

## 多源遥感数据与水文过程模型的土壤 水分同化方法研究

余凡<sup>1</sup>, 李海涛<sup>1</sup>, 张承明<sup>2</sup>, 文雄飞<sup>3</sup>, 顾海燕<sup>1</sup>, 韩颜顺<sup>1</sup>, 鲁学军<sup>4</sup>

(1. 中国测绘科学研究院, 北京 100830;

2. 山东农业大学信息学院, 山东 泰安 271000;

3. 长江科学院, 湖北 武汉 430015;

4. 中国科学院地理科学与资源研究所, 北京 100101)

**摘要:**提出一种基于集合卡尔曼滤波的一维数据同化系统,对不同深度土壤层的水分含量进行同化,该系统的模型算子为分布式水文模型,观测算子是积分方程模型和条件温度指数模型.于2008年6月1日至7月2日在黑河流域进行了同化实验,结果表明,集合卡尔曼滤波能较好地处理强非线性问题,与单独DHSVM模型模拟土壤水分含量相比,同化的表层和根层土壤水分含量精度有明显提高,其中盈科站表层的均方根误差和平均误差分别减小了0.0217和0.0329,根层的均方根误差和平均误差分别减小了0.0193和0.025;临泽站的精度也有明显提高,表明多源遥感数据的同化在地表土壤水分含量的估计中具有较大的潜力.

**关键词:**同化;集合卡尔曼滤波;分布式水文模型;土壤水分含量;多源遥感

中图分类号:P237 文献标识码:A

## Data assimilation on soil moisture content based on multi-source remote sensing and hydrologic model

YU Fan<sup>1</sup>, LI Hai-Tao<sup>1</sup>, ZHANG Cheng-Ming<sup>2</sup>, WEN Xiong-Fei<sup>3</sup>, GU Hai-Yan<sup>1</sup>,  
HAN Yan-Shun<sup>1</sup>, LU Xue-Jun<sup>4</sup>

(1. Chinese Academy of Surveying & Mapping, Beijing 100830, China;

2. Shandong Agricultural University, Taian 271000, China;

3. Changjiang River Scientific Research Institute, Wuhan 430015, China;

4. Institute of Geographic Science and Resources Research, Beijing 100101, China)

**Abstract:** This paper proposed a one-dimensional soil moisture content data assimilation system based on the ensemble Kalman filter (EnKF), the distributed hydrology-soil-vegetation model (DHSVM), microwave radiative transform model (advanced integration equation model, AIEM) and optically semi-empirical model (temperature-vegetation dryness index, TVDI) for soil moisture content retrieval in bare soil. Numerical experiments were conducted at the middle reaches of the Heihe River Basin from June 1 to July 2, 2008. The results indicate that EnKF is an efficient approach to handle the strongly nonlinear problem. By assimilating multi-source remote sensing observations, the assimilation method works successfully with DHSVM and significantly improves the soil surface moisture estimation in the surface layer and root layer, the root mean square error (RMS) and mean bias errors (MBE) decrease 0.0217 and 0.0329 in surface layer and 0.0193 and 0.025 in root layer respectively, both in Yingke station. In the Linze station, the retrieve precision was also improved. It is practical and effective for soil moisture content estimation by assimilation of multi-source remote sensing data.

**Key words:** assimilation, ensemble Kalman filter (EnKF), distributed hydrological model, soil moisture content, multi-source remote sensing

**PACS:** 92.40.Lg

## 引言

土壤水分是描述地表-大气间水热与能量交换最重要的变量之一,在气象、水文、生态、农业等学科中也扮演着关键的角色<sup>[1-2]</sup>. 滑坡、泥石流等是全世界范围内常见的自然地质灾害,给国家建设和人民生命财产造成严重的损失. 土壤水分含量是早期滑坡预警的重要指标之一,对土壤抗剪强度和边坡稳定具有重要作用. 因此,精确获取土壤水分含量对建立区域尺度降雨诱发型滑坡的早期预警,具有非常重要的意义. 数据同化理论提供了一种较好结合观测和模型模拟的方法,在水文模型模拟中,大范围多时相的遥感观测能够有效地降低模型和参数的不确定性. 变分同化和卡尔曼滤波是目前应用最为广泛的数据同化算法,Verlaan等<sup>[3]</sup>比较了这两种卡尔曼滤波的优劣,指出集合卡尔曼滤波处理强非线性问题更为有效. 集合卡尔曼滤波目前在不同学科均取得了广泛的应用<sup>[4-7]</sup>.

