

文章编号:0559-9350(2012)11-1302-10

## 基于集合卡尔曼滤波和HYDRUS-1D模型的土壤剖面含水量同化试验

王文, 刘永伟, 寇小华, 吕海深

(河海大学 水文水资源与水利工程科学国家重点实验室, 江苏 南京 210098)

**摘要:** 本文建立了一个基于集合卡尔曼滤波(EnKF)方法和HYDRUS-1D模型相结合的单点土壤水分同化方案, 以期通过同化表层土壤含水量的实测值来改进土壤剖面含水量预测精度。结果表明, 该同化方案能有效改进表层(4~10cm)及浅层(50cm)的土壤含水量预测精度, 对下层(100~150cm)的土壤含水量预测精度仍有影响, 但其影响有好有坏, 而对于深层( $\geq 200$ cm)的土壤含水量预测精度几乎没有任何影响; 对同化过程中的状态变量及关键状态参数的变化分析表明, 同化效果的产生是由于通过数据同化修正了模型的状态变量及关键参数, 而随着深度的增加, 变量和参数的调整量减小, 同化效果也就随之减弱。

**关键词:** 集合卡尔曼滤波; HYDRUS-1D; 数据同化; 土壤剖面含水量

**中图分类号:** S152.7

**文献标识码:** A

### 1 研究背景

土壤含水量是降雨径流过程模拟中的一个重要状态变量, 具有高度的时空变异性。土壤含水量地面实测只能采集有限地点的数据, 而且采集地表数据比采集深部数据要更容易。遥感可以实现大面积观测, 但仅能提供表层观测数据。因此, 如何根据地表土壤含水量推求土壤剖面含水量信息已成为水文研究的热点之一。

用表层土壤含水量估计土壤剖面含水量的方法大体有4种<sup>[1]</sup>: (1)表层观测值的统计外推法; (2)表层观测值与剖面水量平衡模型相结合的方法; (3)辐射传输模型的反演方法; (4)参数化的剖面模型方法。其中, 基于观测与模型模拟相结合的方法被认为是最有前途的, 而将二者结合最常用的手段是数据同化技术(包括卡尔曼滤波等), 很多研究者进行了相关尝试。例如, Walker等<sup>[2]</sup>利用卡尔曼滤波(KF)方法, 将表层土壤含水量观测数据同化到一个简化的土壤水分模型中, 进行一维土壤水分剖面反演的研究。Crosson等<sup>[3]</sup>也利用KF, 将地表通量-土壤水模型及辐射转换模型和微波遥感观测的与土壤含水量密切相关的亮度温度数据相结合, 实现三维的土壤含水量估计。Crow和Wood<sup>[4]</sup>应用集合卡尔曼滤波(EnKF)将微波遥感观测亮度温度同化到陆面模型中, 估计根层土壤水分。Jia等<sup>[5]</sup>利用EnKF, 将卫星微波遥感 AMSR-E 亮度温度数据同化到土壤水模型中, 结果表明不仅可改进表层土壤含水量估计, 也可以改进10~50cm深度土壤含水量估计。Tian等<sup>[6]</sup>考虑到模型次网格的非均质性和土壤水的冻融过程, 构建了一个双通道路面数据同化系统(DLDAS), 用于同化 AMSR-E 卫星亮度温度(Tb)数据, 对表层10cm的土壤水分估计有明显改进。

HYDRUS-1D模型<sup>[7]</sup>是非饱和流和溶质运移模拟应用最广的模型之一, 在土壤水分运移模拟中也有许多研究<sup>[8-9]</sup>。本文建立一个基于EnKF方法和HYDRUS-1D模型相结合的单点土壤水分同化方案,

收稿日期: 2012-01-06

基金项目: 国家自然科学基金项目(51190091; 40930635)

作者简介: 王文(1967-), 男, 江苏人, 教授, 博士, 主要从事环境变化对水文过程的影响及水文遥感研究。

E-mail: W.Wang@126.com

利用实测多层土壤含水量逐日时间序列数据进行同化试验，分析通过同化表层土壤含水量的实测值来改进土壤剖面含水量估计的效果。

## 2 基于 EnKF 和 HYDRUS-1D 的数据同化方案

**2.1 HYDRUS-1D 模型** HYDRUS-1D 模型主要是应用数值解法解决变饱和土壤中垂向一维水、热及溶质的运移问题。本试验用 HYDRUS-1D 模型模拟变饱和土壤中垂向上的水分运动，模型中采用改进的一维 Richards 方程(水分运动过程中假设气相不起重要作用，并忽略热量梯度的作用)，如式(1)：

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial x} \left[ K \left( \frac{\partial h}{\partial x} + \cos \alpha \right) \right] - S \quad (1)$$

式中： $h$ 为压力水头，cm； $\theta$ 为土壤体积含水量， $\text{cm}^3/\text{cm}^3$ ； $t$ 为时间，d； $x$ 为空间坐标，cm(上为正)； $S$ 为汇项， $\text{cm}^3 \cdot \text{cm}^{-3} \cdot \text{d}^{-1}$ ，表示作为根系吸水； $\alpha$ 为水流方向与垂直方向的夹角(即： $\alpha = 0^\circ$ 代表垂向流动， $90^\circ$ 代表水平流动， $0^\circ < \alpha < 90^\circ$ 代表倾斜流)； $K$ 为非饱和水力传导度，cm/d，如式(2)：

$$K(h, x) = K_s(x) K_r(h, x) \quad (2)$$

式中： $K_r$ 为相对水力传导度，cm/d； $K_s$ 为饱和水力传导度，cm/d。

模型中方程(1)采用伽辽金线状有限元法求解。本试验中，根据该试验站点垂向上的土壤质地分层和实测值分层在整个土壤剖面上设置了21个节点，即将空间在垂向上离散成20个相连的单元。根据该站点的实际情况，将上边界条件设为“Atmospheric”，下边界条件设为“Variable pressure head”。选用单孔隙模型中的 van Genuchten-Mualem(van Genuchten<sup>[10]</sup>)模型处理水分特征曲线，并忽略其滞后作用，模型中的水力特征参数根据土壤质地并结合模型自带的神经网络估计功能确定(详见3.1)。另外，本试验选用 Feddes 模型(Feddes<sup>[11]</sup>)考虑根系吸水作用。

