

# 集合卡尔曼滤波同化估算地表水热通量的研究

刘翔舫<sup>1</sup>, 王鹏新<sup>1</sup>, 唐伯惠<sup>2</sup>, 黄健熙<sup>1</sup>, 陶欣<sup>3</sup>

(1. 中国农业大学信息与电气工程学院 北京 海淀区 100083; 2. 中国科学院地理科学与资源研究所 北京 朝阳区 100101;  
3. 美国马里兰大学帕克分校地理科学系 美国 马里兰帕克大学城 20742)

**【摘要】**采用了顺序同化方法, 利用集合卡尔曼滤波(EnKF)耦合一个简单陆面过程模型, 从而完成了改善地表水热通量估算精度的研究工作。在建立同化系统的过程中, 对同化系统的模型误差进行了探讨和设定, 并通过已建立的同化系统对EnKF中的集合大小设定进行了试验。利用山东禹城试验站提供的站点实测数据与MOD16产品数据, 进行同化系统的驱动和通量结果的验证。结果表明, 以EnKF方法的数据同化系统能较好地完成对地表水热通量的估算, 通过与MODIS ET(MOD16A2)产品的对比试验, 证明该方法具有一定的稳定性和适用性, 能较准确地对地表水热通量进行估算。

**关键词** 集合卡尔曼滤波; 陆面过程模型; 地表温度; 顺序滤波的同化; 地表水热通量  
**中图分类号** TP79 **文献标志码** A **doi**:10.3969/j.issn.1001-0548.2014.03.013

## Estimating Surface Turbulent Fluxes by Using by Ensemble Kalman Filter Method

LIU Xiang-ge<sup>1</sup>, WANG Peng-xin<sup>1</sup>, TANG Bo-hui<sup>2</sup>, HUANG Jian-xi<sup>1</sup>, and TAO xin<sup>3</sup>

(1. College of Information and Electrical Engineering, China Agriculture University Haidian Beijing 100083;  
2. Institute of Geographic Sciences and Natural Resources Research Chaoyang Beijing 100101;  
3. Department of Geographical Sciences, University of Maryland Maryland College Park USA 20742)

**Abstract** Nowadays, more attention has been focused on the estimation of the land surface turbulent fluxes (sensible and latent heat fluxes) with data assimilation method. In this study, a sequential data assimilation scheme is developed based on the concept of the Ensemble Kalman filter (EnKF). It assimilates land surface temperature into a simple land surface model which based on the energy balance theory for the estimation of surface turbulent fluxes. Moreover, from the perspective of error estimation, the simple schemes for estimating model errors and ensemble size are discussed. After construction of the assimilation system, the several numerical experiments tested by Yucheng cropland site in the province of Shandong. Results show that the land surface turbulent fluxes can be retrieved with satisfactory accuracy by using our method(compared to MODIS ET products(MOD16A2), the RMSE of ET results are dropped from 4.18 mm to 2.99 mm), which indicate the availability of our method in the prediction of surface turbulent fluxes.

**Key words** ensemble Kalman filter; land surface model; land surface temperature; sequential assimilation; surface turbulent fluxes

地表水热通量在地面表层范围内的大气、水文和生态等诸多领域研究中具有重要的地位, 其影响着地表水汽和能量的输送过程。在农业大环境中, 地表水热通量的估计对于农业管理、作物估产以及农业生态系统的研究具有十分重要的意义。目前, 在能量平衡原理、空气动力学原理、土壤-植被-大气传输理论的不断展下, 已经建立了多种估算地表水热通量的模型。如研究者利用建立相关的能量和物质交换过程方法的单层或多层模型, 进行水热

通量的估算研究<sup>[1-2]</sup>。

另一方面, 陆面数据同化方法的发展使数据同化技术在陆面过程研究中得到较多的应用<sup>[3]</sup>。其中, 数据同化技术在实际应用中可同时利用多源观测数据结合动态模型的多层次架构的特点, 能够将模型与观测通过数据同化算法有机结合在一起, 共同作用于对目标参数的估算反演中, 从而获得高分辨率的估算结果, 使最终估算的目标数据质量得到提升。因此, 数据同化技术在估算地表水热通量的相关研