在地表土壤水分含量的同化中,人们一般将被动微波的亮温数据作为遥感观测数据,然而被动微波数据的空间分辨率较低,较大地限制了其在地球观测中的应用. 利用合成孔径雷达,主动微波能轻易地解决空间分辨率较低的问题,有研究表明<sup>[8]</sup>,C波段和L波段的主动微波估计的裸露地表土壤水分含量精度误差在5%以内,基本与大多数野外实测的数据相当. 光学遥感数据也能用来估算土壤水分含量,通过地物反射辐射特征变化来模拟地表覆盖类型、地表温度、土壤热惯量、地表蒸散发与土壤水分含量之间的经验关系,从而实现土壤水分含量的反演.

因此,本文提出一种同化微波ASAR与光学ASTER遥感数据获取裸露地表土壤水分含量的算法,利用集合卡尔曼滤波将多源遥感数据同化到DHSVM模型中,同化模型的观测算子为AIEM模型和TVDI模型. 为对本文同化算法进行验证,于2008年6月1日至7月2日在中国甘肃省黑河流域中游进行了土壤水分含量同化实验,并利用实测土壤水分含量对同化结果进行了验证.

## 1 同化算法

一般情况下,数据同化系统应包括模型算子、观测算子、同化算法和观测数据集. Distributed Hydrology-Soil-Vegetation Model(DHSVM)模型作为本文同化系统的模型算子用来描述均匀土壤中不饱和土壤

水分的运动;观测算子(TVDI模型和AIEM模型)主要用来建立模型模拟的状态变量和观测数据之间的关系. 集合卡尔曼滤波具有强大的非线性问题处理能力,用来整合模型模拟与遥感观测的结果. 首先将具有统计特征的伪随机噪声加入土壤水分含量的初始值,产生土壤水分含量的初始集合. 利用土壤水分含量初始集合、DHSVM模型参数和大气驱动参数运行DHSVM模型,获取土壤水分含量的预报值. 当ASAR/ASTER观测数据有效时,将预报的土壤水分含量和其他模型参数作为输入,利用TVDI模型和AIEM模型计算模拟的TVDI指数和后向散射系数,并集合卡尔曼滤波更新土壤水分含量. 然后将更新的土壤水分含量重新初始化模型参数,并进入下一时刻,当遥感观测再次可用时,再进行同化(主要流程见图1).

DHSVM是一个分布式水文-植被模型,清晰地描述了地形和植被对水热通量的影响. 模型精细地考虑了冠层截留、蒸发、蒸散、积雪融化和径流,已广泛应用于不同的流域中<sup>[9-10]</sup>. DHSVM基于流域数字高程模型来描述空间尺度上地表的水文动态过程,以DEM(Digital Elevation Model)的节点为中心,流域被分成若干计算网格单元,每个网格都被赋予各自的土壤特性和植被特性,在整个流域上,这些特性随空间位置的不同而变化. 地形特性用于模拟流域地形对短波辐射吸收、降雨、气温和坡面流的作用. 每一计算时段内,模型对流域内各网格的能量平衡方程和质量平衡方程提供联立解.

本文采用Burgers<sup>[11]</sup>改进的集合卡尔曼滤波算法作为同化算法,该算法基于蒙特卡罗和非线性动力学,能够较好地处理非线性模型中误差协方差的演变. 集合卡尔曼滤波主要分为预报和分析两步,主要步骤可参考文献[11].