**2.2 集合卡尔曼滤波(EnKF)** EnKF 是 Evensen 等<sup>[12]</sup>在 1994 年提出的顺序数据同化算法。它将模型状态预报看成近似随机动态预报，用一个状态总体(设数目为  $N$ ，即集合数)去代表随机动态预报中的概率密度函数，通过向前积分计算下一时刻状态总体的概率密度函数，并得到该时刻的统计特性(如均值与协方差)。本文基于 EnKF 方法的数据同化步骤如下：

(1)初始化。给定  $N$  个符合高斯分布的随机变量(本试验中为土壤剖面上的体积含水量  $\theta$ ，属于状态变量)  $X_{i,0} = (i=1, 2, \dots, N)$

(2)计算每个随机变量在  $k+1$  时刻的预报值  $X_{i,k+1}^f$

$$X_{i,k+1}^f = f_k(X_{i,k}^a) + w_{i,k}; w_{i,k} \sim N(0, Q_k) \quad (3)$$

式中： $X_{i,k+1}^f$ 为随机变量  $i$  在  $k+1$  时刻的预报值； $f_k(\cdot)$ 为非线性模型算子，本文为 HYDRUS-1D 模型， $Q_k$ 为模型误差协方差矩阵； $w_{i,k}$ 为模型误差扰动，一般取高斯白噪声(均值为 0，方差为  $Q_k$  对角线上的元素，即标准差为模型误差标准差)。

(3)计算  $k+1$  时刻的滤波增益矩阵  $K_{k+1}$

$$K_{k+1} = P_{k+1}^f H^T (H P_{k+1}^f H^T + R_{k+1})^{-1} \quad (4)$$

$$P_{k+1}^f = \frac{1}{N-1} \sum_{i=1}^N (X_{i,k+1}^f - \overline{X_{k+1}^f}) \cdot (X_{i,k+1}^f - \overline{X_{k+1}^f})^T \quad (5)$$

$$P_{k+1}^f H^T = \frac{1}{N-1} \sum_{i=1}^N (X_{i,k+1}^f - \overline{X_{k+1}^f}) \cdot \left[ h(X_{i,k+1}^f) - h(\overline{X_{k+1}^f}) \right]^T \quad (6)$$

$$H P_{k+1}^f H^T = \frac{1}{N-1} \sum_{i=1}^N \left[ h(X_{i,k+1}^f) - h(\overline{X_{k+1}^f}) \right] \cdot \left[ h(X_{i,k+1}^f) - h(\overline{X_{k+1}^f}) \right]^T \quad (7)$$

$$\overline{X_{k+1}^f} = \frac{1}{N} \sum_{i=1}^N X_{i,k+1}^f \quad (8)$$

式中： $\overline{X_{k+1}^f}$ 为 $k+1$ 时刻状态变量预报集合的平均值； $R_{k+1}$ 为观测误差协方差矩阵， $h(\cdot)$ 为观测算子，矩阵 $H$ 为其线性表达式。当 $h(\cdot)$ 为高度非线性函数时，矩阵 $H$ 很难显式表示，此时可以把 $P_{k+1}^f H$ 、 $HP_{k+1}^f H^T$ 作为一个整体由式(6)、式(7)计算得来。

(4)计算 $k+1$ 时刻状态变量的分析集合均值 $\overline{X_{k+1}^a}$ 和分析集合协方差 $P_{k+1}^a$

$$X_{i,k+1}^a = X_{i,k+1}^f + K_{k+1} \left[ Y_{k+1} - h(X_{i,k+1}^f) \right] \quad v_{i,k} \sim N(0, R_k) \quad (9)$$

$$\overline{X_{k+1}^a} = \frac{1}{N} \sum_{i=1}^N X_{i,k+1}^a \quad (10)$$

$$P_{k+1}^a = \frac{1}{N-1} \sum_{i=1}^N (X_{i,k+1}^a - \overline{X_{k+1}^a}) \cdot (X_{i,k+1}^a - \overline{X_{k+1}^a})^T \quad (11)$$

式中： $X_{i,k+1}^a$ 为第 $i$ 个随机变量在 $k+1$ 时刻的分析值； $\overline{X_{k+1}^a}$ 为 $k+1$ 时刻的分析集合均值； $Y_{k+1}$ 为 $k+1$ 时刻的观测集合(由实测值加上观测误差生成)； $R_k$ 为观测误差协方差矩阵； $v_{i,k}$ 为观测误差扰动，取高斯白噪声(均值为0，方差为 $R_k$ 对角线上的元素，即标准差为观测误差标准差)。

(5)分析集合均值 $\overline{X_{k+1}^a}$ 成为 $k+1$ 时刻的状态估计量(即土壤剖面含水量)，分析集合协方差 $P_{k+1}^a$ 成为 $k+1$ 时刻的状态误差，根据 $\overline{X_{k+1}^a}$ 和 $P_{k+1}^a$ 生成 $k+1$ 时刻的状态估计集合(即土壤剖面含水量状态集合) $X_{k+1}^{a'}$ ， $X_{k+1}^{a'} \sim N(\overline{X_{k+1}^a}, P_{k+1}^a)$ 。返回步骤(2)。

在以上EnKF的计算步骤中，步骤(2)属于预报部分，方程(3)为状态转移方程；步骤(2)、(4)属于分析部分，方程(9)为状态更新方程。

**2.3 基于EnKF的土壤剖面水分同化方案** 结合2.1与2.2节内容，本文设计了EnKF与HYDRUS-1D相结合的同化方案，其流程如图1所示。同时，在图2中绘出了EnKF与HYDRUS-1D相结合的主要数据流程，以便对比分析土壤剖面上水分的实测值、HYDRUS-1D预报值、同化预报值及分析值之间的差异，分析数据同化的效果。其中， $X_k^f$ 为 $k-1$ 时刻不加观测的条件下由HYDRUS-1D模型模拟得到的 $k$ 时刻土壤剖面含水量，而本文以下HYDRUS-1D预报值是指从模拟开始到结束不加任何观测只用HYDRUS-1D模拟得到的预报值； $X_{k+1}^f$ 为 $k+1$ 时刻的同化预报集合，其集合均值 $\overline{X_{k+1}^f}$ 为同化预报值； $X_{k+1}^a$ 为 $k+1$ 时刻的分析集合，其集合均值 $\overline{X_{k+1}^a}$ 为分析值(同化预报集合与分析集合详述见下文)。

