DOI:10.7702/dlydlxxkx20130119

基于 VIC 模型与集合卡尔曼滤波的土壤水分同化研究

米素娟 1,2 ,唐家奎 1,2 ,张显峰 3 ,于新菊 1,2 ,郭强 2

(1. 中国科学院烟台海岸带研究所,山东烟台 264003;2. 中国科学院大学,北京 100049;3. 北京大学遥感与 GIS 研究所,北京 100871)

摘要:基于可变下渗能力水文模型(Variable Infiltration Capacity, VIC)和集合卡尔曼滤波(Ensemble Kalman Filter,EnKF)算法发展了一种土壤表层水分的数据同化方案,并在新疆维吾尔自治区的森林、高盖草和低盖草3种植 被覆盖类型地区进行试验。在同化过程中,有强降水存在时,3种植被覆盖类型的同化值均比实际测量值高;降水 较少时,高盖草和低盖草两种类型的同化值比实际测量值略低,而在森林地区,同化值比实际测量值略高。总体 上,该同化方案得到的结果比 VIC 模拟得到的结果更接近于实际测量值,取得较好同化效果。

关键词: VIC 模型; 集合卡尔曼滤波; 数据同化; 土壤水分

中图分类号:S152.7 文献标识码:A 文章编号:1672-0504(2013)01-0091-05

0 引言

土壤水分是水资源的重要组成部分,是陆地生态系统最重要的因素之一,准确估计表层土壤水分在环境科学领域具有重要意义^[1]。对于单点的土壤水分,可以通过实地观测和陆面过程模型两种方法获得。实际观测能获得较准确的土壤水分值,但该方法需要大量的人力物力,费时且成本较高。陆面过程主要包括地面上的热力过程(包括辐射及热交换过程)、动量交换过程(如摩擦及植被的阻挡等)、水文过程(包括降水、蒸发、蒸腾和径流等)、地表与大气间的物质交换过程以及地表以下的热量和水分输送过程^[2],陆面过程模型可以对土壤水分进行连续测量,但随着时间的变化,误差不断累积,导致模拟结果较差。常用的模拟土壤水分的陆面过程模型 有 BEPS、VIC 和简单生物圈模型(SiB2)等。

数据同化方法作为一种优化融合地球观测信 息和地学模型信息的重要方法论,在集成复杂地球 系统的模拟经验和融合地球的多源观测信息方面 扮演着重要角色^[3]。在土壤水分数据同化方面,利 用数据同化算法以提供高精度、高分辨率的土壤湿 度时空分布,对于理解陆气相互作用机理和改进气 候模拟和预测有重要意义^[4]。目前,最具有代表性 的陆面数据同化系统,有北美陆面数据同化系统 和全球陆面数据同化系统、欧洲陆面数据同化系统 统和中国西部陆面数据同化系统。近年来,国内 学者对土壤水分同化做了大量研究,如张生雷等 利用 VIC 模型和扩展卡尔曼滤波(Extended Kalman Filter)设计了单点的土壤湿度同化方案,并进 行了单点的同化试验,能完整反演土壤湿度廓线, 对土壤湿度的估计有较大的改善[4]。朱琳等利用 BEPS 模型和集合卡尔曼滤波进行区域的土壤湿度 同化,并提出了"两阶段同化"方法以提高土壤水分 的同化精度^[5]。黄春林等发展了一个基于集合卡 尔曼滤波和简单生物圈模型(SiB2)的单点陆面数 据同化方案,选用"全球能量与水循环亚洲季风之 青藏高原试验"(GAME-Tibet)的观测数据进行 同化试验,结果表明,集合卡尔曼滤波的数据同化 可以明显提高土壤水分的估测精度^[6]。聂肃平等 基于集合卡尔曼滤波方法和 AVIM 陆面模式,利 用理想试验比较了膨胀因子方案、直接随机扰动方 案、误差源扰动方案3种模式误差方案的同化效 果,讨论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^[7]。上述研究主要侧重于单 个观测台站土壤水分的精确测量和区域性土壤水 分的测量,对于不同植被覆盖情况下土壤水分同化 的研究较少。

