

聂肃平, 朱江, 罗勇. 2010. 不同模式误差方案在集合 Kalman 滤波土壤湿度同化中的比较试验 [J]. 大气科学, 34 (3): 580–590. Nie Suping, Zhu Jiang, Luo Yong. 2010. Comparison experiments of different model error schemes in ensemble Kalman filter soil moisture assimilation [J]. Chinese Journal of Atmospheric Sciences (in Chinese), 34 (3): 580–590.

# 不同模式误差方案在集合 Kalman 滤波 土壤湿度同化中的比较试验

聂肃平<sup>1, 2, 3</sup> 朱江<sup>1</sup> 罗勇<sup>3</sup>

1 中国科学院大气物理研究所, 北京 100029

2 中国科学院研究生院, 北京 100049

3 国家气候中心, 北京 100081

**摘 要** 本文主要目的是探讨不同模式误差方案在土壤湿度同化中的性能。基于集合 Kalman 滤波同化方法和 AVIM (Atmosphere-Vegetation Interaction Model) 陆面模式, 利用理想试验对膨胀因子方案 (Covariance Inflation, 简称 CI)、直接随机扰动方案 (Direct Random Disturbance, 简称 DRD)、误差源扰动方案 (Source Random Disturbance, 简称 SRD) 等 3 种模式误差方案的同化效果进行了比较, 讨论了各方案在不同观测误差、观测层数、观测间隔情况下的同化性能。试验结果表明在观测误差估计完全准确的情况下, 3 种方案都能获得较好的同化效果, 并且 SRD 方案相对于真值的均方根误差最小。当观测误差估计不准确时, SRD 方案的同化效果仍能基本得以保持, 而 CI 和 DRD 方案则对观测误差估计更为敏感, 同化效果下降明显。当同化多层观测时, CI 和 DRD 方案由于难以保持不同层观测之间的匹配关系, 同化结果反而变差, 而 SRD 方案能有效协调同化多层观测, 增加观测层后同化结果有了进一步的改善。当观测时间间隔较大时, CI 和 DRD 方案的同化效果显著下降; 而 SRD 方案由于包含了一定的误差订正功能, 在观测稀疏时仍能保持较好的同化效果。

**关键词** 集合 Kalman 滤波 土壤湿度 模式误差 误差源 资料同化

**文章编号** 1006-9895 (2010) 03-0580-11

**中图分类号** P413

**文献标识码** A

## Comparison Experiments of Different Model Error Schemes in Ensemble Kalman Filter Soil Moisture Assimilation

NIE Suping<sup>1, 2, 3</sup>, ZHU Jiang<sup>1</sup>, and LUO Yong<sup>3</sup>

1 *Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing* 100029

2 *Graduate University of Chinese Academy of Sciences, Beijing* 100049

3 *National Climate Center, Beijing* 100081

**Abstract** The purpose of this paper is to explore the performances of different model error scheme in soil moisture data assimilation. Based on the ensemble Kalman filter (EnKF) and the atmosphere-vegetation interaction model (AVIM), point-scale analysis results for three schemes, 1) covariance inflation (CI), 2) direct random disturbance (DRD), and 3) source random disturbance (SRD), are combined under conditions of different observational error

**收稿日期** 2009-04-20, 2009-10-30 收修定稿

**资助项目** 国家自然科学基金资助项目 40905046, 财政部/科技部公益类 (气象) 行业专项 GYHY00706005, 中国气象局科技司“新一代气候系统模式的评估与改进”项目 2009-2012209

**作者简介** 聂肃平, 男, 1981 年出生, 博士生, 主要从事陆面资料同化研究。E-mail: niesp@cma.gov.cn

estimations, different observation layers, and different observation intervals using a series of idealized experiments. The results shows that all these schemes obtain good assimilation results when the assumed observational error is an accurate statistical representation of the actual error used to perturb the original truth value, and the SRD scheme has the least root mean square error (RMSE). Overestimation or underestimation of the observational errors can affect the assimilation results of CI and DRD schemes sensitively. The performances of these two schemes deteriorate obviously while the SRD scheme keeps its capability well. When the observation layers or observation interval increase, the performances of both CI and DRD schemes decline evidently. But for the SRD scheme, as it can assimilate multi-layer observations coordinately, the increased observations improve the assimilation results further. Moreover, as the SRD scheme contains a certain amount of model error estimation functions in its assimilation process, it also has a good performance in assimilating sparse-time observations.

**Key words** ensemble Kalman filter, soil moisture, model error, error source, data assimilation

## 1 引言

土壤湿度是地球气候系统中的一个重要变量,通过控制土壤-植被-大气交界面上的能量和水汽收支平衡,进而对整个气候系统产生显著影响(Shukla and Mintz, 1982)。目前,由于陆面过程模式模拟能力有限,并且受到大气强迫场和地表误差参数不确定性的制约,其模拟的土壤湿度与实际状况相比存在很大不确定性。正因为如此,能将各种不同来源、不同时空分辨率土壤湿度观测资料融合到陆面模式中的陆面资料同化技术,正逐渐在陆面研究和业务预报中扮演越来越重要的角色。

集合 Kalman 滤波方法(Ensemble Kalman Filter, 简称 EnKF)是 Evensen 在 1994 年首次提出的一种利用 Monte Carlo 方法预报误差统计特征的顺序资料同化方法(Evensen, 1994)。该方法继承了传统 Kalman 滤波方法背景场误差协方差随时间演变和流依赖(flow-dependent)的特点,并巧妙利用模式预报集合成员间差异的统计相关信息代表真实模式背景场误差协方差,有效地解决了 Kalman 滤波方法实际应用中背景误差协方差估计和预报困难的问题。与其他同化方法相比,EnKF 在土壤湿度同化中具有较多优势:(1)对于强非线性的陆面过程,能刻画出误差的非线性传播;(2)能方便地考虑不同来源的模式误差;(3)移植性强。基于这些优势和应用的灵活性,EnKF 已逐渐成为国内外土壤湿度同化研究中最有应用前景的方法(Reichle et al., 2002a; Crow and Wood, 2003; Reichle and Koster, 2005; 黄春林和李新, 2006a, 2006b; Ni-Meister et al., 2006; Zhang et al., 2006; 田向军和谢正辉, 2008; Yang and Koike, 2008; 张

生雷等, 2008)。

由于数值模式本身对自然界的刻画存在缺陷,因此,模式预报结果与其对应真实状态之间存在一定的预报误差(包含初始误差和模式误差)。目前,陆面模式中普遍忽略了模式误差的影响,因此,在 EnKF 的应用过程中,通过集合样本统计所得的背景场误差协方差比真实的背景场误差协方差缺少一个模式误差项,从而导致对模式背景场误差的低估。这一低估会削弱同化对观测信息的提取能力,并随着时间的推移最终导致同化中的“滤波发散”(filter divergence)现象。许多学者分析了造成滤波发散的原因(Houtekamer and Mitchell, 1998; van Leeuwen, 1999; Hamill et al., 2001),结果表明,对模式误差的忽视往往是造成滤波发散的主要原因。因此,如何在同化中合理地考虑模式误差的作用已经成为当前集合资料同化中的一个热点问题。