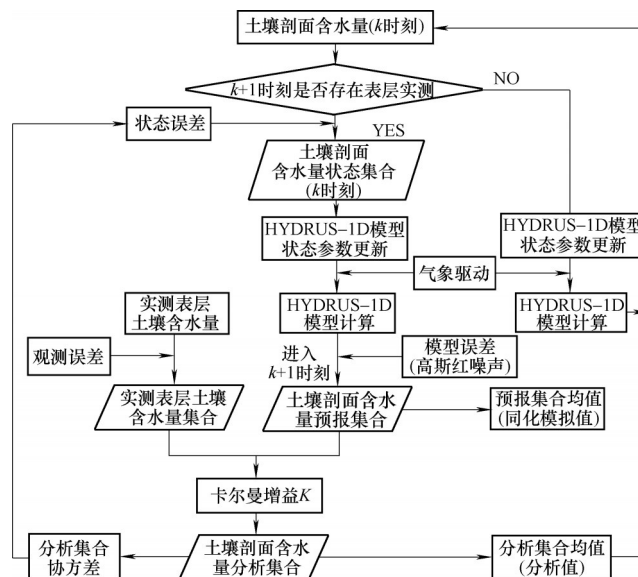


图1 一维土壤水分同化方案流程

根据图 1 和图 2，可以看出本同化方案的处理过程大致如下：

$k$  时刻的土壤剖面含水量有两种情况：(1) 由  $k-1$  时刻只经模型模拟得到的预报值  $X_k^f$ ；(1) 由  $k-1$  时刻经模型模拟并加入  $k$  时刻的表层实测值同化得到的分析集合均值  $\overline{X_k^a}$ 。判断当  $k+1$  时刻不存在表层实测时，直接将  $k$  时刻的土壤剖面含水量 ( $X_k^f$  或  $\overline{X_k^a}$ ) 带入 HYDRUS-1D 模型计算，得到  $k+1$  时刻土壤剖面含水量  $X_{k+1}^f$ ；当  $k+1$  时刻存在表层实测时，由  $k$  时刻的土壤剖面含水量 ( $X_k^f$  或  $\overline{X_k^a}$ ) 加上  $k$  时刻的状态误差 (情况①状态误差根据模型模拟误差和观测误差评经验设定；情况②状态误差标准差取  $k$  时刻分析集合协方差  $P_k^a$  对角线上元素的平方根) 生成  $k$  时刻服从正太分布的土壤剖面含水量状态估计集合  $X_k^{a'}$ 。再根据  $X_k^{a'}$  更新 HYDRUS-1D 模型的状态参数和边界条件，将  $X_k^{a'}$  带入模型计算 (并加入模型误差) 得到  $k+1$  时刻的同化预报集合  $X_{k+1}^f$  (2.2 节中步骤 (2))。对  $k+1$  时刻的表层实测值加上观测误差生成表层观测集合  $Y_{k+1}$ ，然后，根据  $Y_{k+1}$  和  $X_{k+1}^f$  通过步 (3) 计算卡尔曼增益  $K_{k+1}$ ，根据步骤 (4) 得到  $k+1$  时刻的分析集合  $X_{k+1}^a$ 、集合均值  $\overline{X_{k+1}^a}$  (即分析值)、集合协方差  $P_{k+1}^a$ ，并取  $\overline{X_{k+1}^a}$  作为  $k+1$  时刻的土壤剖面含水量。继续判断  $k+2$  时刻是否存在表层实测，由  $k+1$  时刻的土壤剖面含水量 ( $X_{k+1}^f$  或  $\overline{X_{k+1}^a}$ ) 继续同化预报过程。

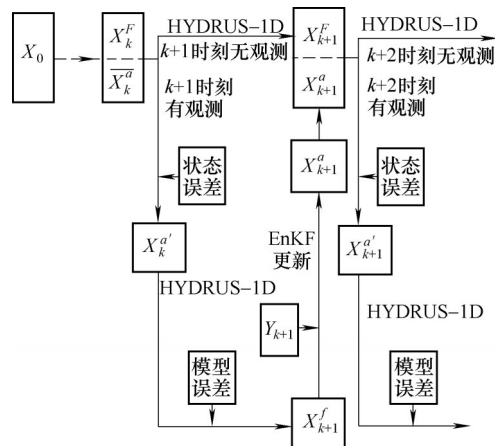


图 2 EnKF 与 HYDRUS-1D 相结合的主要数据流程

### 3 同化试验及结果分析

**3.1 试验数据** 中科院南皮生态农业试验站点 (以下简称南皮站) 位于河北省南皮县 (北纬  $38^{\circ}06'$ ，东经  $116^{\circ}40'$ )，属季风气候，年降水量 400~550mm，地下水埋深 5~7m。本试验选取该站点 2010 年 4 月 29 日—6 月 25 日的观测数据进行了同化试验 (数据由河海大学缴锡云教授提供)，并利用 2010 年 4 月 28 日土壤剖面含水量实测值作为初始状态变量。整个剖面深度为 500cm，实测值为 4~10、50、100、150、200、250、300、350、400 和 450cm 深度处的日平均土壤体积含水量，由 TDR 采集。

HYDRUS-1D 模型所需的驱动数据包括最高温度、最低温度、湿度、日照时数、太阳辐射、净辐射、降雨量及风速，参数集包括土壤、植被等参数。土壤水分特性根据实测数据滤定模型参数，本试验选择的 van Genuchten-Mualem 模型中需要确定的水力特征参数包括残余土壤含水量  $Q_r$ ，饱和土壤含水量  $Q_s$ ，土壤函数中的参数  $\alpha$ ，土壤持水函数中的参数  $n$ ，饱和水力传导度  $K_s$ ，根据 0~300cm 土壤质地和模型自带的神经网络估计功能确定，见表 1。另外，植被根系参照王颖聪<sup>[13]</sup>的设置方式，根据植株高度关系确定根系深度，最大根系深度取 60cm。