本文基于 VIC 模型和集合卡尔曼滤波同化算法 发展了一种土壤表层水分同化方案,并通过森林、高 盖草和低盖草 3 种植被覆盖类型的试验,对其模拟 值、同化值及实测数据进行比较。

1 土壤水分同化方案

陆面数据同化系统主要由数据同化算法、陆面 过程模型组成。本实验采用常用的 VIC 模型和集合 卡尔曼滤波进行同化,同化系统框架如图 1 所示。

收稿日期:2012-02-11; 修回日期:2012-09-19

作者简介:米素娟(1987-),女,硕士研究生,从事遥感数据同化研究。 * 通讯作者 E-mail;jktang@gucas.ac.cn



1.1 VIC 模型

VIC 是一个基于空间分布网格化的分布式水文 模型^[8]。Stamm 等最初构建 VIC 模型时,将土壤分成 两层,习惯称之为 VIC-2L 模型。由于两层 VIC 模型 缺乏对表层土壤水动态变化的描述,且未考虑土层间 土壤水的扩散过程,Liang 等将 VIC-2L 模型分出一 个顶薄层,成为目前常用的 VIC-3L 模型^[8]。

VIC 模型通过网格化考虑每个网格单元的植被 覆盖类型、土壤特性及降水空间分布不均对径流的 影响,每个网格都遵循能量平衡和水量平衡来模拟 水循环的各个过程,包括土壤蒸发、地表截留蒸发、 蒸散发、侧向热通量、感热通量、长波辐射、短波辐 射、地表热通量、下渗、渗透、径流和基流等^[9]。运行 VIC 模型需要的输入文件包括土壤参数文件、植被 参数文件、植被库文件、气象强迫数据和全局控制文 件。气象控制文件描述研究区域内每个网格从起始 时间到终止时间内的气象要素,包括日最高温、日最 低温和日降水量等;土壤参数文件描述每个网格每 个土壤层的特性,包括网格号、网格高程等;植被文 件描述每个网格内各种植被类型的比例和叶面积指 数(LAI)的值;植被库文件描述不同的统计参数与每 种植被覆盖类别的关系,这些统计参数是年内变化 的逐月参数值;全局控制文件是 VIC 模型的核心控 制文件,描述输入输出文件的路径、模型运行的起始 和终止时间、输出参数等[10],是控制模型的接口、对 模型的运行起到设定和引导作用。VIC 模型的输出 参数包括降水量、蒸发量、地表径流量、地下基流量、 雪的水当量和每层土壤的水分含量等。

VIC 模型作为大尺度分布式水文模型之一,模型中包含大量具有明确含义的参数,分为植被参数和土壤参数两大类。植被参数用来确定植被类型并反映某种植被类型的特性,如最小气孔阻抗、叶面积指数、反照率以及根在每一层土壤中所占的比例,这些参数在模型中直接确定且在模型计算中保持不

变。土壤参数中有一部分参数与土壤特性密切相 关,如土壤饱和体积含水量、土壤饱和水势、土壤饱 和力传导度、土壤类型、土壤含砂量及土壤粘土含量 等,在模型中直接确定,不会改动。VIC 模型另外7 个土壤参数与流域产流密切相关,需要用实测水文 数据率定,包括饱和容量曲线形状参数(B)、底层土 壤 24 h内产生基流的最大值(Dsmax)、基流非线性 增长发生时占 Dsmax 的比例(Ds)、基流非线性增长 发生时底层土壤含水量与最大土壤含水量的比值 (Ws)、三层土壤厚度(d1、d2、d3)^[9]。