究也不同程度地开展起来,文献[4]应用变分同化方法,将遥感反演的地表温度数据同化入陆面能量平衡与地表温度为约束条件的动态方程组中,得到了潜热和地表蒸散参数的有效估算。文献[5]应用扩展卡尔曼滤波方法,并通过设计方案中考虑观测和模型误差的不确定性,将卫星反演温度误差与实测表层土壤温度进行拟合,以拟合作为观测误差同化入地表陆面模型中,估计下层土壤温度与地表蒸散。文献[6]以能量平衡方程的方法提出了一种近地表陆面模式,并应用变分同化方法与遥感反演的地表温度耦合,实现对地表热通量的估算。文献[7]应用弱约束的四维变分方法,对简单陆面模式进行了数值试验,将地面站实测地表温度与陆面模式等关键参数共同建立代价函数,实现地表水热通量的估计。文献[8]采用通用陆面模式结合EnKF对地表水热通量进行估算。文献[9]综合考虑了能量传递过程中土壤表层与植被冠层温度之间的差异,并应用变分同化将地表温度同化入一个双层模型,估算了地表水热通量数据。

由于多数地表水热通量的研究工作是基于国外通量观测站数据的支持,对国内研究的参考价值有限。因此本文尝试利用国内试验站点的观测数据,通过搭建一个以同化地表温度改进陆面过程模型中,地表水热通量估计的数据同化系统,为数据同化方法在农田环境中的地表水热通量估计的进一步研究提供借鉴与基础。

## 1 研究站点概况与数据预处理

本文选用的试验数据来自山东禹城农业综合试验站(36.83N, 116.57E),如图1所示,站点位于山东省禹城市,是中国科学院首批的开放式试验站,并于1999年由国家科技部设立为首批国家重点试验站。该站地理环境处于黄淮海平原,地处暖温带半湿润季风气候区。年平均气温13.1℃,年均降水量582 mm,光热资源丰富,雨热同期,是主要的农业生产区,站点农作物主要是小麦、玉米等,该站点同时具有农业生产的有利环境和农业作物类型的代表性。

禹城农业综合试验站通过不断发展,陆续建立了多种类型仪器的监测与研究系统,可具备同时监测气象数据、土壤数据、植物信息等能力。为农业的持续发展和陆面过程中的土壤-植物-大气间的能量循环监测研究,提供完备的科技手段和观测数据支持。

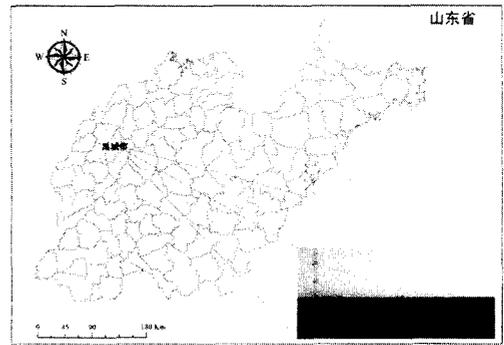


图1 山东禹城试验站地理概况

## 2 EnKF同化系统的构建

图2所示为本文的研究数据同化系统的具体流程图。从图中可以大致看出,数据同化系统的主要结构可分为:外部观测、同化算法、动态模型3个主要部分组成。数值试验研究中的外部观测数据与气象驱动数据由上述国内禹城试验站点的实测数据提供,下面将对数据同化系统的同化算法和陆面过程模型分别进行描述。

