

基于 GLUE方法的植被界面过程模型 (VIP)的不确定性分析*

王 昆^{1,2}, 莫兴国^{1**}, 林忠辉¹, 舒 畅^{1,2}

(1. 中国科学院地理科学与资源研究所陆地水循环及地表过程重点实验室, 北京 100101

2. 中国科学院研究生院, 北京 100039)

摘要: 生态系统模型一般参数较多, 且在应用时存在时空尺度问题, 易产生不确定性。通过模型不确定性分析, 可以加深对模型结构的理解, 提高模型预报的可靠性。植被界面过程模型 (VIP) 是一个综合考虑了陆地生态系统能量收支、水文循环和碳氮等生命元素吸收转化等过程的生态-水文动力学模型。本文采用 GLUE (Generalized Likelihood Uncertainty Estimation) 方法, 以拟合度系数作为似然判据, 利用华北平原冬小麦生长季内的田间观测数据分析 VIP 模型中的作物生长、土壤水分运动以及光合速率模块中 8 个参数以及模型预报的不确定性。研究表明, 最大光合速率 V_{max} 、饱和含水量 w_{sat} 、田间持水量 w_{field} 参数为敏感性参数, 其对似然判据的影响大, 其余参数是相对不敏感参数。在置信度为 95% 水平下, 发现观测值大都接近或者包含在置信预报区域内, 说明可以通过参数校准得到很好的模型模拟效果。

关键词: GLUE 方法; VIP 模型; 预报不确定性; 敏感性参数

Uncertainty Analysis of Vegetation Interface Processes (VIP) Model Based on GLUE

WANG Kun^{1,2}, MO Xing-guo¹, LIN Zhong-hui¹, SHU Chang^{1,2}

(1. Key Laboratory of Water Cycle and Related Land Surface Processes, IGSR, Chinese Academy of Sciences

Beijing 100101, China; 2. Graduate School of Chinese Academy of Sciences, Beijing 100039)

Abstract There are many parameters in the ecosystem model. The variability of parameters at different space-time scales resulted in uncertainties. The uncertainty analysis of parameters could help to understand the structure of the model deeply and improve the reliability of the model predictions. The Vegetation Interface Processes (VIP) model is an ecohydrology dynamic model, which includes energy budget, hydrology cycle, absorption and transformation of carbon and nitrogen in the terrestrial ecosystem. In this paper, the generalized likelihood uncertainty estimation (GLUE) methodology was used to analyze the uncertainty of the parameters. In the VIP model, we chose eight parameters which came from the crop growth module, soil water dynamic process and photosynthesis module. An index of agreement was chosen to be the likelihood weight. The field data (LAI, biomass, soil water content, etc.) in the North China Plain were used. The results showed that maximum catalytic activity of Rubisco, saturated water content and field capacity were sensitive parameters, which influenced the value of likelihood weight greatly. The others were non-sensitive parameters. Almost all the observations approached were included in the confidence interval with the 95% confidence level, which indicated that better simulations could be got by calibrating the model parameters.

Key words GLUE methodology; VIP model; Prediction uncertainty; Sensitive parameters

农业生态系统模型在农业环境调控、农业管理决策、生态系统评估以及气候变化研究等领域发挥着重要作用, 并能够结合实时遥感信息进行大范围长时间

系列的作物长势监测和估产。农业生态系统模型包括土壤水分、光合作用、蒸腾作用和呼吸作用等过程, 每个模块都有若干参数, 通常参数由田间实验测量、

* 收稿日期: 2010-02-02 ** 通讯作者。E-mail: moxg@igsrr.ac.cn

基金项目: 国家“973”计划项目 (2010CB428404); 国家“863”计划项目 (2006AA10Z228); 科技部国际合作项目 (0911)

作者简介: 王昆 (1981-), 女, 山东齐河人, 博士, 主要从事生态模型、陆面过程模式及数据同化的研究。

E-mail: wangk@lreis.ac.cn

经验或者拟合估计给定。但无论经验、观测或率定获得的参数, 都存在不确定性。加上作物模拟从单点到区域研究时, 模型参数的时空尺度效应将限制农业生态系统模型的发展应用。因此, 开展对农业生态系统模型不确定性的研究十分必要。