目前的集合同化中,通常通过将一定形式的噪声(noise)引入到预报集合样本中来代替缺失的模式误差项。对于单模式集合同化,常用的考虑模式误差的方法有:①膨胀背景场误差协方差(covariance inflation)方法,即人为增大背景场误差来促使同化结果向观测逼近(Anderson and Anderson, 1999; Hamill et al., 2001; Whitaker et al., 2004);②针对同化的变量,直接在该变量上加随机扰动,弥补集合样本估计得到的背景误差协方差(Reichle et al., 2002b; Francois et al., 2003);③分析模式误差来源在误差源上建立误差模型,通过各误差源提供模式误差(Reichle et al., 2002a; Crow and Van Loon, 2006)。目前,以上三种考虑模式误差的方法在集合同化研究中都有广泛的应用,但系统

地比较何种考虑模式误差方法更合理的研究,特别是在土壤湿度同化中的相关研究还相对较少。因此,在基于 EnKF 的土壤湿度同化中探讨这三种引入模式误差方法的效果和合理性,对于我们更好地理解 EnKF 方法及其在土壤湿度同化中的应用都具有重要意义。

本文基于一个中等复杂程度陆面过程模式 AVIM (Ji, 1995; 丹利, 2003), 通过理想试验分析, 讨论了利用 EnKF 进行土壤湿度同化时不同引入模式误差方法对同化结果的影响, 试图找出一种土壤湿度同化时合理考虑模式误差的方法, 从而为今后 EnKF 同化方法在实际土壤湿度资料(卫星观测或站点观测)同化中的应用提供一个合理参考。本文的主要内容如下: 第 2 节描述了理想试验的设计流程、陆面模式的简单介绍以及 EnKF 同化中三种不同的考虑模式误差的同化方案; 第 3 节给出同化试验及其结果分析; 第 4 节给出本文的结论和讨论。

## 2 方法和模式介绍

当前, 很多学者都采用理想试验方法来进行土壤湿度同化的相关研究 (Reichle et al., 2002a, 2002b; Crow, 2003)。由于这些试验设计简单, 并且避免了同化实际土壤湿度观测资料时需要各种复杂因素都加以详细考虑的困难(例如, 在实际资料同化中, 不同源模式误差和观测误差的统计特征很难得到准确估计), 因此是同化研究中评估不同方法同化性能的十分有用的工具。在本文中, 为了使研究重点放在我们所关注的问题上, 我们也将采用理想试验的方式来讨论不同模式误差引入方案在 EnKF 土壤湿度同化中的性能。

图 1 给出了理想试验示意图。首先将一组模式运行结果作为“真实场”, 然后在真实场上以特定误差统计特征加入模式误差和观测误差来分别生成模式场和观测场。同化时, 再通过估计模式误差和观测误差信息得到相应背景场和观测场的集合, 最后利用 EnKF 同化后, 得到集合分析场。EnKF 理想试验的显著优点就是同化中观测误差和模式误差的大小从统计意义上可以精确获得。通过给出不同的误差统计特征, 我们就能方便地评估多种误差估计情况下不同方案的同化效果。

### 2.1 陆面模式

植被大气相互作用模式 AVIM 是国内自主研

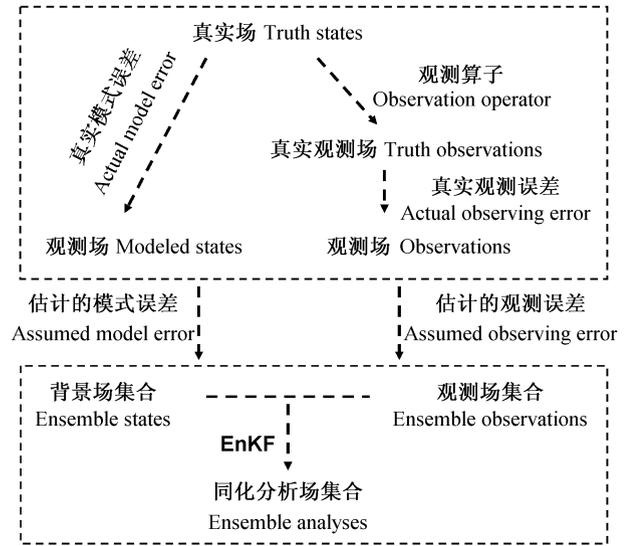


图 1 理想试验示意图

Fig. 1 Schematic diagram of the idealized experiment

发的较成熟的陆面过程模式, 在陆面过程模拟以及陆气相互作用等研究方面都有着广泛的应用 (Li and Ji, 2001; Dan et al., 2002, 2005, 2008)。AVIM 中土壤在垂直方向分为 3 层: 0~0.1 m 为表层, 0.1~0.9 m 为根区层, 0.9~3.6 m 为深层。深层土壤湿度取为气候态常量, 表层和根区层的土壤湿度方程 (丹利, 2003) 为:

$$\frac{d\omega_1}{dt} = \frac{1}{\theta_s d_1} \left[ P_g - \frac{1}{\rho_w} (E_g + E_{tr1}) \right] + \frac{2D_{12}(\omega_2 - \omega_1)}{d_1(d_1 + d_2)} + \frac{K_{ws}}{\theta_s d_1} \omega_{12}^{2b+3} \quad (1)$$

$$\frac{d\omega_2}{dt} = -\frac{E_{tr2}}{\theta_s d_2 \rho_w} - \frac{2D_{12}(\omega_2 - \omega_1)}{d_2(d_1 + d_2)} + \frac{2D_{23}(\omega_3 - \omega_2)}{d_2(d_2 + d_3)} - \frac{K_{ws}}{\theta_s d_2} \omega_{12}^{2b+3} + \frac{K_{ws}}{\theta_s d_2} \omega_{23}^{2b+3} \quad (2)$$

其中,  $\omega_1$  和  $\omega_2$  为上两层土壤湿度 (单位:  $\text{cm}^3/\text{cm}^3$ ),  $d_1$ 、 $d_2$ 、 $d_3$  为三层土壤厚度,  $P_g$  为降水,  $E_g$  是地表的蒸发,  $E_{tr1}$  和  $E_{tr2}$  为两层土壤中由于冠层蒸腾作用而引起的水分流失,  $\theta_s$  为土壤孔隙度,  $k_{ws}$  为饱和导水率, 指数  $b$  为与土壤类型有关的经验参数,  $D_{12}$ 、 $D_{23}$  为土壤层间的水分扩散率。

### 2.2 集合 Kalman 滤波方法

集合 Kalman 滤波 (EnKF) 方法的最大特点是通过完全非线性模式预报得到的一组集合样本来统计同化中的背景场误差协方差矩阵 (Evensen, 1994)。EnKF 的同化过程可主要分为“集合分析”

部分和“集合预报”部分。在有观测时刻, 其分析过程如下 (Evensen, 2003)。

首先, 对于模式变量  $\mathbf{x}$ , 通过一组模式预报场得到该变量的集合样本矩阵  $\mathbf{X}$ :

$$\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) \in \mathfrak{R}^{n \times N}, \quad (3)$$

其中,  $n$  是模式状态向量数,  $N$  是集合样本数。

集合扰动矩阵可定义为:

$$\mathbf{X}' = \mathbf{X} - \bar{\mathbf{X}} = \mathbf{X}(\mathbf{I} - \mathbf{1}_N), \quad (4)$$