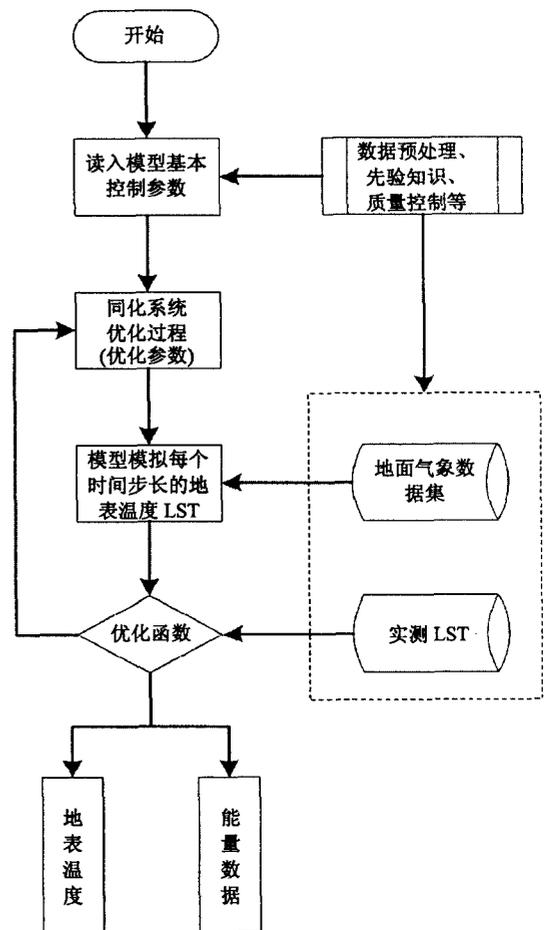


图2 同化系统流程图

### 2.1 EnKF同化方法

为了有效地应对非线性问题,从而合理得到误

差的估计, 一种集合的卡尔曼滤波(ensemble Kalman filter, EnKF)发展起来<sup>[10-18]</sup>。EnKF运用Monte-Carlo方法与集合的思想, 主要解决了一般同化方法在实际应用中背景误差协方差矩阵估计和预报困难的问题; 解决非线性系统的近似问题; 避免使用变分模式下的伴随模式; 较易考虑模式误差问题; 并可以实现并行化计算<sup>[19]</sup>。因此, 本文选择EnKF方法作为数据同化系统中的优化算法。

EnKF算法作为应用最广泛的非线性滤波同化方法, 最初由文献[12]提出并应用于海洋数据同化领域。但是随着近年来陆面数据同化研究的不断发展, EnKF方法也逐步应用于陆面过程研究中, 其EnKF算法由KF算法推衍而来, 主要结构仍分为预报阶段和分析阶段两个组成部分不变, 具体算法描述如下: 将动态模型  $M$  在每个随机变量第  $i$  时刻的模型的预报值  $x_i$  作为背景场的状态集合向量  $X^b$ ,  $X^b = (x_1, x_2, \dots, x_m) \in \mathbb{R}^{n \times m}$ ,  $\mathbb{R}^{n \times m}$  为  $n \times m$  维向量空间, 则有:

$$x_i = M_{i-1}(x_{i-1}) + l_i \quad (1)$$

式中,  $x_{i-1}$  为状态变量在第  $i-1$  时刻的预报值;  $l_i$  表示在第  $i$  时刻的模型中的误差。对应  $\langle l_i \cdot l_i^T \rangle = Q$ ,  $\langle \cdot \rangle$  表示数学期望,  $n$  为模型中的状态变量个数,  $m$  表示集合大小。

观测的表达式为:

$$y_i = H(x_i) + \xi_i \quad (2)$$

式中,  $H$  为观测矩阵, 一般用于将模型的状态变量转变到外部观测空间;  $y_i$  为观测向量, 其组成观测矩阵为  $Y = (y_1, y_{i+1}, \dots, y_m) \in \mathbb{R}^{n \times m}$ ;  $\xi_i$  为观测误差, 其应为无偏的  $\langle \xi_i \rangle = 0$ , 且服从方差为  $\langle \xi_i \cdot \xi_i^T \rangle = R$  的正态分布。

标准Kalman滤波预报阶段的方程组为:

$$x^f = Mx^a \quad (3)$$

$$P^f = MP^a M^T + Q \quad (4)$$

式中,  $x^a$  表示模型前一时刻的分析值;  $P^f$  表示背景(预报)误差协方差矩阵;  $P^a$  表示上一时刻分析场的误差协方差矩阵;  $Q$  表示对应模型的误差协方差; 上标  $a$ 、 $f$  分代表代表同化系统状态变量的分析和预报状态。