目前模型不确定性研究方法包括概率论、泰勒级数展开、蒙特卡洛模拟、普适似然不确定性估计、贝叶斯统计等。其中普适似然不确定性估计法^[1] (以下简称 GLUE 方法) 因其概念简单和使用方便^[2] 在环境模拟模型不确定性研究中应用较广^[3-8]。GLUE 方法在模型参数不确定性分析中取得较好的研究成果。Mo and Beven^[6] 利用小麦灌溉前后的两组实验数据系列, 以碳通量和热通量为目标, 借助 GLUE 方法分析了双源与三源小麦冠层模型在预报不确定性和参数敏感性方面的差异。Franks 等^[9] 利用能量平衡中各种热通量的观测数据结合 GLUE 方法对 SVAT 模型进行了多目标参数率定。而将正统的贝叶斯方法与 GLUE 方法结合以提高水文模型不确定性的研究目前正盛行^[10-11]。GLUE 方法在不确定性分析研究中显示了其优越性^[2, 12], 但在预报不确定性及似然函数选择问题上还存在一定局限性^[4, 13-15]。Varella 等^[16] 利用 GLUE 对作物模型中的土壤参数的估算以及 He 等^[17] 对 GLUE 方法中似然函数选择的影响等研究结果, 显示了 GLUE 方法可以用于作物模型参数的率定。植被界面过程模型 (VIP) 作为一个陆地生态系统模型, 能够模拟系统能量、水分及各种生命物质的循环平衡动态过程。本文基于 GLUE 方法对 VIP 模型中 8 个重要参数进行研究, 以冬小麦生长季内的生

物量、土壤含水量等多种观测数据作为约束条件, 进行多目标模型不确定性分析, 研究结果将有助于理解 VIP 模型的结构和参数, 并对应用、改进与发展该模型具有重要的意义。

1 模型与方法

1.1 VIP 模型

植被界面过程模型 (Vegetation Interface Processes Model, 以下简称 VIP 模型) 对生态系统辐射、水分、能量和 CO₂ 传输过程进行了详细的模拟。蒸散各分量 (冠层蒸腾、截留水蒸发和土壤蒸发) 的计算基于冠层能量平衡的双源模型, 光合作用模拟采用生理生化模式, 植物生长采用植动态模式。模型的详细描述和参数确定见文献 [6, 18], 该模型已经在多个点上得到了很好的应用^[5, 6, 18-19]。VIP 模型耦合了作物干物质形成与水分消耗, 其中涉及太阳辐射、物候、气孔导度、光合速率、呼吸作用 (生长呼吸与维持呼吸)、作物生物量动态变化及土壤水分运动等过程。

选择关键过程中作物生长、土壤水分变化和光合作用中的部分参数进行不确定性分析, 其中包括三类: 第一类, 作物生长过程, SLA_{min} 与 SLA_{max}; 第二类, 土壤水分运动, 这里主要是土壤物理属性的关键参数, kw_{sat}, w_{csat}, ww_{ilt}, w_{cfield}; 第三类, 光合作用模块, V_{max}, Am_{coef}。利用 VIP 模型模拟小麦叶面积指数、生物量、土壤水分以及蒸散的日动态变化。研究数据来源于河北栾城农业生态系统试验站 (37°50'N, 114°40'E, 海拔 50.1m)。表 1 为 VIP 模型主要参数的物理含义, 模型参数的取值范围主要根据实验观测以及文献资料确定。