其中,  $\bar{\mathbf{X}} = \mathbf{X}\mathbf{1}_N$  为集合均值矩阵,  $\mathbf{1}_N \in \mathfrak{R}^{N \times N}$  为每个元素都等于  $1/N$  的矩阵。

这样, 由集合样本统计得到的背景误差协方差矩阵  $\mathbf{P}_e \in \mathfrak{R}^{n \times n}$  可定义为:

$$\mathbf{P}_e = \frac{\mathbf{X}'(\mathbf{X}')^T}{N-1}. \quad (5)$$

然后, 对于给定观测  $\mathbf{y} \in \mathfrak{R}^m$ , 定义  $N$  个扰动观测向量  $\mathbf{y}_i = \mathbf{y} + \boldsymbol{\varepsilon}_i$ ,  $i=1, \dots, N$ , 此时观测集合样本矩阵可表示为:

$$\mathbf{Y} = (\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N) \in \mathfrak{R}^{m \times N}, \quad (6)$$

并且, 由集合样本统计得到的观测误差协方差矩阵  $\mathbf{R}_e$  为:

$$\mathbf{R}_e = \frac{\boldsymbol{\gamma}\boldsymbol{\gamma}^T}{N-1}, \quad (7)$$

其中,  $\boldsymbol{\gamma} = (\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \dots, \boldsymbol{\varepsilon}_N) \in \mathfrak{R}^{m \times N}$  为观测扰动集合。

最后, 得到 EnKF 的分析方程为:

$$\mathbf{X}^a = \mathbf{X}^f + \mathbf{P}_e \mathbf{H}^T (\mathbf{H} \mathbf{P}_e \mathbf{H}^T + \mathbf{R}_e)^{-1} (\mathbf{Y} - \mathbf{H} \mathbf{X}^f), \quad (8)$$

其中,  $\mathbf{X}^f$  为前一时刻预报场,  $\mathbf{X}^a$  为当前时刻分析场,  $\mathbf{K} = \mathbf{P}_e \mathbf{H}^T (\mathbf{H} \mathbf{P}_e \mathbf{H}^T + \mathbf{R}_e)^{-1}$  为 Kalman 增益矩阵,  $\mathbf{H}$  为观测算子。

在无观测时刻, EnKF 主要利用模式动力演变进行多样本的集合预报。下一观测时刻集合平均的背景场通过前一时刻样本分析场的预报得到:

$$\overline{\mathbf{x}_{t+1}^b} = \overline{f(\mathbf{x}_t^a)} = f(\overline{\mathbf{x}_t^a}) + \mathbf{q}, \quad (9)$$

这里, “ $\overline{\quad}$ ” 表示集合样本平均,  $\mathbf{q}$  表示当模式积分算子  $f$  为非线性时产生的非线性误差项。在 EnKF 同化中, 如果忽略模式误差的影响, 集合样本统计得到的背景场误差协方差  $\mathbf{P}_e$  就会由于缺少  $\mathbf{q}$  这一项而被低估。因此, 要更合理地应用 EnKF, 同化过程中就必须考虑模式误差项  $\mathbf{q}$  的影响。

### 2.3 模式误差引入方案

本文采用三种不同方案在 EnKF 中考虑模式误差的影响, 并通过理想试验分析检验各方案的同化性能。

(1) 误差协方差膨胀 (covariance inflation) 方案: 误差协方差膨胀方案 (简称 CI 方案) 是考虑模式误差的一种最简单的方法。该方法通过在背景场误差协方差矩阵中引入一个方差膨胀因子 (inflation factor), 从而人为增大背景场误差的量值以避免其在同化过程中被低估。CI 方案的形式如下:

$$\mathbf{K} = (\beta \mathbf{P}_e + \mathbf{P}_0) \mathbf{H}^T (\mathbf{H} (\beta \mathbf{P}_e + \mathbf{P}_0) \mathbf{H}^T + \mathbf{R}_e)^{-1}, \quad (10)$$

其中,  $\beta$  为方差膨胀因子,  $\mathbf{P}_0$  为背景场误差协方差矩阵中引入的一个表示最小方差值的对角矩阵, 用来在同化过程中始终保持一定的背景场误差。本文不同试验中的  $\mathbf{P}_0$  根据相应观测误差的大小来确定。

(2) 直接随机扰动 (direct random disturbance) 方案: 直接随机扰动方案 (简称 DRD 方案) 通过在模式积分过程中直接在土壤湿度上加随机扰动, 来弥补集合样本统计对背景场误差协方差的低估。我们采用一个线性一阶 Markov 随机模型来模拟模式误差随时间的演变 (Evensen, 2003):

$$\mathbf{q}_t = \alpha \mathbf{q}_{t-1} + \sqrt{1 - \alpha^2} \mathbf{w}_t, \quad (11)$$

其中,  $\mathbf{q}_t$  代表  $t$  时刻在土壤湿度上的模式误差,  $\mathbf{w}_t$  是  $t$  时刻的一组随机扰动场,  $\alpha$  代表随机误差的时间相关系数。

模式积分中土壤湿度的演变可表示成如下形式:

$$\mathbf{x}_t = f(\mathbf{x}_{t-1}) + \sqrt{\Delta t} \sigma \mathbf{q}_t, \quad (12)$$

其中,  $f$  是模式非线性积分算子,  $\Delta t$  是模式误差步长,  $\sigma$  为误差的标准差。

(3) 误差源扰动 (source random disturbance) 方案: 误差源扰动方案 (简称 SRD 方案) 通过在土壤湿度误差源 (大气强迫场、模式参数等) 上引入随机扰动, 利用模式积分过程中的物理约束关系将各误差源上的不确定性传递到土壤湿度中, 从而实现同化中背景场误差协方差矩阵低估的弥补。这里, 根据敏感性试验结果, 我们主要选取对土壤湿度影响较大的大气强迫场 (主要是降水场和辐射场) 和模式参数 (主要是饱和导水率、饱和水势和土壤类型参数  $b$ ) 作为土壤湿度的误差源, 并同样采用一阶 Markov 随机模型来模拟各误差源误差随时间的演变:

$$\boldsymbol{\mu}_t = \alpha_\mu \boldsymbol{\mu}_{t-1} + \sqrt{1 - \alpha_\mu^2} \boldsymbol{\delta}_t, \quad (13)$$

其中,  $\boldsymbol{\mu}$  代表了大气强迫场或模式参数上引入的误差,  $\alpha_\mu$  表示其时间相关系数,  $\boldsymbol{\delta}_t$  则表示强迫场或模

式参数在  $t$  时刻引入的随机扰动场。

此时，模式积分过程中考虑误差后的各误差源可表示为：

$$\mathbf{x}_{\mu,t} = \mathbf{x}_{\mu,t}^i + \sqrt{\Delta t_{\mu}} \sigma_{\mu} \boldsymbol{\mu}_t, \quad (14)$$

其中， $\mathbf{x}_{\mu,t}^i$ 、 $\mathbf{x}_{\mu,t}$ 、 $\sqrt{\Delta t_{\mu}}$ 、 $\sigma_{\mu}$  表示加误差前的原始值（大气强迫场或参数场）、引入误差后的值、引入误差的步长、相应的误差标准差。