采用AIEM模型来模拟微波数据的后向散射系数,该模型是目前最为广泛的真实地表电磁散射模拟模型,AIEM中的单次散射项由下式给出<sup>[12]</sup>:

$$\sigma_{pq}^o = \frac{k_1^2}{2} \exp(-s^2(k_z^2 + k_{sz}^2)) \quad (1)$$

$$\sum_{n=1}^{\infty} \frac{s^{2n}}{n!} |I_{pq}^n|^2 W^n(k_{sx} - k_x, k_{sy} - k_y)$$

$$I_{pq}^n = (k_{sz} + k_z)^n f_{pq} \exp(-s^2 k_z k_{sz}) +$$

$$\frac{(k_{sz})^n F_{pq}(-k_x, -k_y) + (k_z)^n F_{pq}(-k_{sx}, -k_{sy})}{2} \quad (2)$$

式中, $k_1$ 是介质1中的自由空间波数, $s$ 是土壤表面均方根高度, $W^n(k_{sx} - k_x, k_{sy} - k_y)$ 是地表相关函数

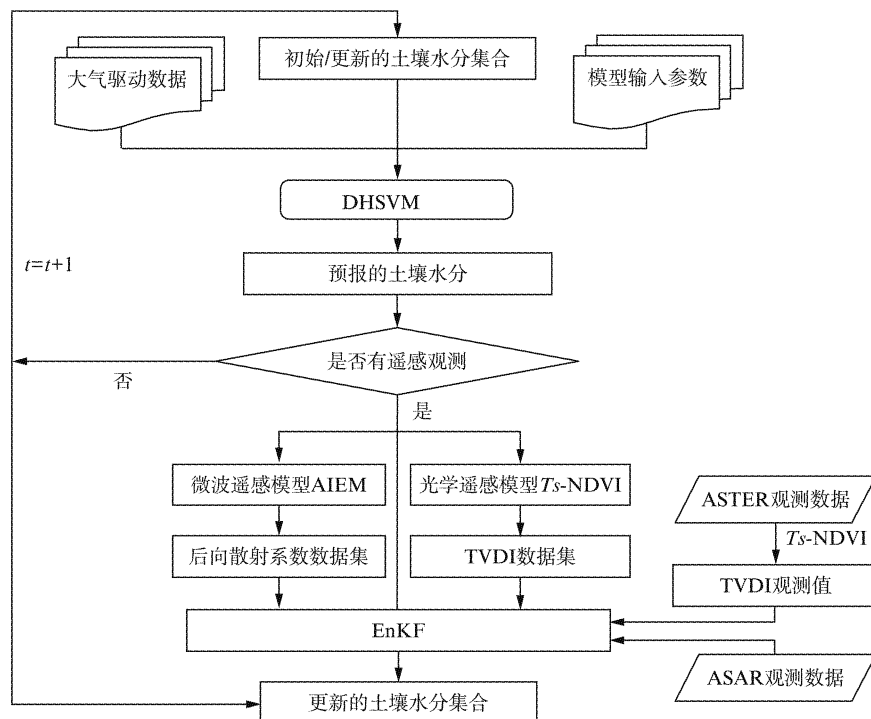


图1 集合卡尔曼滤波同化遥感观测的流程图

Fig. 1 Schematic description of the soil moisture assimilation algorithm

的  $n$  阶傅里叶变换,  $k_x = k \cos \theta_i$ ,  $k_{xz} = k \cos \theta_s$ ,  $k_x = k \sin \theta_i \cos \varphi$ ,  $k_{sx} = k \sin \theta_s \cos \varphi_s$ ,  $k_y = k \sin \theta_i \sin \varphi$ ,  $k_{sy} = k \sin \theta_s \sin \varphi_s$ ,  $\theta_i$  是入射角,  $\varphi$  是入射方位角,  $\theta_s, \varphi_s$  分别是散射角和散射方位角.  $F_{pq}$  和  $f_{pq}$  都是菲涅尔反射系数相关的函数.