1.2 集合卡尔曼模型

集合卡尔曼滤波(EnKF)本质上是一种基于蒙 特卡罗采样方法的传统卡尔曼滤波的近似^[5],EnKF 数据同化就是集合滤波和 Kalman 滤波的有机结 合^[11]。EnKF 自 20 世纪 90 年代中期问世至今逐渐 成熟并已成为数据同化领域的一个热点,利用 En-KF 进行数据同化,最初的工作由 Evensen 于 1994 年完成。它利用了蒙特卡罗方法的思想,用符合高 斯分布的一组随机变量(*N*)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,从而 避免了背景误差方差是已知的或不随时间变化的假 设^[6]。在 EnKF 中,观测值的集合是通过在观测值 的基础上产生均值为 0、方差等于观测误差的高斯分 布随机扰动,否则,更新的集合将产生一个非常小的 协方差^[12]。EnKF 的计算步骤如下:

(1)初始化背景场。给定 k 时刻 N 个符合高斯 分布的随机变量 $X_i(i=1,\dots,N)$,本文状态变量为 土壤表层体积含水量。

(2)计算随机变量在 k+1 时刻的预报值 X_i^{k+1} 。 非线性系统可以用时间上离散的状态和观测方程表达,因此,k+1 时刻的预报值可以用下式计算:

 $X_i^{k+1} = f(X_i^k) + w(i,k), w(i,k) \sim N(0,Q_k)$ (1) 式中: X_i^{k+1} 代表 k+1 时刻的预报值; $f(X_i^k)$ 代表过 程模型,本文中代表 VIC 模型;w(i,k)代表 X_i^{k+1} 的 不确定性; Q_k 为模型误差矩阵;w(i,k)为均值为 0、 方差为 Q_k 的高斯白噪声。

(3)计算 k+1 时刻的卡尔曼增益矩阵 K_{k+1} :

$$K_{k+1} = P_{k+1}^{-} H^{\mathrm{T}} (HP_{k+1}^{-} H^{\mathrm{T}} + R_{k+1})^{-}$$

 $=C_{k+1}^{xy}[C_{k+1}^{yy}+R_{k+1}]^{-1}$ (2)

式中:H 是线性的观测转换矩阵,本文中 H 为 1; R_{k+1} 为观测误差; $HP_{k+1}^- = C_{k+1}^{SV}$,表示模型预测值的 方差; P_{k+1}^- 表示 k+1 时刻状态变量集合的误差方差 矩阵,表示 k+1 时刻模型状态变量和模型预测值之

间的协方差,有如下关系式:

$$P_{k+1}^{-} = E[e_{k+1}^{-}e_{k+1}^{-}] = \frac{1}{n-1}e_{k+1}^{-}e_{k+1}^{-}$$
(3)

$$e_{k+1}^{i-} = \{x_{k+1}^{i} - \bar{x_{k+1}}\} \quad i=1,2,\cdots,N$$
(4)

$$e_{k+1}^{-} = \{ e_{k+1}^{1-}, e_{k+1}^{2-}, \cdots, e_{k+1}^{N-} \}$$
(5)

$$P_{k+1}^{-} H^{\mathrm{T}} = C_{k+1}^{\mathrm{xy}} \tag{6}$$

(4)计算 k+1 时刻分析场的状态变量平均值
 X^a_{k+1}。利用 k+1 时刻的卡尔曼增益矩阵计算 k+1
 时刻的状态变量 X^{a,i}_{k+1}。

$$X_{i,k+1}^{a} = X_{i}^{k+1} + K_{k+1} \left(O_{k+1} - H X_{i}^{k+1} + \nu_{i,k+1} \right),$$

$$\nu_{i,k+1} \simeq N(0, R_{i+1})$$
(7)

$$\overline{X}_{k+1}^{a} = \frac{1}{N} \sum_{i,k+1}^{N} X_{i,k+1}^{a} \tag{8}$$

$$P_{k+1}^{a} = \frac{1}{N-1} \sum_{i=1}^{N} (X_{i,k+1}^{a} - \overline{X}_{k+1}^{a}) (X_{i,k+1}^{a} - \overline{X}_{k+1}^{a})^{\mathrm{T}}$$
(9)