此时，模式积分过程中土壤湿度的演变可表示为：

$$\mathbf{x}_t = \mathbf{f}(\mathbf{x}_{t-1}, \mathbf{x}_{\mu,t}^i) + \mathbf{f}(\sqrt{\Delta t_{\mu}} \sigma_{\mu} (\alpha_{\mu} \boldsymbol{\mu}_{t-1} + \sqrt{1 - \alpha_{\mu}^2} \boldsymbol{\delta}_t)). \quad (15)$$

(15) 式表明，各误差源上引入的误差通过模式物理约束作用后传递到土壤湿度中，同化中考虑模式误差的同时也保证了各个状态变量之间的平衡关系。

### 3 同化试验及结果分析

#### 3.1 试验背景和设计

作为检验不同模式误差方案效果的初步研究，为简单起见，我们在“点尺度”上进行试验，以避免水平空间同化中对水平相关性考虑方式的不同可能对结果分析产生的影响。选取淮河流域能量与水循环试验 HUBEX (Fujiyoshi et al., 2006) 中蒋集站作为理想试验的背景站点，试验时间为 3 年（1998 年 1 月 1 日至 2000 年 12 月 31 日），试验前进行 2 年的 spin-up 积分。模式的大气强迫场来自 NCEP 再分析资料集 1 (Kalnay et al., 1996)，利用 NCEP 强迫场和 AVIM 在蒋集站对应的默认参数

运行模式，得到的各层土壤湿度值作为理想试验的“真实场”。随后，分别在土壤湿度初始场上引入初始误差，以及降水场、辐射场、饱和导水率、参数  $b$ 、饱和水势等模式误差源上引入模式误差生成“模拟场”。表 1 给出了真实场和模拟场的设置。模拟场中降水和辐射由在对应真实场加相对于时间的随机扰动误差得到，即每年都生成一组零均值的高斯随机扰动，通过将扰动对应分配到每天得到考虑随机误差后的模拟强迫场。图 2 给出了试验中模拟相对于真值的误差。从图 2 可以看出，在前 3 个月，模拟偏小，这主要是受模拟场初值估计偏小的影响，随后，模拟则偏大并随季节波动，这主要是由于后期初值影响减弱，参数误差和强迫场误差起主要作用的原因。理想试验中的土壤湿度“观测场”通过在真实场上加标准差为 5% (0.02 cm<sup>3</sup>/cm<sup>3</sup>) 的白噪声随机扰动得到，试验中集合样本个数取为 100。

本文主要采用均方根误差 (RMSE) 和集合样本离散度 (Spread) 对试验结果进行评估。因理想试验中真值已知，我们定义  $R_t$  和  $R_o$  来分别衡量土壤湿度同化值相对于真值和观测的 RMSE； $S_d$  用来衡量试验中各集合样本相对于集合均值的离散度。

$$R_t = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - x^t)^2}, \quad (16)$$

$$R_o = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - y)^2}, \quad (17)$$

表 1 理想试验中真实场和模拟场的设置

Table 1 Setup of truth states and simulation states in the idealized experiments

变量 & 参数 (单位)	真实场	模拟场
降水 (mm/d)	NCEP 再分析数据集 1	加零均值高斯扰动，标准差为真实值的 50%
长波、短波辐射 (W/m <sup>2</sup> )	NCEP 再分析数据集 1	加零均值高斯扰动，标准差为真实值的 30%
初始土壤湿度 (cm <sup>3</sup> /cm <sup>3</sup> )	2 年 spin-up 值	0.12
饱和导水率 (m/s)	5.23 × 10 <sup>-6</sup>	5.0 × 10 <sup>-5</sup>
参数 $b$	4.74	12.0
饱和水势 (m)	-0.218	-0.7

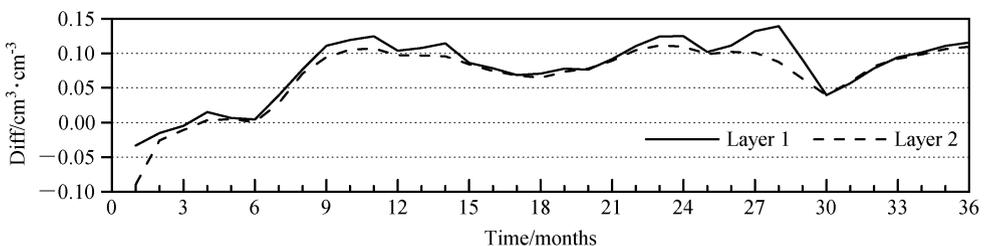


图 2 两层土壤湿度模拟值相对真实值的误差序列

Fig. 2 Time series of differences of soil moisture in the top two layers between simulation and truth

$$S_d = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2}, \quad (18)$$

以上三个式子中,  $x_i$  代表同化后土壤湿度的集合样本值,  $x^t$  代表真值,  $y$  代表观测值,  $\bar{x}$  代表集合样本均值。本文比较了 CI 方案、DRD 方案和 SRD 方案在不同观测误差、观测层数、观测间隔情况下的同化结果。试验中, CI 方案中最优膨胀因子通过敏感性试验得到 (图 3), 当其为 1.1 时, 得到的  $R_t$  最小, 本文中取 (10) 式中的  $\beta$  为 1.1; DRD 方案考虑了一定的土壤湿度误差时间相关性, 表层取为 10 天, 深层取为 20 天, 即 (11) 式中  $\alpha$  对表层取 0.9, 对深层取 0.95; SRD 方案中则不考虑误差源误差的时间相关性, 即 (13) 式中  $\alpha_\mu$  取为 0。试验共有五组, 每组试验都包含 3 种模式误差方案的试验, 具体设计参见表 2。

### 3.2 试验结果分析

#### 3.2.1 基准试验

理想试验中, 当同化中给定的观测误差统计特征和观测本身的误差特征一致时, 得到的同化结果最好。试验一首先比较了 3 种方案在此情况下的同化结果, 并以该试验作为后续试验的参考试验, 在此称为基准试验。

图 4 给出了基准试验的结果。从图 4a 的相对于真值的  $R_t$  时间序列可以看出, 当观测误差准确估计时, 3 种方案的同化值都能很好地接近真值。

SRD 方案的  $R_t$  在开始几个月内相对于其他两个方案略微偏大, 随后逐渐减小到  $0.01 \text{ cm}^3/\text{cm}^3$  以下, 在 9 个月以后, SRD 方案的结果相对于 CI 和 DRD 方案都更接近真值。

按照 Kalnay (2003) 的定义, 一个好的集合预报样本需要保证其均值相比模拟更接近真值, 且样本的范围 (离散度) 要能合理包含均值相对真值的偏差。从图 4b 的样本离散度可以看出, 3 种方案的  $S_d$  与  $R_t$  在整个时间序列上都能较好匹配, 说明在观测误差特征准确估计的情况下, 基于这 3 种方案的 EnKF 同化都能给出很好的同化结果。

从图 4c 的  $R_o$  结果可以看出, CI 和 DRD 方案的结果都十分接近观测 ( $0.01 \text{ cm}^3/\text{cm}^3$  左右), 而 SRD 方案则与观测存在较大偏差 ( $0.02 \text{ cm}^3/\text{cm}^3$  左右)。由于观测相对真值包含了  $0.02 \text{ cm}^3/\text{cm}^3$  的随机误差, 这一结果表明, CI 和 DRD 方案在同化时对观测中包含的“信息”和“误差”的区分能力不足, 使得同化结果中包含了大量的观测误差。而 SRD 方案采用的在误差源引入模式误差的做法, 充分利用模式自身物理约束关系来对观测中的有用信息和误差噪声进行有效区分, 在同化中既能有效提取观测资料中的有用信息, 又能有效避免对观测误差的过多引入, 因此使得同化结果不是简单的向观测靠近, 而是更合理地去接近真值。