采用光学 TVDI 模型来模拟 ASTER 数据的土壤水分反演. Sandholt 等利用简化的温度-植被指数 ( $T_s$ -NDVI) 特征空间提出了温度植被干旱指数 (temperature vegetation drought index, TVDI) 概念, 表示为<sup>[13]</sup>:

$$\text{TVDI} = \frac{T_s - T_{s_{\min}}}{T_{s_{\max}} - T_{s_{\min}}}, \quad (3)$$

式中,  $T_{s_{\min}}$  为 NDVI 对应的最低表面温度, 对应着  $T_s$ -NDVI 特征空间的湿边;  $T_{s_{\max}}$  为 NDVI 对应的最高表面温度, 对应着  $T_s$ -NDVI 特征空间的干边. TVDI 值越大, 土壤水分含量就越接近于萎焉含水量, 土壤就越干旱; 反之, 土壤就越湿润.

## 2 研究区和数据

### 2.1 研究区与地面观测数据

在中国西北的甘肃省张掖市盈科气象试验站的农田中进行了遥感同化实验. 该实验区高程在海拔 1 000 ~ 2 000 m 之间, 年平均气温为 6 °C, 该地区较干旱, 年均降雨仅 121.5 mm, 但年潜在蒸发量超过 2 340 mm, 为降雨量的 20 余倍. 研究区主要是由人

工灌溉的农田组成, 属于中国西北的低平原部分. 盈科灌区为荒漠中的一片绿洲, 土壤质地为沙土占 16.7%, 泥沙占 74.8, 粘土占 8.5%. 主要农作物为玉米. 如图 2 所示, 研究区共建立了若干自动气象站用于长期观测, 所有的地面实测数据都来自于中国科学院西部行动计划项目“黑河流域遥感——地面观测同步试验”, 详见相关网站<sup>[14]</sup>. 土壤类型、植被类型数据均来源于国家自然科学基金委员会中国西部环境与生态科学数据中心 (DCRES)<sup>[15]</sup>.

### 2.2 遥感数据

在实验进行期间 (2008 年 6 月 1 日至 7 月 2 日), 共有 3 景 ASAR 观测数据. 获取时间为 2008 年 6 月 9 日, 6 月 25 日, 6 月 28 日. 地面分辨率为 12.5 m × 12.5 m. 在对 ASAR 数据进行预处理时, 利用 EnviView 读取 ASAR 相关参数, 并进行辐射定标, 得到 ASAR 后向散射系数, 并用 Lee 滤波对 ASAR 影像进行滤波, 抑制斑点噪声. 将 ASAR 数据与已有的 2.5 m 全色 SPOT 正射影像进行几何配准, 配准精度在 1 个像素以内.

收集的 ASTER 数据获取时间为 2008 年 6 月 4 日, 6 月 13 日. 利用 ASTER 可见光波段的第 2、第 3 波段获取 NDVI, 热红外的第 13、14 波段计算  $T_s$ . 将所有遥感影像都采样到 90 m × 90 m, 保持分辨率一致, 并分别对 ASTER 和 ASAR 进行几何校正与配准.