式中: O_{k+1} 为 k+1 时刻的观测值; $\nu_{i,k+1}$ 为 k+1 时刻 均值为 0、方差为 R_{k+1} 的高斯白噪声。

(5)进入下一时刻计算,返回步骤 $(2)^{[6,13]}$ 。

集合卡尔曼滤波用蒙特卡罗集合预报方法来估 计预报误差协方差的顺序数据同化算法,其最大的 优点是不需要预报算子的切线性模式和伴随模式, 对于非线性很强、不连续的动态模型也能取得较好 的同化结果^[13-15],已被应用在加拿大气象中心、 NASA 的季节预报系统以及各种水文同化系统中^[16]。

2 试验描述

本试验中土壤水分实测数据取自 WatchDog 观 测仪观测的新疆地区 2009 年 5 月 11-30 日森林、 高盖草和低盖草 3 种植被类型的土壤表层水分,时 间分辨率为 1 h,每天的土壤水分值由 24 h 的实际 测量值平均得到。测定上述 3 种覆盖类型的站点坐 标分别为:43 88131667N,85 6676E;43 9359444N, 85 77847E;44 15177778N,85 83933E。由于数据限制, 试验采用 10 km×10 km 网格对整个新疆区域进行 划分,3 个观测台站位于 3 个不同的网格中。 VIC 的可执行程序是从华盛顿大学网站下载后 在 linux 系统中编译得到,VIC 模型的驱动数据包括 气象参数、植被参数、土壤参数等。按照 VIC 模型要 求的数据格式准备各个输入文件。其中,气象数据 采用中国气象科学数据共享网提供的日气象数据, 由反距离权重法插值得到;植被覆盖类型文件数据 参考 Maryland 大学发展的 $1 \text{ km} \times 1 \text{ km}$ 的植被覆盖 数据;土壤类型文件采用 Reynolds 等发展的 10 km $\times 10 \text{ km}$ 的土壤数据库。

3 结果分析

3.1 同化结果

取 2009 年 5 月 10 日的实测土壤水分值作为初 始的状态变量,对 VIC 的 7 个参数分别用符合高斯 分布、均值为参数预估计值、方差为预估计值的 20% 进行扰动,以产生集合数目为 200 的集合;并行将 200 组数据输入 VIC 模型,得到每天 200 组 0~10 cm 土壤水分的模拟值,并用模拟值集合的平均值代 表当日的土壤水分模拟值。

同化过程中,以5d为一同化周期,即在5月15 日、20日、25日、30日,将该天的土壤湿度观测值进 行200次扰动,得到200个该天的观测值集合,将土 壤水分观测值集合加入EnKF模型中进行同化。经 过计算,观测误差和模型误差分别为0.0001和 0.001。整个同化过程如下:首先,用初始数据驱动 VIC模型得到第一个周期即5月11-15日的模拟 值集合;其次,将5月15日的观测值集合、模拟误 差、观测误差和模拟值集合输入EnKF,得到第一个 周期的同化结果;最后,将5月15日的同化结果作 为下一个周期的初始值再次驱动VIC模型进行下一 个周期的同化。重复上述步骤,直到完成所有的周 期为止,得到整个时间段的同化结果。3个台站 2009年5月11-30日的实测、模拟和同化的土壤水 分结果如图2所示。





由图 2a 可得,森林覆盖地区土壤表层水分随降 水变化较大。在 5 月 15 日、20 日、25 日、30 日,同化 结果值比模拟值更接近于实测值,说明基于集合卡 尔曼滤波取得较好的同化效果。在整个模拟过程 中,模拟值均高于实测值;在 5 月 26 日强降水时,同 化值比模拟值更接近于实际观测值。

由图 2b 可得,高盖草地区土壤表层水分受降水 变化较大,在5月24-27 日更明显。整个模拟过程 中模拟值均小于实测值,说明模拟的土壤水分值偏 低。在5月15日、20日、25日、30日,同化结果接近 于实际观测值,说明同化效果较好。在5月11-24 日降水较少时,同化值更接近于实际观测值,但5月 26-29日强降水,过高估计土壤水分值,同化值比模 拟值更偏离实测值。总之,同化取得了较好结果。

