小于 EnKF 所耗费的计算时

间. 从上面分析不难看出 POD-E4DVAR 方法的优势 较大, 潜力不小.

致谢 感谢审稿专家提出的宝贵意见.

参考文献

- 1 Lewis L, Deber J. The use of adjoint equations to solve a variational adjustment problem with advective constraint. *Tellus Ser A-Dyn Meteorol Oceanol*, 1985, 37: 309—322
- 2 LeDimet F X, Talagrand O. Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. *Tellus Ser A-Dyn Meteorol Oceanol*, 1986, 38: 97—111
- 3 李新, 小池俊雄, 程国栋. 一个基于模拟退火法的陆面数据同化算法. *地球科学进展*, 2003, 18(4): 632—636
- 4 Kalnay E, Park S K, Pu Z X, et al. Application of the quasi-inverse method to data assimilation. *Mon Weather Rev*, 2000, 128: 864—875[DOI]
- 5 王斌, 赵颖. 一种新的资料同化方法. *气象学报*, 2005, 63: 694—700
- 6 Courtier P, Thepaut J N, Hollingsworth A. A strategy for operational implementation of 4D-Var using an incremental approach. *Quart J R Meteorol Soc*, 1994, 120: 1367—1388[DOI]
- 7 Qiu C J, Chou J F. Four-dimensional data assimilation method based on SVD: theoretical aspect. *Theor Appl Climatol*, 2006, 83: 51—57[DOI]
- 8 邱崇践, 张蕾, 邵爱梅. 一种显式四维变分资料同化方法. *中国科学 D 辑: 地球科学*, 2007, 37(5): 698—704
- 9 Pearson K. On lines and planes of closet fit to systems of points in space. *Phil Mag*, 1901, 2(6): 559—572
- 10 Loève M. *Probability Theory*. Princeton, NJ: van Nostrand, 1995
- 11 Ly H V, Tran H T. Modeling and control of physical processes using proper orthogonal decomposition. *Math Comp Model*, 2001, 33: 223—236[DOI]
- 12 Berkooz G, Holmes P, Lumley J. The proper orthogonal decomposition in the analysis of turbulent flows. *Ann Rev Fluid Mech*, 1993, 25: 777—786
- 13 Evensen G. Sequential data assimilation with a non-linear geostrophic model using Monte Carlo methods to forecast error statistics. *J Geophys Res*, 1994, 99(C5): 10143—10162[DOI]
- 14 Oleson K W, Dai Y, Bonan G, et al. Technical description of the community land model (CLM). NCAR Technical Note, NCAR/TN-461+STR, 2004. 186
- 15 Selker J S, Keller C K, McCord J T. *Vadose Zone Processes*. Boca Raton: Lewis Publishers, 1999
- 16 Clapp R B, Homberger G M. Empirical equations for soil hydraulic properties. *Water Resour Res*, 1978, 14: 601