表 2 试验设计列表

Table 2 The detailed setup of the idealized experiments

试验	观测时间间隔/d	观测层次	估计的观测误差标准差	试验目的
试验一	1	1	5.0%	检验在观测误差准确估计情况下, 3 种模式误差方案的同化效果
试验二	1	1	2.5%	检验在对观测误差低估的情况下, 3 种模式误差方案的同化效果
试验三	1	1	10.0%	检验在对观测误差高估的情况下, 3 种模式误差方案的同化效果
试验四	1	2	5.0%	检验 3 种模式误差方案在同化多层观测时的同化效果
试验五	10	1	5.0%	检验观测稀疏时 3 种模式误差方案的同化效果

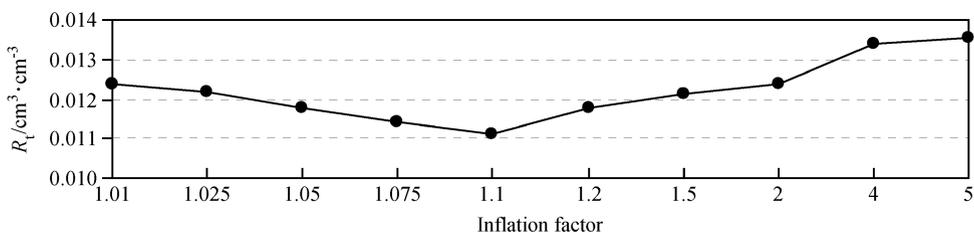
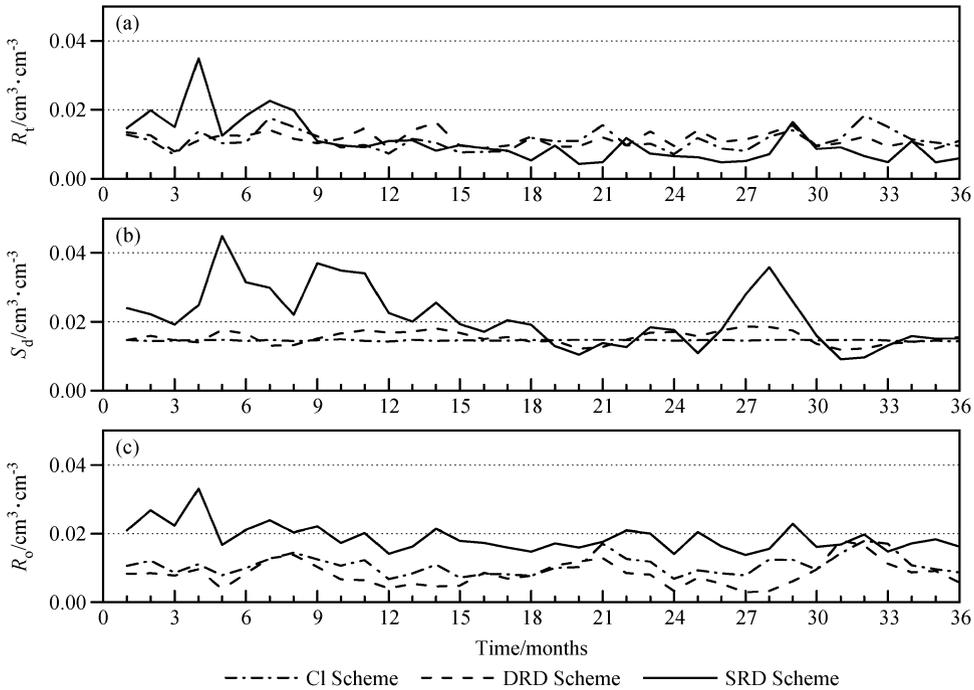


图 3 CI 方案中不同膨胀因子的敏感性试验

Fig. 3 Sensitivity tests for different inflation factors in the Covariance Inflation (CI) scheme

图4 基准试验结果: (a)  $R_t$ ; (b)  $S_d$ ; (c)  $R_o$ 。Fig. 4 The results for the benchmark experiment: (a)  $R_t$ ; (b)  $S_d$ ; (c)  $R_o$ 。表3 试验一、试验二、试验三中3种模式误差方案的  $R_t$ ,  $S_d$  和  $R_o$ 。Table 3  $R_t$ ,  $S_d$ , and  $R_o$  for different model error schemes in Expt 1, Expt 2, and Expt 3

	试验一 (基准试验)			试验二 (观测误差估计不足)			试验三 (观测误差估计过度)		
	CI	DRD	SRD	CI	DRD	SRD	CI	DRD	SRD
$R_t/\text{cm}^3 \cdot \text{cm}^{-3}$	0.0111	0.0115	0.0097	0.0136	0.0131	0.0109	0.0155	0.0112	0.0091
$S_d/\text{cm}^3 \cdot \text{cm}^{-3}$	0.0145	0.0154	0.0188	0.0092	0.0088	0.0068	0.0187	0.0249	0.0145
$R_o/\text{cm}^3 \cdot \text{cm}^{-3}$	0.0108	0.0083	0.0188	0.0042	0.0047	0.0112	0.0199	0.0129	0.0169

### 3.2.2 观测误差估计不准确试验

EnKF 进行土壤湿度同化时需要估计观测误差, 在基准试验中我们假设对观测误差统计特征的估计完全准确。然而在实际应用中, 我们仅能够对观测误差进行猜测, 不可避免会有猜测不准的情况。因此在试验二和试验三中, 我们考察在当对观测误差猜测不准确时, 不同引入模式误差方案的效果。这两组试验中, 除同化中对观测误差的估计不同以外, 其他都与基准试验相同 (同化的观测资料相同)。试验二假设对观测误差估计不足, 试验三则假设过分估计了观测误差。

图5和图6给出了试验二和试验三的结果。试验二中, 由于对观测误差估计不足, 3个方案中EnKF都给予观测更多权重, 因此3种方案的结果相对于基准试验都更加靠近观测 (图5c), 且包含

的不确定性也更小 ( $S_d$  更小, 图5b)。在试验三中, 由于对观测误差估计过大, EnKF会更多的偏向背景场: 对于CI和DRD方案, 由于对观测资料中过高估计的虚假误差不能准确辨认, 只能通过增加背景值的权重来平衡增大的观测误差, 因此使得同化效果明显下降 (图6a), 而且分析值的不确定性也显著增加 ( $S_d$  增大, 图6b); 而SRD方案由于能在一定程度上区分观测中的有用信息和误差 (噪声), 过高估计的虚假观测误差对其的影响并不明显, 因此该方案同化结果的误差和基准试验相比相差不大 (见表3)。从表3中还可以看出, 在观测误差估计同样不准确的情况下, SRD方案得到  $R_t$  都是最小, 这一结果表明, SRD方案在观测误差估计不准确时的同化效果仍然优于CI和DRD方案, 显示出良好的性能。