表 1 VIP 模型主要参数及取值范围

| 模块名称 | 参数名称 | 物理意义 | 取值范围 | 校准时取值 | 单位 |
|--------|---------------------|----------|-------------|-------|---------------------------------------|
| 作物生长过程 | SLA _{min} | 最小比叶面积 | 0.005-0.025 | 0.016 | m ² ·g ⁻¹ |
| | SLA _{max} | 最大比叶面积 | 0.03-0.08 | 0.057 | m ² ·g ⁻¹ |
| 土壤水分运动 | kw _{sat} | 饱和导水率 | 0.01-40.00 | 34.10 | 10 ⁻⁴ cm·s ⁻¹ |
| | w _{csat} | 饱和含水量 | 0.35-0.5 | 0.435 | cm ³ ·cm ⁻³ |
| | ww _{ilt} | 萎蔫含水量 | 0.05-0.15 | 0.114 | cm ³ ·cm ⁻³ |
| | w _{cfield} | 田间持水量 | 0.16-0.34 | 0.195 | cm ³ ·cm ⁻³ |
| 光合作用模块 | V _{max} | 最大羧化速率 | 30-120 | 75 | μmol·m ⁻² ·s ⁻¹ |
| | Am _{coef} | 气孔导度经验系数 | 5-30 | 11 | 无量纲 |

1.2 GLUE 方法

GLUE 方法主要是针对模型预报中存在的确定性, 特别是模型参数的等效性问题, 即多组不同取值的参数组合驱动下模型模拟效果相似的现象。目前该方法已非常成熟, 基于蒙特卡洛方法对模型参数值进行随机采样组合来驱动模型运行, 依据事先选定

的似然目标函数计算似然值。目标函数反映模型模拟结果与观测值之间的拟合程度。GLUE 方法通常认为设定一个似然临界值来标定模型的行为特征, 令低于该临界值的参数组似然值为零, 再将其他参数组合重新归一化, 按照似然值大小得到一定置信度下模型预报的不确定性范围。具体的计算过程见文献

[1]。本文随机均匀采样 10000 组 (参数取值范围见表 1) 作为 VIP 模型参数组, 以冬小麦叶面积指数、地上茎、叶、穗的生物量及地上总生物量、第三、四层土壤体积含水量和逐日的蒸散量等变量的观测作为约束, 选择 GLUE 方法应用中较常见的拟合度系数 $R^{2[20]}$ 为似然判据, 每一个观测变量对应着一个目标变量。

拟合度系数

$$R^2 = 1 - \frac{\sum_{i=1}^n (X_i - \hat{X}_i)^2}{\sum_{i=1}^n (|\hat{X}_i - \bar{X}| + |X_i - \bar{X}|)^2} \quad (1)$$

其中, X_i 、 \hat{X}_i 分别表示某一目标变量在第 i 时刻的观测值与模拟值, \bar{X} 是该变量所有观测的均值, n 是观测数据的个数。

针对不同的目标变量, 首先根据选定的临界值, 对大于临界值的参数组合予以保留, 并将其似然判据数值归一化, 归一化的方法是对每一个目标变量, 每一个参数组合对应的 R^2 与所有参数组合计算的 R^2 之和求比值, 然后根据贝叶斯理论综合考虑得到总似然值。

$$L = \frac{L_1 L_2 \dots L_n}{C} \quad (2)$$

其中, L_1 、 L_2 、 \dots 、 L_n 分别表示各目标变量的归一化后的似然值, L 表示总似然值, C 表示归一化系数, 这里取 1, 表示各变量的权重是相同的。

1.3 应用区域与实验简介

田间试验数据收集于 2006 年 10 月 - 2007 年 6 月在中国科学院栾城农业生态试验站。该站属于暖温带半湿润、半干旱季风气候, 多年平均降水量 48.3mm。试验中的观测项目主要有: ①冬小麦生育期调查与土壤水分动态监测, 包括冬小麦生物量、叶面积、物候观测与土壤含水量等, 每隔 10d 观测 1 次; 其中土壤含水量采用土钻法, 叶面积指数和生物量取样后采用 LB3100 测量叶面积, 各器官生物量先在 105°C 下杀青 2h 再在 80°C 下烘干 20h, 用百分之一电子天平测量干物重; ②气象资料, 来自河北石家庄气象站 (38°2'N, 114°25'E, 海拔 81m) 的观测数据。