标准Kalman滤波分析阶段的方程组为:

$$K = P^f H^T (HP^f H^T + R)^{-1} \quad (5)$$

$$P^a = (I - KH) P^f (I - KH) T + KRK^T = (I - KH) P^f \quad (6)$$

$$x^a = x^f + K(y - Hx^f) \quad (7)$$

式中,  $K$  为Kalman增益, 其计算过程主要取决于  $P^f$ 。而在EnKF中,  $P^f$  是由模型预报集合的取样误差协方差近似得到。集合均值  $\bar{x}$  表示为:

$$\bar{x} = \frac{1}{m} \sum_{i=1}^m x_i \quad (8)$$

并将其集合成员的扰动  $x'_i$  定义为:

$$x'_i = x_i - \bar{x} \quad (9)$$

则  $P^f$  计算过程变为:

$$P^f = \frac{1}{m-1} \sum_{i=1}^m x'_i x'^T_i \quad (10)$$

在协方差估计的过程中, 只要用到  $\bar{x}$ , 分母则为  $(m-1)$ , 因为这样的估计才确保无偏。以上即为EnKF算法的主要计算流程。

### 2.2 动态模型描述

本文选用的估算地表水热通量的简单陆面过程模式是由地表能量平衡原理和土壤-植被-大气能量传输原理基础上发展而来<sup>[6]</sup>。利用前向微分方程随时间变化给出地表温度的估计:

$$\frac{dT_s}{dt} = \frac{2\sqrt{\pi\omega}}{P} (R_n - H - LE) - 2\pi\omega(T_s - T_{deep}) \quad (11)$$

$$T_{deep} = \sum_{i=1}^{DN} T_{si} / DN \quad (12)$$

式中,  $R_n$  是地表净辐射;  $H$  和  $LE$  分别是显热通量和潜热通量;  $P$  是地表的热惯量;  $\omega$  为每日频率 (1/86 400 s);  $T_{si}$  表示前一天第  $i$  时刻的地表温度值;  $T_{deep}$  表示土壤温度;  $DN$  表示一天内的时间步长。式(11)为该模式的主体部分, 其中热通量数据(潜热、显热与净辐射)依赖于地表温度  $T_s$  的状态估计。因此, 准确估算与优化当前时刻的地表温度  $T_s$ , 对下一刻的地表水热通量数据的估计至关重要。所以, 可认为通过改变地表温度的预报, 进而可改变地表水热通量的估计精度。模型中的参数具体求解与设定参考文献[6-7, 20]。

### 2.3 同化系统中误差的确定

在数据同化系统的研究中, 由于系统中模型表述状态变量的复杂性, 在实际应用中同化系统的误差估计与处理是受到广泛关注的问题。数据同化的目的是在动态模型的动力框架内, 通过综合不同来源的不同时空分辨率的常规和非常规观测, 将动态模型和各种观测信息集成为不断依靠观测而调整优化模型的轨迹, 减小模型预报误差而获得更高精度的预报值。因此, 误差的研究应始终贯穿于数据同化的研究之中<sup>[21]</sup>。

通常情况认为动态模型在运行一段时间后将自

动达到一种稳定和平衡的状态,此时的模型模拟结果在一定程度上具有模型自身误差的代表性。本文根据该方法,利用运行一段时间后的模型模拟结果与地面试验站点的实测值进行比较,可获得统计上的模型误差。

如图3所示,利用禹城试验站点提供的气象数据使模型从第91天开始,持续运行90天的模拟结果。图中所示从第100天左右开始,模型模拟的LST结果高于地表观测的LST结果,之后在一段时间内(第100天~160天)的观测值与模拟值的差别较大,此外,图中显示在第165天左右,试验站点实测的地表温度数值有部分超过320 K,认为该部分数据存在较大的测量误差,在以后计算中以320 K为阈值,对观测站

点实测的地表温度数据进行质量控制。

为了研究并得到模型的误差,认为模型自身运行10天后的模拟结果即可达到稳定。将模型在一段时间内的(第100天~160天)对应试验站点的观测值与模型独立运行的模拟值进行以天为单位的差值平均,得到该时段一天内的模型误差变化趋势,如图4所示。图中的结果表示,地表温度的误差主要体现在白天阶段,并且白天的误差相比夜间变化的波动较大。综合考虑,模型地表温度平均高于地面实测地表温度约为1.5 K,本文在试验中假设模型运行时的模型误差为定值,不随其他变量相关,通过上述统计方式得出平均每天模型误差的均值,作为同化系统中模型误差的初始设定。