表 1 不同深度土壤水力参数值

土层/cm	质地	$Q_r/(cm^3/cm^3)$	$Q_s/(cm^3/cm^3)$	$\alpha/cm^{-1}$	$n$	$K_s/(cm/d)$
0~170	粉壤	0.1	0.35	0.01	1.617 7	13.26
170~230	粉壤	0.022 6	0.30	0.005 2	1.767 4	35.0
230~275	粉壤	0.034 1	0.29	0.005 9	1.453 9	9.23
275~315	粉黏土	0.080 9	0.45	0.017 4	1.520 3	1.23
315~350	粉壤	0.035 2	0.49	0.005 0	1.974 9	8.06
350~450	砂壤	0.018 1	0.47	0.031 6	1.455 5	71.99
450~500	砂土	0.043 3	0.43	0.039 4	2.077 7	179.22

**3.2 同化结果分析** 本文只同化表层(4~10cm)土壤含水量的实测值,并只在表层加误差(包括初始状态误差、模型误差、观测误差),其它各层误差均设为0。本试验中初始状态变量 $X_0 = [\theta_1, \theta_2, \dots, \theta_n]$  ( $n=10$ )为2010年4月28日实测土壤剖面上4~10、50、100、150、200、250、300、350、400和450cm处的日平均土壤体积含水量。用初始土壤剖面含水量加上高斯白噪声(均值为0,标准差凭经验设为0.01)生成初始土壤剖面含水量的集合 $X_{i,0} = [\theta_1, \theta_2, \dots, \theta_n]$  ( $i=1, 2, \dots, N$ )。

随集合数的增加,集合成员的代表性增强、分析误差减小,但计算负担加重,适度的集合数即能满足集合成员的代表性。例如Houtekamer<sup>[14]</sup>中认为100个集合成员足以精确地描述其局地非均质性。Mitchell等<sup>[15]</sup>的同化试验中应用64个集合成员取得了很好的效果。本试验为消除局地化的影响并综合考虑计算量问题取集合数 $N$ 为100。

模型误差的确定参考Evensen<sup>[16]</sup>,根据式(12)、式(13)计算。

$$\begin{bmatrix} q_k \\ X_k \end{bmatrix} = \begin{bmatrix} \alpha q_{k-1} \\ f(X_{k-1}) + \sqrt{\Delta t} \sigma \rho q_k \end{bmatrix} + \begin{bmatrix} \sqrt{1-\alpha^2} \varphi_{k-1} \\ 0 \end{bmatrix} \quad (12)$$

$$\rho^2 = \frac{1}{\Delta t} \frac{(1-\alpha)^2}{n - 2\alpha - n\alpha^2 + 2\alpha^{n+1}} \quad (13)$$

式中: $q_k$ 为模型误差项; $\alpha$ 为相邻两个模型误差项的相关系数且 $\alpha \in [0, 1)$ ; $\varphi_k$ 为随机变量,且 $\varphi_k \sim N(0, 1)$ ; $\Delta t$ 为同化时间步长; $n$ 为从初始状态到模拟状态的时间步长数; $\sigma$ 为模型误差标准差。本试验中 $\Delta t=1$ , $n=1$ ,取 $q_0 \sim N(0, 1)$ ,根据HYDRUS-1D预报值相对于实测值的均方根误差(RMSE)确定 $\sigma=0.026$ 。则根据公式(13)可得 $\rho^2=1$ ,因参数 $\alpha$ 很难确定,本试验中分别取 $\alpha=0.05, 0.25, 0.5, 0.75, 0.95$ 时,发现 $\alpha$ 的取值对同化结果影响不大,最终取 $\alpha=0.5$ 。本试验中,观测误差取均值为0,标准差为0.01的高斯白噪声。另外,因状态变量为土壤剖面上各层的体积含水量,而观测只为表层实测值,故观测算子 $h(\cdot)$ 取线性算子, $H_{1 \times 7}=(1, 0, 0, 0, 0, 0, 0, 0, 0, 0)$ 。

根据以上设计,南皮站4~10、50、100、150、200、250和300cm深度处的土壤水分同化结果如图3(350~450cm处同化效果不显著,为节省篇幅未列出)。图中的分析值、同化预报值、HYDRUS-1D预报值、实测值的含义如2.3所述。

本文采用均方根误差(RMSE)和相对误差(RE)比较预测结果与实测值之间的拟合程度。RMSE和RE分别表达为:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (P-O)^2}{n}} \quad (14)$$

$$RE = \frac{\sum_{i=1}^n (P-O)}{\sum_{i=1}^n O} \quad (15)$$

式中: $P$ 为土壤含水量的预测值; $O$ 为土壤含水量的实测值; $n$ 为观测数。误差统计结果如表2。

从图3与表2可以得出如下结论。

(1)在表层(4~10cm),HYDRUS-1D预报值与实测值存在一定差距,有降水时,二者差距更大。分析值明显优于同化预报值和HYDRUS-1D预报值,同化预报值又优于HYDRUS-1D预报值,尤其是在有降水的情况下,同化效果更加明显。从表2同样可以看出,分析结果较同化预报结果、同化预报结果较HYDRUS-1D预报结果RE、RMSE的绝对值都变小了。

(2)50~300cm深度范围的分析值与同化预报值基本重合,分析值不一定优于同化预报值。

(3)在50cm深度处,同化预报值总体上比HYDRUS-1D预报值更接近实测值,有降水时更明显,说明同化表层实测值对50cm深度土壤含水量模拟预测效果较明显。同样,从表2也可以看出,50cm深度处,分析结果略优于同化预报结果,同化预报结果略优于HYDRUS-1D预报结果。

(4)在100cm、150cm深度处,有降水时,同化预报值比HYDRUS-1D预报值偏大,且都大于实测值,此时该两层的同化效果为负;无降水时,同化预报值与HYDRUS-1D预报值基本重合,同化效果较小。由表2从总体上看,同化表层实测值,对该两层的效果为负。