2 结果与分析

2.1 有效参数组合的临界值选择

对于拟合度系数 R^2 , 这里 R^2 考虑的是总似然值大小, 采用下式计算 VIP 模型的模拟效率

$$E = \frac{N_{R^2 > m}}{N_{total}} \times 100\% \quad (3)$$

式中, E 表示模拟效率系数 (%); $N_{R^2 > m}$ 表示满足 $R^2 > m$ 的参数组数目; N_{total} 表示用于 GLUE 方法的参数组总数目; m 为 R^2 的临界取值。不同临界值对应的模拟差异见图 1。

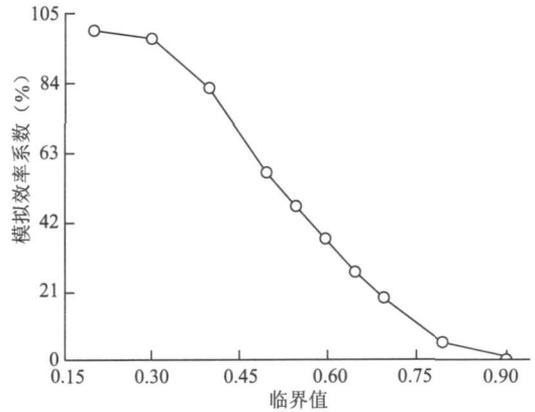


图 1 模型模拟效率系数随临界值的变化

由图 1 可见, 随着临界值 m 的增大模拟效率下降, 当 m 取 0.5 时模拟效率为 56.64%, 随着 m 取值的增大, 减少的速率增大, 这是由多目标观测变量对模型模拟的约束造成的。本研究采用的观测数据较丰富, 包括叶面积指数、生物量、蒸散和土壤水分等变量。根据模拟效率大小, 即参数组合的有效性情况, 选定似然判据可接受的临界值是 0.5 此时模型模拟效率保证在 50% 以上。

2.2 参数的不确定性分析

以 0.5 为临界值, 得到参数 - 似然值的散点分布图 (图 2 和图 3), 以此分析参数的选取与似然值之间的关系。根据结果将参数分为两类。第一类, 参数 - 似然值散点图分布均匀, 无变化趋势。这类参数有 $Am\ coef$ 、 SLA_{min} 、 SLA_{max} 、 $kw\ sat$ 、 $ww\ ill$ 它们的变化对似然判据 R^2 影响小, 属于不敏感参数, 如图 2。

第二类为敏感参数, 其似然散点图有明显的变化趋势 (图 3)。这类参数有 V_{max} 、 $w\ c\ f\ i\ e\ l\ l$ 属于相对敏感的参数。其中, V_{max} 峰值出现在 $50\ \mu\ mol \cdot m^{-2} \cdot s^{-1}$ 附近, V_{max} 表示单位面积叶面中 Rubisco 的最大催化能力, 主要取决于该羧化酶的浓度, 而酶的浓度与叶片氮含量有较强相关性^[21], 显然, 叶片氮含量不是固定不变的, 随植物与土壤的氮循环过程而变化, 受气候和环境各因素的影响, 所以最大羧化速率在植物生长季内不同阶段是动态的, V_{max} 峰值与诸多最大羧化速率的均值必定是相近的; $w\ c\ f\ i\ e\ l\ l$ 峰值出现在 $0.16 \sim 0.23\ cm^3 \cdot cm^{-3}$, 且似然值峰值自左向右呈迅速下降趋势; $w\ c\ s\ a\ t$ 与 $w\ c\ f\ i\ e\ l\ l$ 呈反方向的倾斜分布, 在