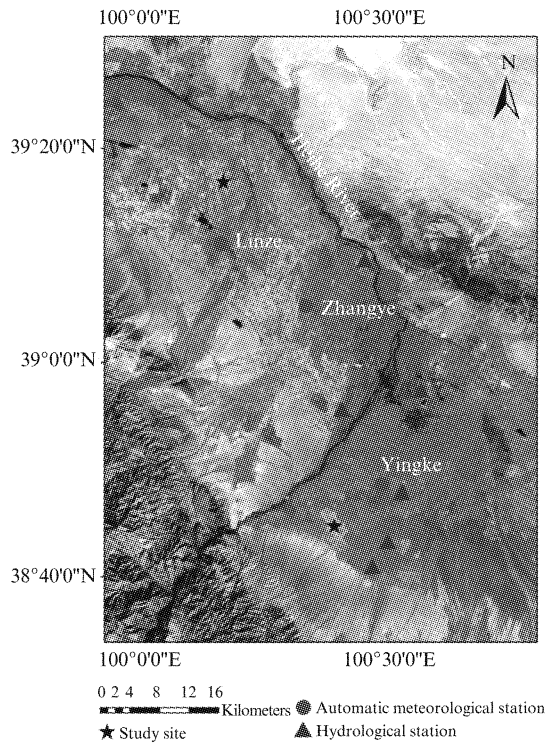


图2 黑河中游实验区  
Fig. 2 Location of the study site, middle reaches of Heihe River

### 3 实验和结果

#### 3.1 实验

为评估遥感数据在土壤水分含量同化中的影响,于2008年6月1日(儒略日152)至7月2日(儒略日183)之间进行了一些同化实验,实验样地分别位于盈科站与临泽站(图2),每个样地同化表层(0~10 cm),根层(10~40 cm),深层(40~80 cm)的土壤水分含量。

在运行集合卡尔曼滤波时需要先验知识,以便决定模型与观测误差的方差。误差的方差的选择往往影响同化的结果,而模型和观测的误差在不同时刻是变化的,因此难以同时标定模型和观测的误差。假设野外观测数据是真实数据。在同化实验进行前,即2008年6月1日前,DHSVM模型先运行一个月,获取模型误差和背景场误差,并在初始的土壤水分含量集合中加入高斯白噪声。本文中,经实测数据标定,观测数据的方差设置为观测值的5%,模型误差设为模型模拟状态变量的10%。利用大气驱动数据和地面观测的数据,DHSVM以3h为步长进行运行。依据收敛测试的结果,集合的数目设定为120个,与其他文献相关实验的设置较为一致<sup>[6]</sup>。此外,

在同化实验期间,进行一个单独运行DHSVM模型,不同化遥感数据的模拟实验。

#### 3.2 结果

图3和图4分别显示了盈科站和临泽站2008年6月1日(儒略日152)到7月2日(儒略日183)表层、根层和深层土壤水分含量的模拟结果和同化结果。可以清楚地看到通过数据同化,土壤水分含量的估计精度有明显提高。在表层土壤水分含量的同化遥感数据实验中,两个站点的模拟土壤水分含量明显过大,同化结果更接近于地表实测结果,当遥感观测数据有效时,大多数情况下,数据同化能将曲线拉近地表实测值。对于根层土壤水分含量,两个站点的模拟结果也高估了土壤水分含量,而同化结果明显好一些。不同于对表层和根层土壤水分含量的改进,同化深层的土壤水分含量并没有太大变化,这可能是由于深层土壤水分含量本身变化就很小,难以受到影响。