由图 2c 可得,低盖草覆盖地区土壤表层水分受 降水影响比高盖草更大。整个模拟过程中模拟值均 小于实测值。在 5 月 15 日、20 日、25 日、30 日,加入 集合卡尔曼滤波同化后,同化值更接近于实际观测 值。在整个同化过程中,同化值均比模拟值更接近 于实测值。

3.2 误差统计

为了进一步定量描述土壤水分的同化情况,本 文采用均方根误差(*RMSE*)和平均误差(*AVE*)比较 3 个站点的实测值、模拟值和同化值(表 1),均方根 误差和平均误差越小,说明更接近于实测值。

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^{N} (O_t - X_t)^2}$$
(10)

$$AVE = \frac{1}{N} \sum_{t=1}^{N} (O_t - X_t)$$
(11)

式中:N 为整个观测周期的天数; O_t 为t 时刻的实测 值; X_t 为t 时刻的模拟值或同化值, $t = 1, 2, \dots, N^{[6]}$ 。

表1 3种植被覆盖类型的均方根误差和平均误差比较 Table 1 The RMSE and AVE comparison among three different land cover types

方法	森林		高盖草		低盖草	
	模拟	同化	模拟	同化	模拟	同化
RMSE	0.04	0.024	0.034	0.018	0.04	0.012
AVE	-0.034	-0.016	0.025	-0.002	0.032	0.001

从表 1 可得,对于 3 种植被覆盖类型,同化均使 RMSE 和AVE 的值降低。对于森林,同化结果的均 方根误差为模拟结果的 2/3,平均误差为模拟结果的 1/3;对于高盖草,同化结果的均方根误差为模拟结 果的 1/2,平均误差为模拟结果的 1/10;对于低盖 草,同化结果的均方根误差为模拟结果的 1/3,平均 误差为模拟结果的 1/20。试验证明,基于 VIC 和集 合卡尔曼滤波组合的数据同化结果比 VIC 模型的模 拟结果更接近实际测量结果。

4 结语

基于 VIC 模型和集合卡尔曼滤波的同化系统, 能降低过程模拟的均方根误差和平均误差。5 月 15 日、20 日、25 日、30 日加入实际观测值进行同化,同 化结果比模拟值更接近实际测量值,说明同化取得 较好的效果。

森林、高盖草和低盖草的土壤表层水分对降水 比较敏感,森林覆盖地区 VIC 模拟值比实际测量值 偏大,高盖草和低盖草 VIC 模拟值均比实际测量值 偏小。在同化过程中,有强降水存在时,3 种植被类 型表层水分扰动较大,使同化值均比实际测量值偏 高;降水较少时,高盖草和低盖草同化值比实际测量 值略低,森林地区同化值比实际测量值略高。

试验中没有涉及集合数、模型误差、观测误差等 对同化结果的影响研究,另外,仅限于对森林、高盖 草、低盖草3种类型土壤表层水分的变化进行分析, 深层的土壤水分尚未考虑,这是进一步研究的方向。

参考文献:

- [1] 王斌,王伟锋,窦慧芳.基于分数阶植被归一化系数的表层土壤
 湿度分析模型[J].安徽农业科学,2009,37(9):4330-4332.
- [2] 汪薇,张瑛.陆面过程模式的研究进展简介[J]. 气象与减灾研 究,2010,33(3):1-6.
- [3] 摆玉龙,李新,韩旭军.陆面数据同化系统误差问题研究综述 [J].地球科学进展,2011,26(8):795-804.
- [4] 张生雷,谢正辉,田向军,等.基于土壤水模型及站点资料的土 壤湿度同化方法[J].地球科学进展,2006,21(12):1350-1362.
- [5] ZHU L,CHEN J M,QIN Q M,et al. Optimization of ecosystem model parameters using spatio-temporal soil moisture information[J]. Ecological Modelling, 2009, 220: 2121-2136.
- [6] 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.
- [7] 聂肃平,朱江,罗勇.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,2010,34(3):580-590.
- [8] LIANG X, LETTENMAIER D P, WOOD E F. Surface soil moisture parameterization of the VIC-2L model: Elevation and modification[J]. Global Plant Change, 1996, 13(1): 195-206.
- [9] 张续军. VIC 模型在中国湿润地区的应用研究[D]. 河海大学, 2006.12.
- [10] 陈小凤. VIC 模型在白莲河流域的应用研究[J]. 人民黄河, 2010,32(1):37-38.
- [11] 高山红,吴增茂,谢红琴. kalman 滤波数据同化中的发展和应 用[J]. 地球科学进展,2000(5):571-575.