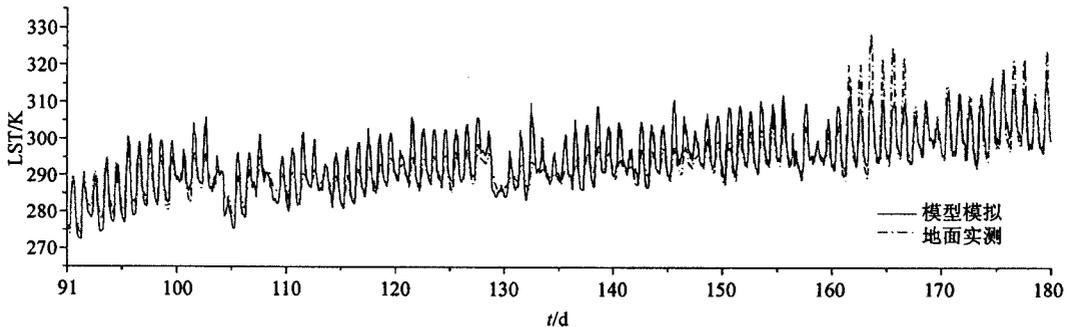


图3 山东禹城站点模拟地表温度结果与观测结果比较图

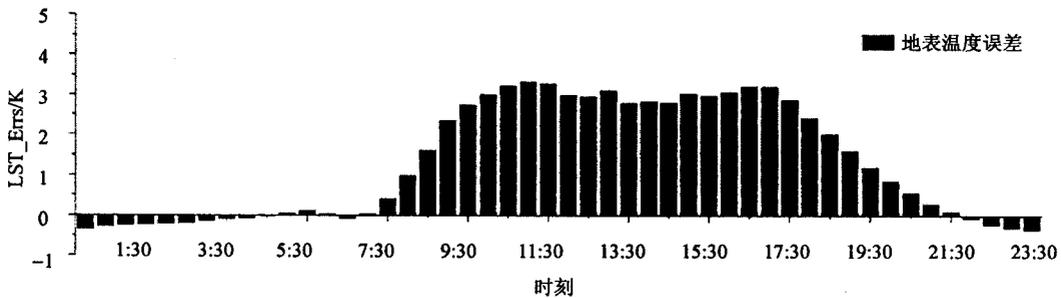


图4 一天内模拟值与观测值的差值变化

### 3 结果分析与讨论

#### 3.1 集合大小的确定

EnKF集合的多少决定最终整个数据同化系统的效率和精度。集合数目过多影响系统执行效率,过少却影响最终的同化精度。因此在本文试验中,考虑利用调整集合 $m$ 的大小,通过系统运行结果的均方根误差(root mean square error, RMSE)变化,试确定适合本同化系统的集合数目 $m$ 。

在确定集合大小的试验中,同化步长,模型误差、背景场误差、模型驱动参数等均取值相同,只

调整集合 $m$ 的大小,将得到的同化后地表温度结果与试验站观测的地表温度对比,利用误差RMSE变化分析EnKF的集合数目的影响,如图5所示。

由图5可以看出, RMSE变化趋势随着集合数目的增大显著减小,当集合大小为30时RMSE最小,之后逐渐趋于稳定。因此取 $m=30$ 。

#### 3.2 同化结果

利用已设定的数据同化系统,对山东禹城2009年4~6月内进行EnKF的数据同化试验。考虑数据的连续性和演示效果的清晰性,选择试验站点的第91天开始之后30天的数据同化结果进行展示。