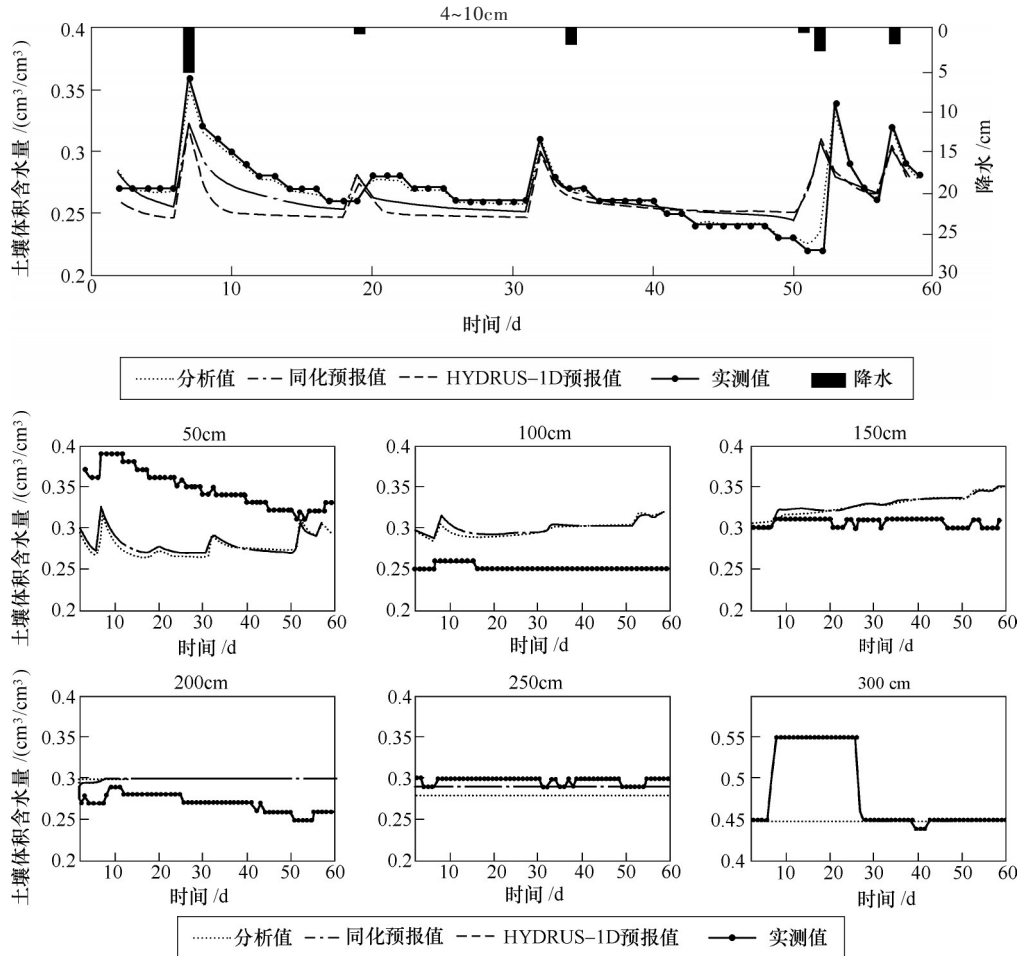


图3 不同深度土壤含水量的分析值、同化预报值、HYDRUS-1D预报值、实测值比较

表2 不同深度土壤含水量分析值、同化预报值、HYDRUS-1D预报值误差统计

土壤深度/cm	分析值		同化预报值		HYDRUS-1D预报值	
	<i>RE</i>	<i>RMSE</i>	<i>RE</i>	<i>RMSE</i>	<i>RE</i>	<i>RMSE</i>
4~10	-0.003	0.002 7	-0.017 5	0.020 6	-0.037 0	0.025 6
50	-0.194 7	0.071 7	-0.194 8	0.071 8	-0.202 4	0.075 8
100	0.191 2	0.048 8	0.191 1	0.048 8	0.184 3	0.047 6
150	0.068 1	0.024 0	0.068 1	0.024 0	0.064 3	0.023 8
200	0.108 4	0.031 0	0.108 4	0.031 0	0.107 8	0.030 9
250	-0.025 5	0.008 7	-0.025 5	0.008 7	-0.059 1	0.018 1
300	-0.069 5	0.058 0	-0.069 5	0.058 0	-0.069 5	0.058 0

(5)在200cm以上深度处(200、250、300cm),分析值、同化预报值、HYDRUS-1D预报值已基本或十分接近。可以认为,同化表层实测值对于深层( $\geq 200\text{cm}$ )的土壤含水量模型预测无影响。

总之,同化表层(4~10cm)土壤含水量的实测值能有效改进表层(4~10cm)及浅层(50cm)土壤含水量预测精度;对深度100~150cm的土壤含水量预测精度仍有影响,但其影响有好有坏;而在深度 $\geq 200\text{cm}$ 的部位,分析值、同化预报值、HYDRUS-1D预报值基本重合,同化表层土壤含水量对该深度土壤含水量预测精度几乎没有任何影响。

**3.3 同化效果产生的原因分析** 为说明同化效果产生的原因,下面对本试验中同化与未同化条件下HYDRUS-1D模型公式(1)、式(2)中的主要状态参数(非饱和水力传导度 $K$ 、水头 $h$ )及状态变量(土壤体积含水量 $\theta$ )的变化情况进行对比分析。这里,同化是指加入表层实测更新状态变量,再利用HY-

DRUS-1D模型预报的过程；未同化是指用HYDRUS-1D模型预报，而未加入表层实测值的过程。我们采用以下几个指标来反映这几个量在同化过程中的调整量：(1) $K$ 的相对差值。 $K$ 相对差值= $(K_{\text{同化}} - K_{\text{未同化}}) / K_{\text{未同化}} \times 100\%$ ；(2)水头 $h$ 的差值。 $h$ 差值= $h_{\text{同化}} - h_{\text{未同化}}$ ；(3)土壤体积含水量 $\theta$ 的差值。 $\theta$ 差值=分析值-未同化预报值；(4) $\theta$ 在同化、未同化条件下的预报误差。 $\theta$ 同化预报误差=同化预报值-实测值； $\theta$ 未同化预报误差=未同化预报值-实测值。以上未同化预报值，即HYDRUS-1D预报值，对比结果如图4所示。