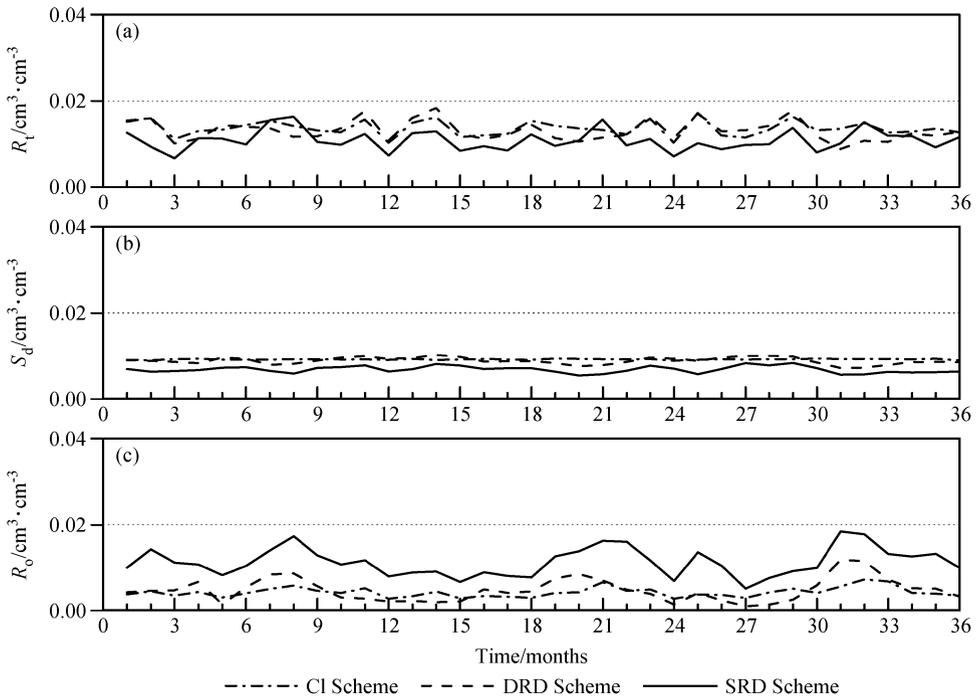


图 5 试验二的结果：(a)  $R_l$ ；(b)  $S_d$ ；(c)  $R_o$

Fig. 5 The results for Expt 2: (a)  $R_l$ ；(b)  $S_d$ ；(c)  $R_o$

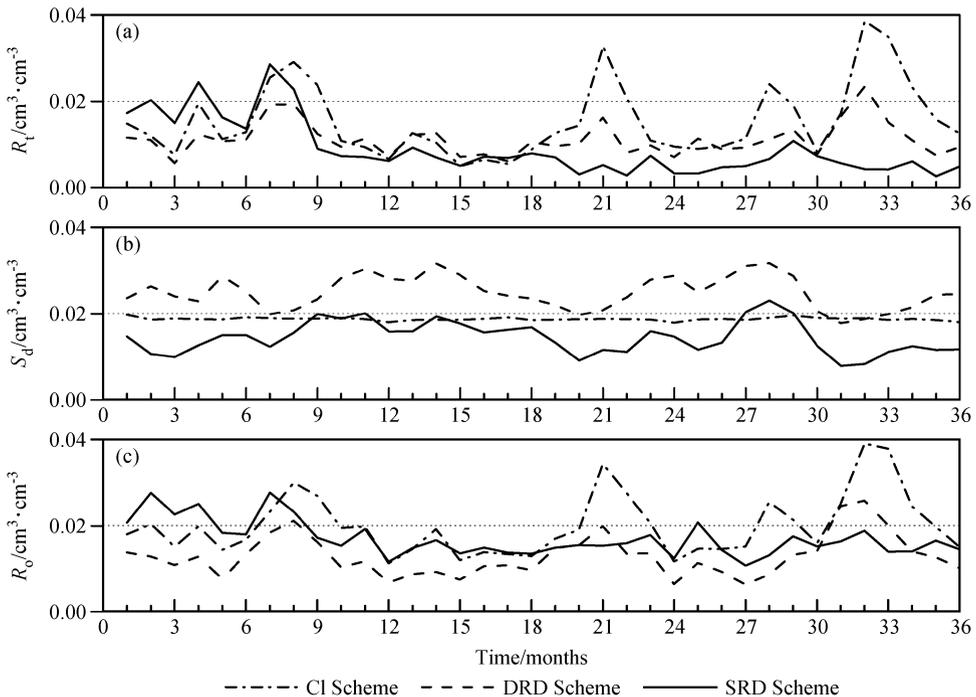


图 6 试验三的结果：(a)  $R_l$ ；(b)  $S_d$ ；(c)  $R_o$

Fig. 6 The results for Expt 3: (a)  $R_l$ ；(b)  $S_d$ ；(c)  $R_o$

### 3.2.3 同化两层观测试验

试验四检验了 3 种方案在同化多层土壤湿度观测时的性能。这里对表层和根区层两层土壤湿度观

测进行同化，且根区层的真实观测误差取为和表层相同。同化两层观测时，CI 方案考虑了对根区层也引入膨胀因子来考虑该层模式误差，DRD 方案

增加了对根区层土壤湿度的随机扰动, SRD 方案由于是在误差源考虑模式误差, 在此不引入额外扰动。除此以外, 3 个方案的设置都和基准试验相同。

表 4 给出了试验四的同化结果。与基准试验类似, SRD 方案得到的  $R_t$  在两层土壤湿度中都是最小。并且相对于其他两个方案, SRD 方案所得的  $R_t$  在表层减小了 50% 左右, 在根区层更是减小 60% 以上, 依然反映出更好的同化性能。而通过与基准试验(表 3)比较可以发现, CI 方案和 DRD 方案在同化了根区层观测后, 得到的  $R_t$  反而更加偏大, SRD 方案则有了明显改善。由于 SRD 方案能充分利用模式物理约束关系, 保证了同化时新引入的根区层观测中的有用信息能更好地被同化系统所利用, 从而使得此时的同化结果在仅同化表层观测的基础上有了进一步的改善。而 CI 和 DRD 方案由于很难区分观测中的有用信息和噪声, 因此使得同化多层观测时不同层观测中的误差反而相互影响,

导致在同化了根区层观测后, 表层结果反而由于受到根区层观测中误差的影响而变得更差。 $S_d$  和  $R_o$  的结果也能反映出这一影响, CI 和 DRD 方案的  $R_o$  与基准试验相比都偏小, 说明同化结果更多地向有误差的观测靠近; 而集合离散度则更大, 反映出集合样本均值中具有的不确定性相对于基准试验也变得更大。

以上结果表明, 在同化多层观测时, CI 和 DRD 方案都存在明显缺陷, 而 SRD 方案由于能利用模式本身来约束引入的模式误差, 因此能合理同化不同层的观测信息, 得到更好的同化结果。

3.2.4 稀疏观测试验

由于实际土壤湿度站点观测在时间上的分辨率都比较低 (Nie et al., 2008), 因此在试验五中, 我们还考察了 3 种方案在观测间隔较大情况下的同化效果。试验五中除了 10 天才进行一次同化外, 其他设置与基准试验相同。