模型中, 土壤饱和含水量决定着降水和灌溉的入渗分配, 田间持水量控制着冬小麦的灌溉时间与灌溉量, 不同的土壤分层中, 土壤饱和含水量与田间持水量因

各层土壤质地的差异而不同, 加之根系分布也有差别, 综合因素作用下模型选定单一数值来表达导致这两个参数较为敏感, 如图 3。

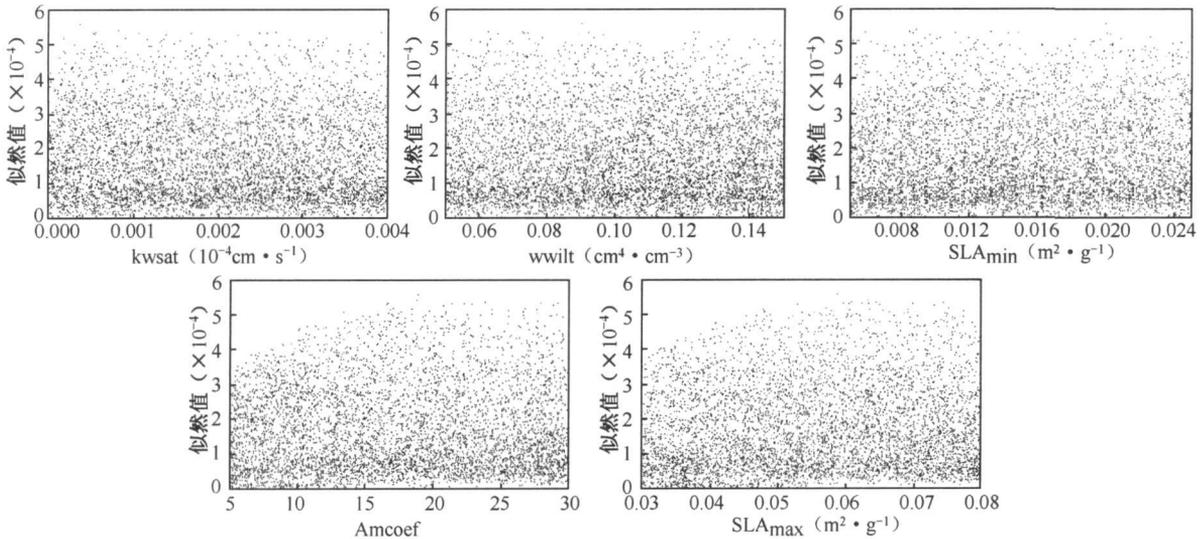


图 2 不敏感参数

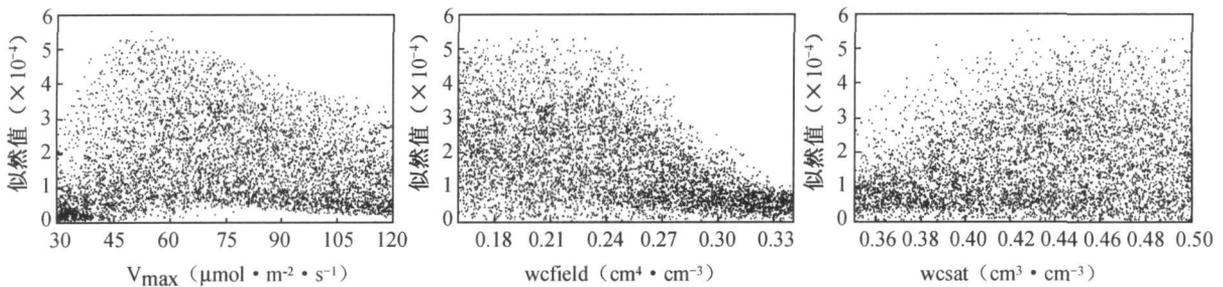


图 3 敏感性参数

2.3 参数组的等效性分析

从图 2 和图 3 中可以看到, 在很大的参数空间内, 模型的模拟效果是可以接受的, 模拟结果中出现大量“异参等效”现象即参数等效性, 依据舒畅等^[22]采用的总和标准化计算方法得到“等效”参数组合归一化数值对比图 (图 4), 选中的三组参数对应的似然值是 0.000514 (最大似然值是 0.000557)。从图 4 中看出, 同一个参数在各组合中取值存在差异, 其中 V_{max} , w_{csat} , w_{field} 数值变化比较小, 其他参数如 k_{wsat} , SLA_{min} , $wwilt$ 的变化范围则相对较大, 说明模型中各类参数的不确定性不同。三组参数组合中各参数原始取值见表 2 (单位与表 1 一致), 其中参数的数值大小与模型校准时取值 (表 1) 不尽相同, 但大部分数值较接近。而从图 5 的例子中可以看到, 对于单一的参数值, 所对应的似然值也不是唯一的, 存在较宽的变幅, 这依赖于参数组合中其他参数的取值, 证实了模型模拟结果的优劣取决于参数组合, 而并非单个参数。可见, 参数组合选择存在极大的不确定性, 是模型应用参数校准中不可避免的。