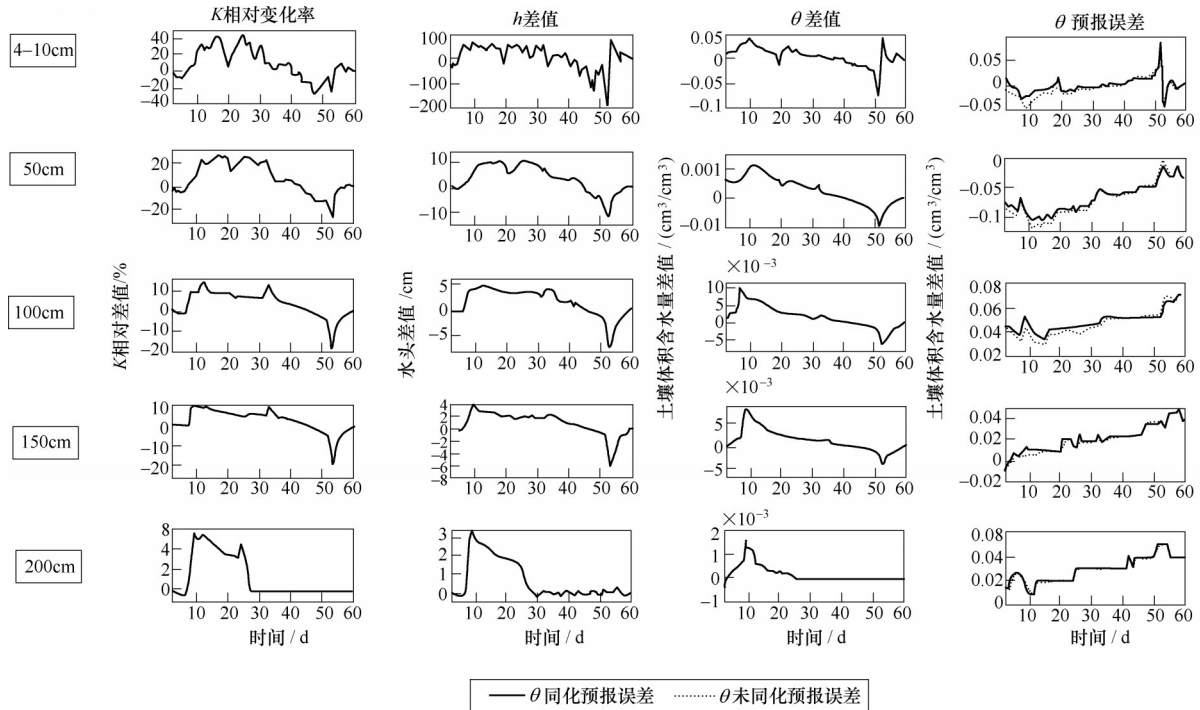


图4 土壤剖面(0~200cm)非饱和水力传导度 $K$ 、水头 $h$ 、土壤体积含水量 $\theta$ 在同化与未同化过程中的差值变化以及同化与未同化条件下 $\theta$ 的预报误差变化情况

由图4从表层(0~4cm)到深层200cm，同化与未同化条件下非饱和水力传导度 $K$ 相对差值的变化范围逐渐减小，表示土壤剖面上随深度增加同化过程中参数 $K$ 的调整量逐渐减小；同样，同化过程中随深度增加水头 $h$ 、土壤体积含水量 $\theta$ 的调整量也逐渐减小。由图4中的各土壤深度来看，每一层上状态参数 $K$ 、 $h$ 和状态变量 $\theta$ 在同化过程中的调整量随时间变化大体一致。另外，在表层(0~4cm)和浅层50cm， $\theta$ 同化预报误差小于未同化预报误差，同化有效果；在100cm、150cm， $\theta$ 同化预报误差大于未同化预报误差，同化效果为负；在200cm， $\theta$ 同化预报误差基本等于未同化预报误差，同化已无效果。这与3.3节的同化结果相符合。

由此可以看出：同化效果的产生在于在同化表层土壤含水量的实测值得到分析集合后，更新了模型状态，同时更新了HYDRUS-1D模型中的状态参数，即对模型的状态变量和状态参数都进行了调整，从而使得同化过程中模型预报误差减小。但随深度增加模型状态变量和状态参数参数的调整量逐渐减小，导致整个土壤剖面上的同化效果随深度逐渐减弱，到200cm深度处，模型状态变量和状态参数的调整量几乎为0，同化对预报结果已无影响。

## 4 讨论

### 4.1 EnKF数据同化流程

目前，虽然EnKF已广为使用，但并无统一的实现流程，在具体计算步骤、集合数目、模型误差的设定等方面均有所差异。较为典型的流程是Houtekamer<sup>[14]</sup>提出的EnKF同化方案，其中设计了两个成员数目均为 $N$ 的集合同时进行并行计算，同化过程中，两个集合分别向

对方提供预报误差协方差  $P_k^f$ ，用于计算对方的卡尔曼增益  $K_k$  (通过式(4))，再分别利用自己的预报集合  $X_k^f$  通过状态更新方程(9)计算分析集合  $X_k^a$ 。这样避免了只用一个成员数目为  $N$  的集合预报时，由于只有一个预报集合，状态更新方程中决定分析集合的卡尔曼增益和同化预报值完全由同化预报集合决定，导致同化过程中分析集合的不确定性被低估，即分析集合分布过于集中，集合均值无法代表系统真值。同样，为避免以上情况，不同于 Houtekamer 的做法，本方案在流程上进行以下设计：选择一个成员数目为  $N$  的集合，但在由  $k$  时刻加入观测集合进行 EnKF 更新生成分析集合之后，根据分析集合均值和协方差生成了符合高斯分布的  $k$  时刻土壤剖面含水量状态集合。(2.2 中步骤(5))。避免了同化过程中分析集合分布集中代表性降低的现象，同时减弱了同化过程中误差的积累。