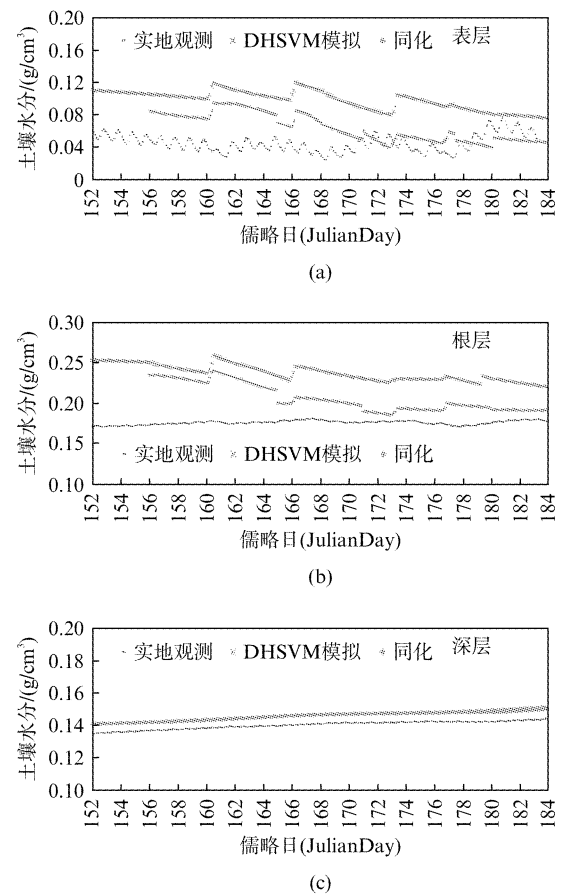


图3 实地测量、DHSVM模型模拟和同化结果的比较,2008年6月1日(儒略日152)到7月2日(儒略日183)张掖盈科站

Fig. 3 Results of assimilation of soil moisture in surface layer, root layer and deep layer at Yingke station from June 1 (Julian day 152) to July 2 (Julian day 183), 2008

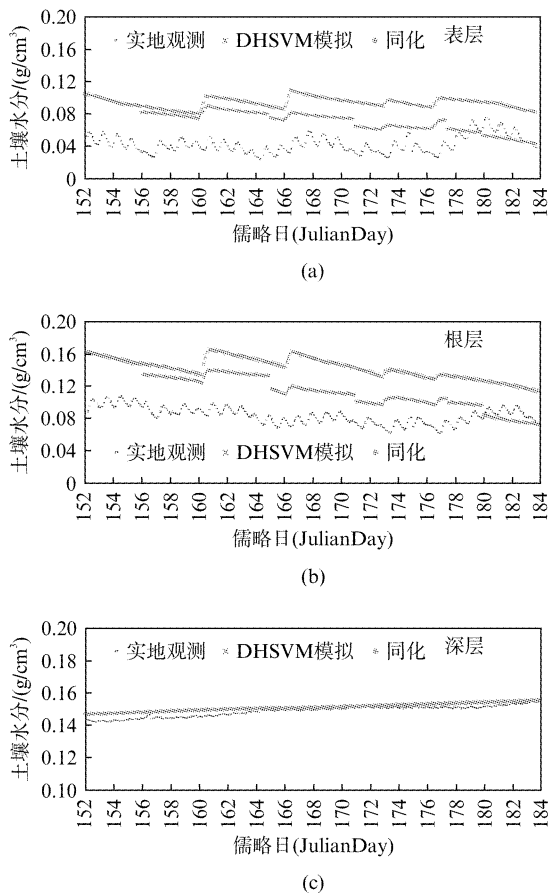


图4 实地测量、DHSVM 模型模拟和同化结果的比较, 2008年6月1日(儒略日152)到7月2日(儒略日183)张掖临泽站

Fig. 4 Results of assimilation of soil moisture in surface layer, root layer and deep layer at Linze station from June 1 (Julian day 152) to July 2 (Julian day 183), 2008

从图3和图4可以看到,表层模拟结果大于实测土壤水分含量时,同化系统会减少此时表层的土壤水分含量,使其慢慢地靠近表层实地观测的结果;进而,减少土壤水分的信息会通过同化误差矩阵传递到根层,同化系统也会减少根层土壤水分含量的模拟结果,使其接近地面实测值.对于深层土壤水分含量,由于土壤水分含量变化不大,且模型误差设定相对较小,根层土壤水分含量的变化对其影响很小,因此模拟和同化的结果都很接近深层土壤水分含量的观测值.

为对比模拟结果、同化结果和地表实测数据,采用统计学中的均方根误差(RMS)和平均误差(MBE)来表达精度(表1).对于盈科站,表层土壤水分含量模拟的均方根误差和平均误差分别为0.0466和0.0534.当同化遥感观测后,均方根误差和平均误差变成0.0249和0.0205,分别减小了0.0217和0.0329.根层的土壤水分含量,同化结果

表1 同化与模拟土壤水分含量误差分析

Table 1 Error statistics of simulation and assimilation of surface soil moisture

土壤水分含量	模拟土壤水分含量		同化土壤水分含量		
	均方根误差	平均误差	均方根误差	平均误差	
盈科站	表层	0.0466	0.0534	0.0249	0.0205
	根层	0.0524	0.0513	0.0331	0.0263
	深层	0.0058	0.0057	0.0054	0.0053
临泽站	表层	0.0519	0.0507	0.0363	0.0328
	根层	0.0498	0.0485	0.0315	0.0287
	深层	0.0029	0.0027	0.0027	0.0023