表 2 三组“等效性”参数组

| | 参数名称 | | | | | | | |
|-----|------------|-----------|-------------|-------------|------------|-------------|---------|----------|
| | k_{wsat} | V_{max} | SLA_{max} | SLA_{min} | w_{csat} | w_{field} | $wwilt$ | $Amcoef$ |
| 组 1 | 11.57 | 58.53 | 0.0478 | 0.0094 | 0.4706 | 0.1821 | 0.0748 | 24.3315 |
| 组 2 | 18.79 | 49.82 | 0.0691 | 0.0182 | 0.4632 | 0.1795 | 0.0634 | 16.4752 |
| 组 3 | 13.46 | 52.74 | 0.0635 | 0.0209 | 0.4489 | 0.1924 | 0.124 | 16.8579 |

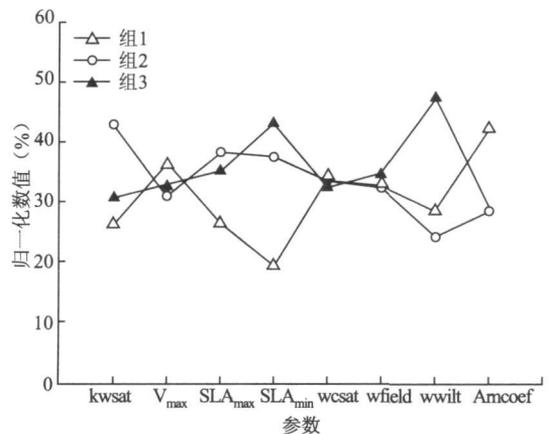


图 4 等效参数组合标准化数值对比图

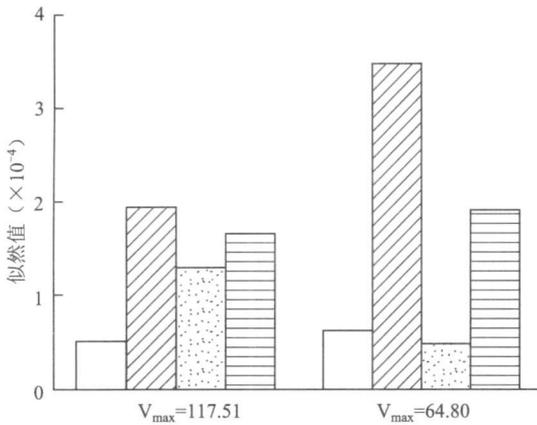


图 5 同一参数值对应似然值对比图

2.4 模型预测的不确定性分析

根据总似然值的大小, 挑选出较大的 500 组参数, 对所选的参数进行相关性分析, 除 V_{max} 与 SLA_{max} 之间的相关系数是 -0.62 外, 其他参数相关性很小, 相关系数都在 0.1 之下。接着, 对模型单位步长上的各状态变量, 如叶面积指数、地上生物量、茎和叶生物量以及土壤含水量的预报值做不确定性分析, 分位数选定 95% , 得到置信度 95% 下模型预报的不确定性范围 (图 6 与图 7)。从图 6 和图 7 中可以看出, 各变量的观测值几乎全部落在不确定性预报上下限内, 只有极少观测值落在 95% 的置信区间之外, 这可能是多种因素共同作用的结果, 例如参数的先验分布和模型本身的影响。