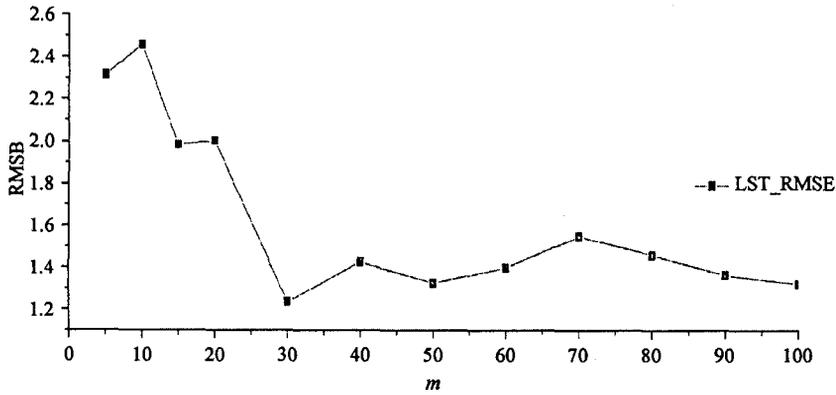
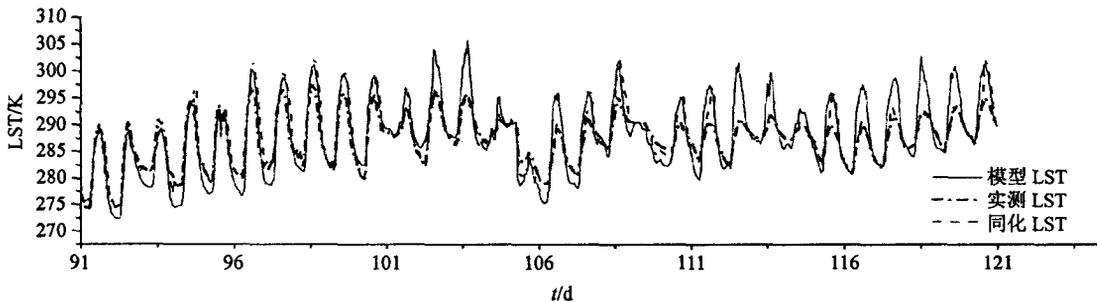


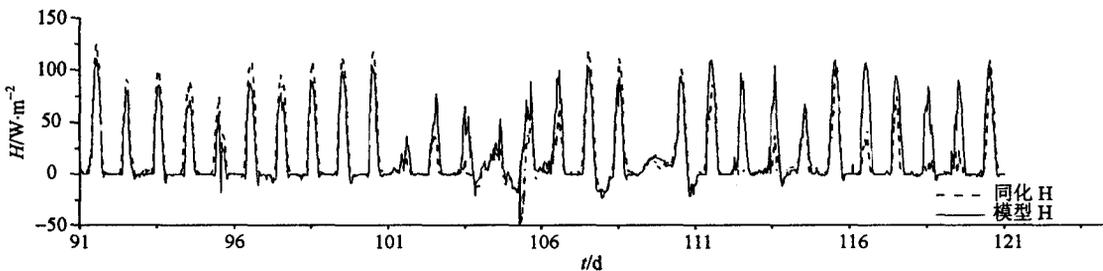
图5 集合大小对同化系统的影响

图6所示为模型模拟的LST和同化后的LST的趋势变化, 两者总体趋势与实测的LST变化相同, 但是从图中可以明显看出, 同化后的LST趋势更接近实测的LST曲线, 说明相较于模型模拟的LST, 同化后的LST精度得到了有效的改善。其中, 模型模拟

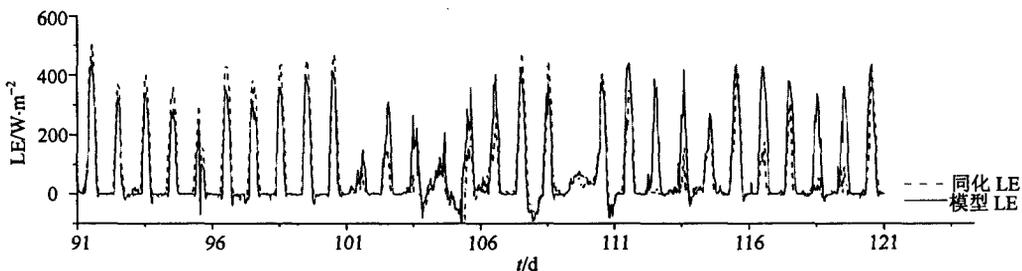
LST与地面实测LST的RMSE为4.85 K, 经过同化后LST的RMSE则下降到1.57 K。证明在同化过程中对于LST估计精度的改善较为明显。因此, 伴随在整个同化的过程中, 改善的LST将使模型最终估计的地表水热通量的精度同时得到提高。



a. 地表温度同化结果



b. 显热通量同化结果



c. 潜热通量同化结果

图6 山东禹城站点地表温度、地表水热通量的同化结果

另一方面, 由于试验中缺少对地面观测通量数据的支持, 为了验证数据同化后的通量结果, 利用

MODIS的数据产品进行辅助验证。将同化后得到的水热通量中的潜热通量与MODIS的ET产品

(MOD16A2)进行对比分析。MODIS的ET产品分为8天和月的ET/LE的全球产品,分辨率为1 km×1 km。验证试验选择MODIS对应禹城站点的8天ET产品与陆面模型独自模拟的ET(未同化)和同化系统估算的ET进行对比验证,验证试验将分别从ET变化趋势和RMSE的结果两方面做进一步的判断与分析。