集合数越多，集合成员的代表性越好，但是随着集合数的增多计算量也增大。本同化方案中取集合数 100，保证不会由于集合数不足导致集合成员代表性不够，甚至滤波发散的情况<sup>[14]</sup>。模型误差的设定是 EnKF 的关键，对同化方案的成败有很大的影响<sup>[15]</sup>。目前，较为简单的是将模型误差设为高斯白噪声，而本方案模型误差的设定参考 Evensen<sup>[16]</sup>，在状态转移方程中加入误差扰动为高斯红噪声的模型误差，使得模型误差与模型状态变量相关，同化过程中更新模型状态变量的同时也更新了模型误差，这样既解决了未加入模型误差时出现的滤波发散问题(即集合卡尔曼滤波同化循环中分析误差协方差很快收敛为 0，最终完全排斥观测资料，同化不再起作用)，又使之更符合实际。

**4.2 同化结果的比较分析及其与实时校正的关系** EnKF 方法数据同化的作用可以通过比较同化与未经同化的土壤含水量的模拟或预测值的差异来说明。已发表文献中(如文献[4、17-19]等)的比较分析主要是针对 EnKF 分析值、模型模拟值和实测值三者的比较，一般把实测值近似看做真值，若同化结果比模型模拟结果更接近实测值，则说明同化有效果。但是，基于 EnKF 方法同化表层实测值来估计土壤剖面含水量时，加入实测之后的表层分析值一定比模拟值更接近实测值。实际应用(如水文预报)中，在当前时刻无法得到下一时刻的实测值，我们所关注的是当前时刻的同化过程能否改进下一时刻的预报。于是本研究的分析比较对象是加入实测后的同化预报结果与 HYDRUS-1D 预报结果：若同化预报结果比 HYDRUS-1D 预报结果更接近实测值，则说明当前时刻的同化过程对于改进下一时刻的预报起作用，即同化有效果。反之，同化无效果。

数据同化,从狭义而言,指在数值模型的动态运行过程中,将模型计算结果与观测数据相融合,通过不断更新系统状态变量与模型参数以不断调整模型轨迹,提高物理过程模拟或预报精度的方法<sup>[20]</sup>;从广义而言,也包括利用多源信息对模型输入变量进行更新以及对输出结果进行校正。也就是说,广义上,水文预报中常用到的实时校正技术(即根据某一时刻的预报值及相应的实测值,求出该预报值的误差,并利用这个误差调整预报参数,提高下一预报时段预报精度的方法)也是数据同化的一部分。本文的研究是通过更新系统状态变量与模型参数来实现以预测精度的提高的,因而属于狭义的数据同化。

**4.3 由表层土壤含水量推求土壤剖面含水量的有效深度** 由表层土壤含水量推求土壤剖面含水量的有效深度(以下称同化有效深度)受许多因素影响。现有方法很多,推求精度也依不同情形而有较大差异。

经验公式法,即利用先验的土壤剖面水分变化特点结合表层观测预测土壤剖面含水量的方法,较为简单易用。鹿洁忠<sup>[21]</sup>利用表层以下各层土壤水分与表层(0~5cm)土壤水分的关系建立经验公式,得到了 0~150cm 以 10cm 为间隔层次的任一厚度的水分储量,其计算的平均误差 <5%,最大误差为 11%。Ragab<sup>[22]</sup>建立了一个简单的回归方程,由表层(0~5cm 和 0~10cm)土壤含水量得到了根层(0~50cm)土壤水分储量。

利用数据同化技术将表层观测数据与各种水文模型、土壤水热传输模型相结合来推求土壤剖面含水量是近年来的研究热点。Heathman 等<sup>[23]</sup>的研究表明用直接插值法同化表层(0~5cm)土壤含水量来估算土壤剖面(0~60cm)含水量时,只有(0~30cm)的土壤剖面水分同化有效果。由于卡尔曼滤波(KF)方法可以利用不同层的土壤水分的相关性来修正观测深度以下的土壤含水量,其成为目前获取土壤剖面含水量最适当的方法。Galantowicz 等<sup>[24]</sup>基于 EKF 方法同化 L-波段射电亮度数据得到了(0~



100cm)的土壤水分剖面,基于不同的土壤类型反演误差为2.2%~8%。Crow和Wood<sup>[4]</sup>通过同化微波遥感地表(5cm)亮度温度,可以获得40cm深度土壤水的时空分布信息。Jia等<sup>[5]</sup>在同化AMSR-E亮度温度进行根部土含水量估算时,对10~50cm深度的土壤含水量的RMSE可以减少20.9%。Li和Islam<sup>[25]</sup>应用四层陆面模型通过顺序同化表层土壤含水量来估计土壤剖面含水量,结果表明直到第三层72cm处同化有效果,第四层189cm处同化基本无影响。

同化有效深度受多方面因素(包括观测深度、土壤质地、表层观测与模拟的偏离程度及模型算子等)的影响。根据本文研究结果与他人已有成果可以看出,从表层推求深层土壤含水量的有效深度一般为50~100cm。

## 5 结论

本文建立了基于EnKF方法和HYDRUS-1D模型相结合的单点数据同化方案,并利用中科院南皮生态农业试验站点进行了通过同化表层(4~10cm)土壤含水量,改进HYDRUS-1D模型对剖面土壤含水量预测精度的试验研究。研究表明:(1)对于表层土壤水分的估计效果最好,同化预报值表层土壤含水量能有效改善HYDRUS-1D预报精度,尤其是在有降水时,效果更显著;(2)同化表层土壤含水量也能有效改进50cm深度土壤含水量的预测,尤其是当表层HYDRUS-1D预报值与实测值偏差较大(本试验中为有降水时)时;对100cm及150cm深度的土壤含水量预测有一定影响,但作用有正有负;同化表层土壤含水量对改进深层土壤含水量的影响随深度的增加而减小,对于深度 $\geq 200$ cm处已几乎无影响;(3)对同化过程中的状态变量及关键状态参数的变化分析表明,同化效果的产生是由于通过数据同化修正了模型的状态变量及关键参数,而随着深度的增加,变量和参数的调整量减小,同化效果也就随之减弱。