图 7 给出了试验五的结果。可以看出, 由于同

表 4 同表 3, 但为试验四和试验五  
Table 4 Same as Table 3 but for Expt 4 and Expt 5

	试验四(表层)			试验四(根区层)			试验五(稀疏观测)		
	CI	DRD	SRD	CI	DRD	SRD	CI	DRD	SRD
$R_t/\text{cm}^3 \cdot \text{cm}^{-3}$	0.0120	0.0144	0.0064	0.0105	0.0150	0.0043	0.0397	0.0333	0.0107
$S_d/\text{cm}^3 \cdot \text{cm}^{-3}$	0.0147	0.0172	0.0095	0.0137	0.0185	0.0044	0.0132	0.0467	0.0213
$R_o/\text{cm}^3 \cdot \text{cm}^{-3}$	0.0097	0.0058	0.0137	0.0081	0.0029	0.0156	0.0425	0.0348	0.0186

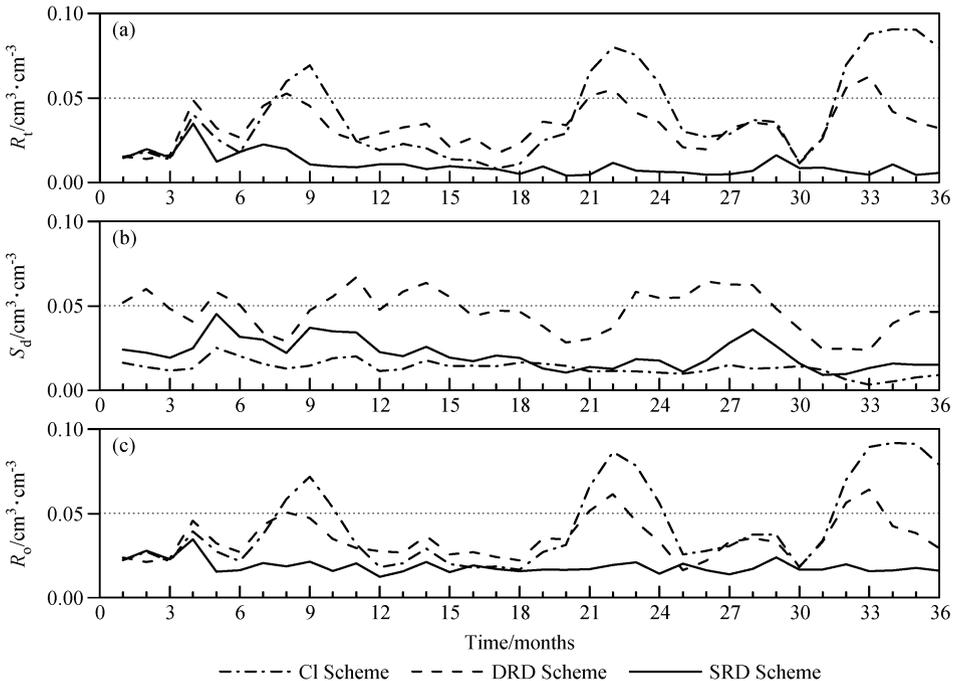


图 7 试验五的结果: (a)  $R_t$ ; (b)  $S_d$ ; (c)  $R_o$   
Fig. 7 The results for Expt 5: (a)  $R_t$ ; (b)  $S_d$ ; (c)  $R_o$

化次数减少, CI 和 DRD 方案得到的分析结果相对于真值和观测的误差都明显增加, 同化性能明显下降; 而 SRD 方案的结果则和基准试验的类似, 仍然保持了较好的同化效果。从表 4 的均方根误差也可以看出来, 观测间隔的增加使得 CI 和 DRD 方案的误差都有了明显增加, 而 SRD 方案却没有太大变化。由于 CI 和 DRD 方案都仅在同化时刻才对背景场进行订正, 如果没有观测的持续引入, 同化后的结果很快就会重新回到背景状态, 从而使得误差显著增加。而 SRD 方案是从源上考虑模式误差的特性, 在同化过程中还能够一定程度上对特定误差源中的误差进行订正(如参数误差订正), 从而能在同化后进一步改善模式预报效果, 因此, 即使在观测稀疏时仍然能有效地保持结果向真值靠近。

利用 SRD 方案在 EnKF 同化过程中考虑对特定误差源误差进行订正是目前土壤湿度同化中一个值得关注的问题, 对于如何更好地同化实际土壤湿度稀疏站点资料有着重要的意义, 由于篇幅所限, 对该问题的讨论将另文详细论述。

## 4 结论与讨论

作为一种以 Monte Carlo 方法为基础的同化方法, EnKF 能更方便地同化过程中考虑各种不同来源模式误差的影响, 因此在陆面土壤湿度同化的应用中具有明显优势。随着 EnKF 方法的发展, 如何在同化中合理考虑各种误差源模式误差影响的问题正逐渐得到重视和关注。本文基于 EnKF 方法, 利用一系列理想试验, 对 3 种模式误差引入方案(CI 方案直接在协方差矩阵上乘膨胀因子、DRD 方案直接在土壤湿度上加随机扰动、SRD 方案在土壤湿度误差源上引入随机扰动)进行了比较分析, 探讨了各方案在不同观测误差、观测层数、观测间隔情况下的同化性能, 具体结论如下:

(1) 在观测误差估计准确的情况下, 3 种方案都能获得良好的同化效果, 并且 SRD 方案所得结果相对于真值的均方根误差最小。

(2) 当对观测误差的估计不准确时, CI 方案和 DRD 方案的同化效果都有显著的下降, 而 SRD 方案的效果则基本能得以保持。CI 方案和 DRD 方案对观测误差估计比 SRD 方案更为敏感, SRD 方案在观测误差估计准确性上具有更好的“容错”能力。

(3) 当同化多层观测时, CI 方案和 DRD 方案

由于很难保持不同层观测之间的匹配关系, 在增加观测层数后同化效果反而变差; 而 SRD 方案则可以协调同化多层观测, 使得同化根区层观测后对表层的结果有了进一步的改善。

(4) 当观测时间间隔较大时, CI 方案和 DRD 方案的同化效果明显下降; 而 SRD 方案由于在误差源上考虑模式误差时还能够对特定源(如模式参数)进行误差订正, 因此其效果明显好于 CI 方案和 DRD 方案, 在观测稀疏时仍能得到较好的同化结果。

基于以上理想试验的结论, 我们认为 SRD 方案目前看来是一个相对有优势的模式误差扰动方案, 基于该方案的 EnKF 土壤湿度同化方法所得结果相对其他两种方案更优。但是我们也应该注意到, 由于只用了 AVIM 模式进行同化试验, 本文对三种方案比较所得的结果可能还与所用陆面模式和所选模式参数有一定关系。另外, 这里的结论主要都是来自理想试验的结果, 而实际土壤湿度资料的同化研究中会遇到许多来自各方面的复杂问题, 如观测间隔太大缺测较多、观测站点的稀少和分布的空间不均匀性、观测误差的确定、不同误差源误差统计特征和空间相关性的确定、模式参数订正问题等等, 这些都会影响 SRD 方案在实际应用中的效果, 并且也是实际土壤湿度同化应用中不可避免。本文只是 SRD 方案在土壤湿度同化中应用的一个初步探讨, 在今后的工作中, 我们将继续利用其他陆面模式来验证 SRD 方案的有效性和通用性, 并对该方法在实际土壤湿度同化中的性能和应用潜力, 以及对这些实际同化中会遇到的相关问题做进一步的探讨。