MOD16的ET/LE产品采用8天累加的ET值。文中的数据均为30 mins,因此需要对数据进行时间尺

度上的匹配。根据MODIS产品的定义,本文将模型模拟的ET与同化结果的ET,在8天内的数据累加生成8天的数据产品(mm/8 d),与MOD16产品对比结果。

图7所示为2009年第89天~177天(4~6月)间的8天的ET结果随时间的变化的趋势情况,图中的3个方法得到的ET的总体变化趋势相同。由图中可以明显看出,模型模拟得到的ET值总体偏高(与MOD16的ET之间的RMSE=4.18 mm)。

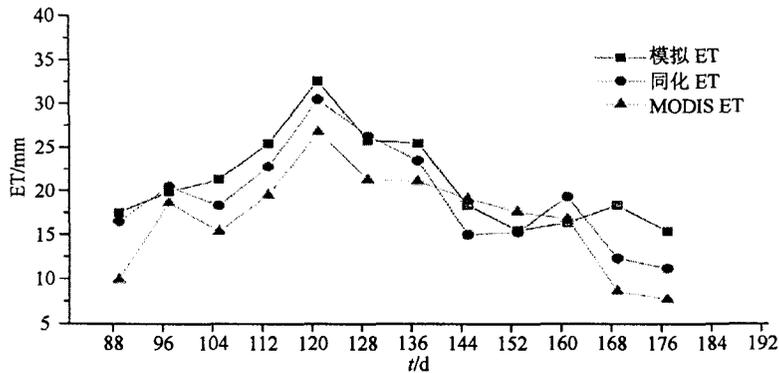


图7 同化后ET结果对比图

而经过同化后的ET整体趋势与模型相比略低,但ET值仍高于MODIS的ET产品(与MOD16的ET之间的RMSE=2.99 mm)。综合LST与ET结果的RMSE对比试验,认为经过数据同化后的模型模拟结果(LST、H和LE)均得到了有效改善。因此,本文构建的利用地表温度改进地表水热通量估计数据同化系统具有可应用性与适用性。

## 4 结论

本文利用山东禹城试验站点数据通过EnKF同化方法,改进模型地表水热通量估算精度的尝试是成功的,具有较高的可重复性与稳定性,并具备一定的适用性。在EnKF的数据同化系统中,探讨了系统中陆面模型误差的设定问题,认为一定时间内模型误差的变化可用统计的方法标定,进而运用于整个数据同化过程。同时,探讨了EnKF算法中集合大小对同化结果的影响,通过对集合大小变化的误差协方差RMSE的比较分析,确定已建立同化系统的集合数目。最后,利用该同化系统同化地表温度数据,提高了模型的地表水热通量的最终估计精度,认为该同化系统方案可有效地改善陆面过程模型对地表水热通量的估计。该研究方法和结果在农田地表水热通量的估算方面具有参考价值。