也有较大地改进,均方根误差和平均误差分别减小了0.0193和0.025.表明本文同化算法是有效的,大幅提高了土壤水分含量的估算精度.同样,对于临泽站的根层和深层土壤水分含量,同化结果较模拟结果也有较大地改进,表层的均方根误差和平均误差分别减小了0.0156和0.0179,根层的均方根误差和平均误差分别减小了0.0183和0.0189.而对于深层土壤水分含量,两个站点的同化与模拟的结果相差都很小.

#### 4 讨论

同化后的模型精度提高了,但是土壤水分含量依然被高估,可能是因为模型模拟的结果一直大于实测的土壤水分含量,同化虽然能将模拟的土壤水分含量和实测的土壤水分含量拉近,但是并不能完全消除模拟值过高的影响.模型模拟值过高可能是由于模型参数(比如植被参数和土壤参数)和大气驱动参数(空气湿度、大气温度和大气辐射)的不确定性造成的,比如饱和土壤导水率和土壤孔隙度等土壤物理参数在DHSVM模型的三个土壤层中都设置成同样的值,这种假设可能会对模拟的土壤水分含量造成较大的误差.因此,模型的参数应该依据土壤的特点进行调整,但是受制于实验条件,并没有获得足够的参数,一些模型参数都用经验参数替代了.地表实测的土壤水分含量在一天中会有一些波动,而在模型模拟时很难看到这种变化.这可能也是由模型参数与大气驱动数据的不确定性造成的,真实的地表环境是物理模型难以精确模拟的,而且是随时间变化的,远比模型模拟的情况复杂.而本文的经验参数一般在一段时间中是固定不变的,因此,模型模拟的结果不会像地表实测土壤水分含量那样波动.在图3的表层土壤水分含量同化中,有时存在模拟、同化结果和实测数据的趋势相反的情况,模拟和同化结果增大,而实测数据却减小.主要可能是由于DHSVM模型模拟的误差造成的,包括大气驱动数据

和模型参数的误差,另外模型本身存在误差,这种误差会随运行时间积累,会造成模拟结果与实测数据的不一致。

此外,同化结果受到遥感观测的影响也很大。遥感观测主要受到遥感卫星的数据质量和观测算子的影响。因此,为了提高同化的精度,一般尽量选择质量好的遥感数据,同时提高观测算子的运算效率和精度。

## 5 结论

研究了将主动微波数据同化到 DHSVM 分布式水文模型的单点同化算法,采用集合卡尔曼滤波作为同化算法。观测算子采用 AIEM 模型精确描述地表散射状况,利用在 2008 年 6 月 1 日到 7 月 2 日获取的研究区地面实测的土壤水分含量对本文同化方案进行了验证,结果表明,首先集合卡尔曼滤波能有效地同化遥感观测到水文模型中,以获取不同深度土壤层的水分含量估计,同化方法也易于实现。与模型模拟的结果比较,表层和根层同化的土壤水分含量有较大程度的改进,相对平均误差和均方根高度也有明显的降低。但是深层土壤水分含量结果变化不大,同化结果与模拟结果非常接近,这可能是由于土壤深层水分比较稳定,很难受根层土壤水分含量的影响;其次整个同化系统涉及的数据较多,例如长时间连续的气象驱动数据、土壤和植被参数等。受实验条件所限,本文部分参数使用经验参数代替,会给同化结果带来一些误差。因此,期望在今后的研究中能获取更多高可靠的驱动数据和模型参数;其次是对高分辨率遥感数据在土壤水分含量同化应用中的一次尝试,同化系统虽然一定程度上提高了土壤水分含量的精度,但是遥感观测数据量偏少,可能会影响同化结果的可靠性。因此,在今后的研究中应综合利用其它高分辨率传感器的遥感数据,利用尽可能多的多源遥感观测数据对土壤水分含量进行同化,提高同化系统的稳定性和可靠性。对同化主动微波观测估算土壤水分含量的单点同化算法进行了初步研究,在以后的研究中,需要在以下两个方面进行深入研究,首先就是将同化的尺度由单点扩展到流域尺度,另一方面就是将本文的同化算法应用到更多不同的地表下垫面区域,尤其针对滑坡、泥石流等地质灾害易发区,综合考虑地形、坡度等对土壤水分的影响,反演区域内的地表土壤水分含量,为滑坡、泥