## 参考文献 (References)

- Anderson J L, Anderson S L. 1999. A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts [J]. *Mon. Wea. Rev.*, 127: 2741-2758.
- Crow W T. 2003. Correcting land surface model predictions for the impact of temporally sparse rainfall rate measurements using an ensemble Kalman filter and surface brightness temperature observations [J]. *J. Hydrometeor.*, 4: 960-973.
- Crow W T, Wood E F. 2003. The assimilation of remotely sensed soil brightness temperature imagery into a land surface model using ensemble Kalman filtering: A case study based on ESTAR measurements during SGP97 [J]. *Advance in Water Resources*, 26: 137-149.
- Crow W T, Van Loon E. 2006. Impact of incorrect model error as-

- sumptions on the sequential assimilation of remotely sensed surface soil moisture [J]. *J. Hydrometeor.*, 7: 421–432.
- Dan L, Ji J, Li Y. 2002. Climate simulations based on a different-grid nested and coupled model [J]. *Adv. Atmos. Sci.*, 19 (3): 487–499.
- 丹利. 2003. 大气和生物圈双向耦合模式中的多尺度变化模拟研究 [D]. 中国科学院大气物理研究所博士学位论文, 22–37.
- Dan Li. 2003. Multi-scale variation study on a two-way atmosphere and biosphere coupled model [D]. Ph. D. dissertation (in Chinese), Institute of Atmospheric Physics, Chinese Academy of Sciences, 22–37.
- Dan L, Ji J, Zhang P. 2005. The soil moisture of China in a high resolution climate-vegetation model [J]. *Adv. Atmos. Sci.*, 22 (5): 720–729.
- Dan L, Ji J, Liu H. 2008. Use of a land surface model to evaluate the observed soil moisture of grassland at the Tongyu reference site [J]. *Adv. Atmos. Sci.*, 25 (6): 1073–1084.
- Evensen G. 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics [J]. *J. Geophys. Res.*, 99 (C5): 10143–10162.
- Evensen G. 2003. The ensemble Kalman filter: Theoretical formulation and practical implementation [J]. *Ocean Dyn.*, 53: 343–367.
- Francois C, Quesney A, Ottlé C. 2003. Sequential assimilation of ERS-1 SAR data into a coupled land surface-hydrological model using an extended Kalman filter [J]. *J. Hydrometeor.*, 4: 473–487.
- Fujiyoshi Y, Ding Y H, Zhang Y. 2006. Outline of GAME/HUBEX [R]. Final Report of GAME/HUBEX. Fujiyoshi Y, Ding Y H, GAME/HUBEX Project Office, 1–6.
- Hamill T M, Whitaker J S, Snyder C. 2001. Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter [J]. *Mon. Wea. Rev.*, 129: 2776–2790.
- Houtekamer P L, Mitchell H L. 1998. Data assimilation using an ensemble Kalman filter technique [J]. *Mon. Wea. Rev.*, 126: 796–811.
- 黄春林, 李新. 2006a. 基于集合卡尔曼滤波的土壤水分同化试验 [J]. *高原气象*, 25 (4): 665–671. Huang Chunlin, Li Xin. 2006a. Experiments of soil moisture data assimilation system based on ensemble Kalman filter [J]. *Plateau Meteorology (in Chinese)*, 25 (4): 665–671.
- 黄春林, 李新. 2006b. 土壤水分同化系统的敏感性试验研究. *水科学进展* [J]. 17 (4): 457–465. Huang Chunlin, Li Xin. 2006b. Sensitivity analysis on land data assimilation scheme of soil moisture [J]. *Advances in Water Science (in Chinese)*, 17 (4): 457–465.
- Ji J. 1995. A climate-vegetation interaction model: Simulating physical and biological processes at the surface [J]. *Journal of Biogeography*, 22: 445–451.
- Kalnay E, Kanamitsu M, Kistler R, et al. 1996. The NCEP/NCAR 40-year reanalysis project [J]. *Bull. Amer. Meteor. Soc.*, 77: 437–471.
- Kalnay E. 2003. *Atmospheric Modeling, Data Assimilation and Predictability* [M]. Cambridge University Press, 341 pp.
- Li Y, Ji J. 2001. Model estimates of global carbon flux between vegetation and the atmosphere [J]. *Adv. Atmos. Sci.*, 18 (5): 807–818.
- Nie S, Luo Y, Zhu J. 2008. Trends and scales of observed soil moisture variations in China [J]. *Adv. Atmos. Sci.*, 25 (1): 43–58.
- Ni-Meister W, Houser P R, Walker J P. 2006. Soil moisture initialization for climate prediction: Assimilation of scanning multifrequency microwave radiometer soil moisture data into a land surface model [J]. *J. Geophys. Res.*, 111: D20102, doi: 10.1029/2006JD007190.
- Reichle R H, McLaughlin D B, Entekhabi D. 2002a. Hydrologic data assimilation with the ensemble Kalman filter [J]. *Mon. Wea. Rev.*, 130: 103–114.
- Reichle R H, Walker J P, Koster R D, et al. 2002b. Extended versus ensemble Kalman filtering for land data assimilation [J]. *J. Hydrometeor.*, 3: 728–740.
- Reichle R H, Koster R D. 2005. Global assimilation of satellite surface soil moisture retrievals into the NASA Catchment land surface model [J]. *Geophys. Res. Lett.*, 32: L02404, doi: 10.1029/2004GL021700.
- Shukla J, Mintz Y. 1982. Influence of land-surface evapotranspiration on the earth's climate [J]. *Science*, 215: 1498–1501.
- 田向军, 谢正辉. 2008. 考虑次网格变异性与土壤冻融过程的土壤湿度同化方案 [J]. *中国科学 (D辑)*, 38 (6): 741–749. Tian Xiangjun, Xie Zhenghui. 2008. A land surface soil moisture data assimilation framework in consideration of the model subgrid-scale heterogeneity and soil water thawing and freezing [J]. *Sci. China (Ser. D) (in Chinese)*, 38 (6): 741–749.
- van Leeuwen P J. 1999. Comment on “Data assimilation using an ensemble Kalman filter technique” [J]. *Mon. Wea. Rev.*, 127: 1374–1377.
- Whitaker J S, Compo G P, Wei X, et al. 2004. Reanalysis without radiosondes using ensemble data assimilation [J]. *Mon. Wea. Rev.*, 132: 1190–1200.
- Yang K, Koike T. 2008. Satellite monitoring of the surface water and energy budget in the central Tibetan Plateau [J]. *Adv. Atmos. Sci.*, 25 (6): 974–985.
- 张生雷, 谢正辉, 师春香, 等. 2008. 集合 Kalman 滤波在土壤湿度同化中的应用 [J]. *大气科学*, 32 (6): 1419–1430. Zhang Shenglei, Xie Zhenghui, Shi Chunxiang, et al. 2008. Applications of ensemble Kalman filter in soil moisture assimilation [J]. *Chinese Journal of Atmospheric Sciences (in Chinese)*, 32 (6): 1419–1430.
- Zhang S, Li H, Zhang W, et al. 2006. Estimating the soil moisture profile by assimilating near-surface observations with the ensemble Kalman filter (EnKF) [J]. *Adv. Atmos. Sci.*, 22 (6): 936–945.