- [12] MORADKHANI H, HSU K, HONG Y, et al. Investigating the impact of remotely sensed precipitation and hydrologic model uncertainties on the ensemble streamflow forecasting[J]. Geophysical Research Letters, 2006(L12107):5.
- [13] EVENSEN G. Sequential data assimilation with a non-linear quasigeostrophic model using Monte Carlo methods to forecast error statistics[J]. Geophysical Research, 1994, 99(C5):10143-10162.
- [14] HOUTEKAMER P L, MITCHELL H L. Data assimilation using an ensemble Kalman filter technique[J]. Monthly Weather Review, 1998, 126(3):769-811.
- [15] ANDERSON J L. An ensemble adjustment Kalman filter for data assimilation [J]. Monthly Weather Review, 2001, 129: 2884-2903.
- [16] REICHLE R H. Data assimilation methods in the Earth sciences[J]. Advances in Water Resources, 2008, 31:1411-1418.

Soil Moisture Assimilation Based on Variable Infiltration Capacity and Ensemble Kalman Filter

MI Su-juan^{1,2}, TANG Jia-kui^{1,2}, ZHANG Xian-feng³, YU Xin-ju^{1,2}, GUO Qiang²

(1. Yantai Institute of Coastal Zone Research, Chinese Academy of Sciences, Yantai 264003;

2. University of Chinese Academy of Sciences, Beijing 100049; 3. Institute of Remote

Sensing and GIS, Peking University, Beijing 100871, China)

Abstract: Soil moisture is not only an important part of water resource, but also one of the most important factors of the land ecosystem. Accurate soil moisture is significant in environmental science field. This paper presents a soil water data assimilation scheme which based on Variable Infiltration Capacity (VIC) and the Ensemble Kalman Filter (EnKF). The scheme was experimented on three land cover types, which are forestry, high cover grass land and low cover grass land, in the Xinjiang Uygur Autonomous Region. During the assimilation process, when heavy rain, the estimated value of new scheme was larger than field surveys. However, when little rain, the estimated value was less than the field survey in high cover grass land and low cover grass land, but larger than field survey in forestry. In general, the experiment results showed that the estimated values of soil moisture by this new scheme were more close to field surveys than the simulated values of VIC, which showed that the proposed soil moisture assimilation scheme based on VIC and EnKF has a preferable result.

Key words: VIC model; EnKF; data assimilation; soil moisture

(上接第 63 页)

A New Method to Measure the Neighboring Relations of Landscapes. Boundary Adjacency Index

XIE Zheng-feng

(School of Geography Science and Tourism, Jiaying University, Meizhou 514015, China)

Abstract: The landscape adjacency is an important feature of landscape structure. An index named boundary adjacency index (BAI) was put forward in this paper based on the common boundary of patches. The index can be used to measure the landscape boundary adjacency characteristic at patch level and class level. To construct BAI three preconditions were given:1) landscapes consist of mosaic patches;2) to a certain patch, the values of the BAI of others patches respectively adjacent to it are not necessarily equal and the larger value indicates the closer spatial relation to it, and it is the same at class level;3) the BAI of the other patch to the patch and the BAI of the patch to the other one is not the same one. And then the boundary adjacency indexes at patch level and class level were structured. With the land use of Zengcheng City in 2009 as a case, the BAI was applied at patch level and class level. Research shows that BAI can be used to measure the adjacency relation of landscape at patch level and class level. Further research is needed on the ecology significance of BAI and the mechanism of spatial adjacency of landscape. Key words: landscape structure; landscape index; boundary adjacency index; patch level; class level