感谢中国科学院地理科学与资源研究所提供山东禹城试验站观测数据。

## 参考文献

- [1] 刘绍民, 李小文, 施生锦, 等. 大尺度地表水热通量的观测、分析与应用[J]. 地球科学进展, 2010, 25(11): 1113-1127.  
LIU Shao-min, LI Xiao-wen, SHI Sheng-jin, et al. Measurement, analysis and application of surface energy and water vapor fluxes at large scale[J]. Advances in Earth Science, 2010, 25(11): 1113-1127.
- [2] 辛晓洲, 田国良, 柳钦火. 地表蒸散定量遥感的研究进展[J]. 遥感学报, 2003(3): 233-240.  
XIN Xiao-zhou, TIAN Guo-liang, LIU Qin-huo. A review of researches on remote sensing of land surface evapotranspiration[J]. Journal of Remote Sensing, 2003(3): 233-240.
- [3] LIANG S. Advances in land remote sensing, system, modeling, inversion and application[M]. New York, USA: Springer, 2008.
- [4] BONI G, ENTEKHABI D, CASTELLI F. Land data assimilation with satellite measurements for the estimation of surface energy balance components and surface control on evaporation[J]. Water Resources Research, 2001, 37(6): 1713-1722.
- [5] KUMAR P, KALEITA A L. Assimilation of near-surface temperature using extended Kalman filter[J]. Advances in Water Resources, 2003, 26(1): 79-93.
- [6] CAPARRINI F, CASTELLI F, ENTEKHABI D. Estimation of surface turbulent fluxes through assimilation of radiometric surface temperature sequences[J]. Journal of Hydrometeorology, 2004, 5(1): 145-159.
- [7] JUN Q, LIANG S L, LIU R G, et al. A weak-constraint-based data assimilation scheme for estimating

- surface turbulent fluxes[J]. *Geoscience and Remote Sensing Letters*, 2007, 4(4): 649-653.
- [8] XU T R, LIANG S L, LIU S M. Estimating turbulent fluxes through assimilation of geostationary operational environmental satellites data using ensemble Kalman filter[J]. *Journal of Geophysical Research-Atmospheres*, 2011: 116(D09): 109-125.
- [9] BATENI S M, LIANG S. Estimating surface energy fluxes using a dual-source data assimilation approach adjoined to the heat diffusion equation[J]. *Journal of Geophysical Research-Atmospheres*, 2012, 117(D17): 118-132.
- [10] EVENSEN G. Sequential data assimilation with a nonlinear quasi-geostrophic model using monte-carlo methods to forecast error statistics[J]. *Journal of Geophysical Research-Oceans*, 1994, 99(C5): 10143-10162.
- [11] EVENSEN G. Advanced data assimilation for strongly nonlinear dynamics[J]. *Monthly Weather Review*, 1997, 125(6): 1342-1354.
- [12] EVENSEN G. The ensemble Kalman filter: theoretical formulation and practical implementation[J]. *Ocean Dynamics*, 2003, 53(4): 343-367.
- [13] BURGERS G, VAN LEEUWEN P J, EVENSEN G. Analysis scheme in the ensemble Kalman filter[J]. 1998(126): 1719-1724.
- [14] BISHOP C H, TOTH Z. Ensemble transformation and adaptive observations[J]. *Journal of the Atmospheric Sciences*, 1999, 56(11): 1748-1765.
- [15] BOCQUET M. Ensemble Kalman filtering without the intrinsic need for inflation[J]. *Nonlinear Processes in Geophysics*, 2011, 18(5): 735-750.
- [16] WANG X G, BISHOP C H, JULIER S J. Which is better, an ensemble of positive-negative pairs or a centered spherical simplex ensemble?[J]. *Monthly Weather Review*, 2004, 132(7): 1590-1605.
- [17] WHITAKER J S, HAMILL T M. Ensemble data assimilation without perturbed observations[J]. *Monthly Weather Review*, 2006, 134(6): 1722-1722.
- [18] WANG X G, HAMILL T M, WHITAKER J S, et al. A comparison of the hybrid and EnSRF analysis schemes in the presence of model errors due to unresolved scales[J]. *Monthly Weather Review*, 2009, 137(10): 3219-3232.
- [19] 黄健熙, 武思杰, 刘兴权, 等. 基于遥感信息与作物模型集合卡尔曼滤波同化的区域冬小麦产量预测[J]. *农业工程学报*, 2012, 28(4): 142-148.  
HUANG Jian-xi, WU Si-jie, LIU Xing-quan, et al. Regional winter wheat yield forecasting based on assimilation of remote sensing data and crop growth model with ensemble Kalman method[J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2012, 28(4): 142-148.
- [20] CAPARRINI F, CASTELLI F, ENTEKHAB D. Variational estimation of soil and vegetation turbulent transfer and heat flux parameters from sequences of multisensor imagery[J]. *Water Resources Research*, 2004, 40(12): 5-20.
- [21] 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.  
BAI Yu-long, LI Xin, HAN Xu-jun. A review of error problems for land data assimilation systems[J]. *Advances in Earth Science*, 2011, 26(8): 795-804.

编辑 黄 莘