总之,利用数据同化技术在获取土壤剖面含水量信息方面前景广阔,还有待于进一步研究。

## 参 考 文 献:

- [ 1 ] Jackson T J . III . Measuring surface soil moisture using passive microwave remote sensing[J] . Hydrological Processes, 1993, 7(2): 139-152 .
- [ 2 ] Walker J P, Willgoose G R, Kalma J D . One-dimensional soil moisture profile retrieval by assimilation of near-surface observations: A comparison of retrieval algorithms[J] . Advances in Water Resources, 2001, 24(6): 631-650 .
- [ 3 ] Crosson W L, Laymon C A, Inguva R, et al . Assimilating remote sensing data in a surface flux-soil moisture model[J] . Hydrological Processes, 2002, 16(8): 1645-1662 .
- [ 4 ] Crow W T, Wood E F . 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] . Advances in Water Resources, 2003, 26(2): 137-149 .
- [ 5 ] Jia B H, Xie Z H, Tian X J, et al . A soil moisture assimilation scheme based on the ensemble Kalman filter using microwave brightness temperature[J] . Science in China Series D: Earth Sciences, 2009, 52(11): 1835-1848 .
- [ 6 ] Tian X, Xie Z, Dai A, et al . A microwave land data assimilation system: Scheme and preliminary evaluation over China[J] . J. Geophys Res, 2010(115): D21113 .
- [ 7 ] USDA-ARS U.S . Salinity Laboratory . HYDRUS-1D Model [EB/OL] . [Http://www.ars.usda.gov/services/docs/html?docid=8921](http://www.ars.usda.gov/services/docs/html?docid=8921) .
- [ 8 ] Šimunek J, Jarvis N J, M Th van Genuchten, et al . Review and comparison of models for describing non-equilibrium and preferential flow and transport in the vadose zone[J] . Journal of Hydrology, 2003, 272: 14-35 .
- [ 9 ] Šimunek J, van Genuchten, M Th, Sejna M . Hydrus: model use, calibration, and validation[J] . Transactions of the ASABE, 2012, 55(4): 1261-1274 .
- [ 10 ] van Genuchten M T . A Closed-form Equation for Predicting the Hydraulic Conductivity of Unsaturated Soils[J] .

- Soil . Sci . Soc . Am . , 1980(44) : 892–898 .
- [ 11 ] Feddes R A , Kowalik P J , Zaradny H . Simulation of Field Water Use and Crop Yield[M] . John Wiley & Sons , New York , NY , 1978 .
- [ 12 ] Evensen G . Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics[J] . Journal of Geophysical Research–All Series– , 1994(99) : 10143–10162 .
- [ 13 ] 王颖聪 . 灌溉施肥对土壤水分溶质运移的影响[D] . 南京 : 河海大学 , 2010 .
- [ 14 ] Houtekamer P L , Mitchell H L . Data assimilation using an ensemble Kalman filter technique[J] . Monthly Weather Review , 1998 , 126(3) : 796–811 .
- [ 15 ] Mitchell H L , Houtekamer P , Pellerin G . Ensemble Size , Balance , and Model–Error Representation in an Ensemble Kalman Filter\*[J] . Monthly Weather Review , 2002 , 130(11) : 2791–2808 .
- [ 16 ] Evensen G . The ensemble Kalman filter : Theoretical formulation and practical implementation[J] . Ocean dynamics , 2003 , 53(4) : 343–367 .
- [ 17 ] Dunne S , Entekhabi D . An ensemble-based reanalysis approach to land data assimilation[J] . Water Resour Res , 2005 , 41(2) : W02013 .
- [ 18 ] 张生雷 , 谢正辉 , 师春香 . 集合 Kalman 滤波在土壤湿度同化中的应用[J] . 大气科学 , 2008 , 32(006) : 1419–1430 .
- [ 19 ] 黄春林 , 李新 . 基于集合卡尔曼滤波的土壤水分同化试验[J] . 高原气象 , 2006 , 25(4) : 665–671 .
- [ 20 ] 王文 , 寇小华 . 水文数据同化方法及遥感数据在水文数据同化中的应用进展[J] . 河海大学学报 : 自然科学版 , 2009 , 37(005) : 556–562 .
- [ 21 ] 鹿洁忠 . 根据表层数据估算深层土壤水分[J] . 中国农业气象 , 1987(3) : 60–62 .
- [ 22 ] Ragab R . Towards a continuous operational system to estimate the root-zone soil moisture from intermittent remotely sensed surface moisture[J] . Journal of Hydrology , 1995 , 173(1–4) : 1–25 .
- [ 23 ] Heathman G C , Starks P J , Ahuja L R , et al . Assimilation of surface soil moisture to estimate profile soil water content[J] . Journal of Hydrology , 2003 , 279(1–4) : 1–17 .
- [ 24 ] Galantowicz J F , Entekhabi D , Njoku E G . Tests of sequential data assimilation for retrieving profile soil moisture and temperature from observed L-band radiobrightness[J] . Geoscience and Remote Sensing , IEEE Transactions on , 1999 , 37(4) : 1860–1870 .
- [ 25 ] Li J , Islam S . On the estimation of soil moisture profile and surface fluxes partitioning from sequential assimilation of surface layer soil moisture[J] . Journal of Hydrology , 1999 , 220(1–2) : 86–103 .

### **EnKF and HYDRUS-1D based data assimilation experiments for improving soil moisture profile prediction**

WANG Wen , LIU Yong-wei , KOU Xiao-hua , LU Hai-shen

(State Key Laboratory of Hydrology–Water Resources and Hydraulic Engineering , Hohai University , Nanjing 210098 , China)

**Abstract:** A one-dimensional soil moisture assimilation scheme at point scale based on the combination of the ensemble Kalman filter (EnKF) and HYDRUS-1D model is developed to improve the soil moisture profile prediction accuracy by assimilating the surface soil moisture. The results show that the assimilation scheme has a remarkable effect on the surface layer of 4~10cm and the shallow layer of 50cm. For the layers at depth 100~150cm, it still has certain influence, but the influence may be negative for the layer in the depth more than 200cm, which assimilation has no effect on the soil moisture prediction. The analysis based on changes of state variable and key parameters shows that the assimilation effect is caused by adjusting the state variable and key parameters in the model during assimilation process, and along with the increase of depth the amount of adjustment decrease, which leads to the weakening of assimilation effect.

**Key words:** ensemble Kalman filter (EnKF); HYDRUS-1D; data assimilation; soil moisture profile

(责任编辑: 吕斌秀)