石流早期预警提供一定的科学和技术支撑。

## References

- [1] Brubaker K L, Entekhabi D. Analysis of mechanisms in land-atmosphere interaction[J]. *Water Resources Research*, 1996, **32**(5): 1343 - 1357.
- [2] YU Fan, ZHAO Ying-Shi, LI Hai-Tao. Research on soil moisture retrieval based on Genetic-BP Neural networks algorithm using active and passive remote sensing data[J]. *Journal of Infrared and Millimeter Waves*(余凡、赵英时、李海涛. 基于遗传 BP 神经网络的主被动遥感协同反演土壤水分. *红外与毫米波学报*), 2012, **31**(2): 283 - 288.
- [3] Verlaan M, Heemink A W. Nonlinearity in data assimilation applications; A practical method for analysis[J]. *Monthly Weather Review*, 2001, **129**: 1578 - 1589.
- [4] Kourgli A, Ouarzeddine M, Oukil Y, et al. Texture modeling for land cover classification of fully polarimetric SAR images[J]. *International Journal of Image and Data Fusion*, 2012, **3**(2): 129 - 148.
- [5] Li X, Koike T, Mahadevan P. A very fast simulated re-annealing (VFSA) approach for land data assimilation[J]. *Computer & Geosciences*, 2004, **30**(3): 239 - 248.
- [6] Huang C H, Li X, Lu L, et al. Experiments of one-dimensional soil moisture assimilation system based on ensemble Kalman filter[J]. *Remote Sensing of Environment*, 2008, **112**: 888 - 900.
- [7] Anna B, Eyal B D, Rudolf R. Modelling and monitoring urban built environment via multi-source integrated and fused remote sensing data[J]. *International Journal of Image and Data Fusion*, 2013, **4**(1): 2 - 32.
- [8] Mancini M R, Hoeben R, Troch P. Multifrequency radar observations of bare surface soil moisture content; A laboratory experiment[J]. *Water Resources Research*, 1999, **35**: 1827 - 1838.
- [9] Bowling L, Storck P, Lettenmaier D. Hydrologic effects of logging in western Washington, United States[J]. *Water Resources Research*, 2000, **36**(11): 3223 - 3240.
- [10] Cuo L, Lettenmaier D P, Mattheussen B V. Hydrologic prediction for urban watersheds with the Distributed Hydrology-Soil-Vegetation Model[J]. *Hydrological Processes*, 2008, **22**(21): 4205 - 4213.
- [11] Brubaker K L, Entekhabi D. Analysis of mechanisms in land-atmosphere interaction [J]. *Water Resources Research*, 1996, **32**(5): 1343 - 1357.
- [12] Wu T D, Chen K S, Shi J C. A study of an AIEM model for bistatic scattering from randomly rough surfaces[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2008, **46**(9): 2584 - 2598.
- [13] Sandholt I, Rasmussen K, Andersen J. A simple interpretation of the surface temperature/vegetation index space for assessment of surface moisture status[J]. *Remote Sensing Environment*, 2002, **79**: 213 - 224.
- [14] <http://water.westgis.ac.cn/indexenglish.asp>
- [15] Hughes C G, Ramsey M S. A radiometrically-accurate super-resolution approach to thermal infrared image data[J]. *International Journal of Image and Data Fusion*, 2013, **4**(1): 52 - 74.