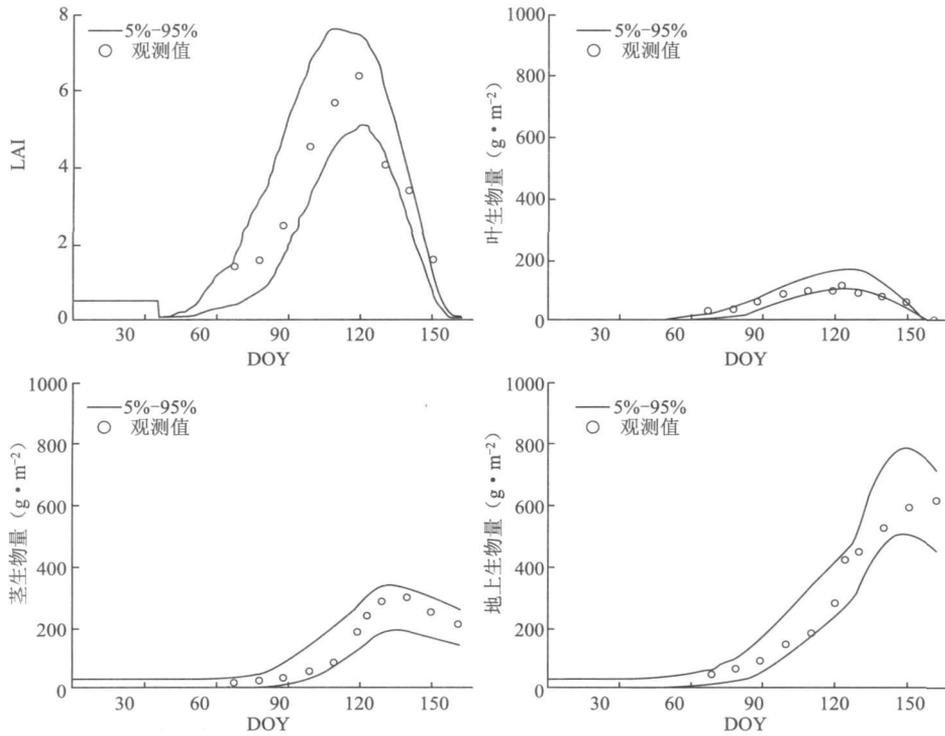


图 6 叶面积指数与生物量预报的边界与观测值对比图

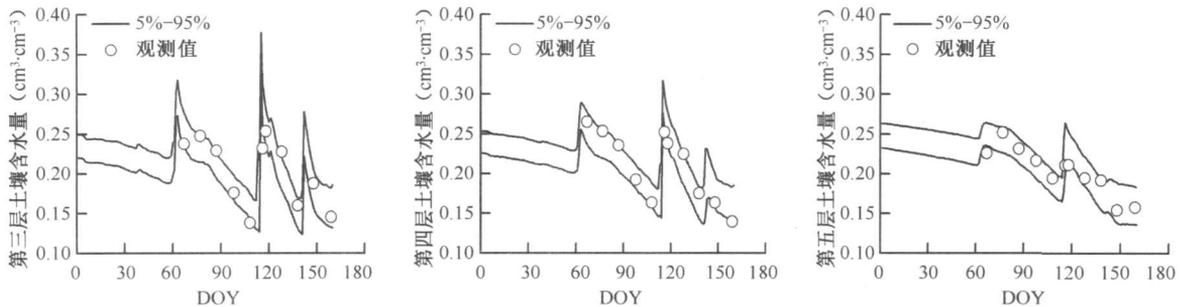


图 7 土壤含水量预报的边界与观测值对比图

3 结论与讨论

(1) 本研究使用的观测数据较丰富, 包括叶面积指数、蒸散发、各器官生物量以及两层的土壤含水量,

以似然函数临界值为 0.5 得到对应的模型模拟效率是 56.64% , 结果说明采用的参数组合能够较好地模拟出真实的动态过程。

(2) 对模型参数不确定性分析的结果显示, 最大光合速率 V_{max} 、饱和含水量 w_{csat} 以及田间持水量 w_{cfield} 属于敏感参数, 在模型应用校准时应给予重点考虑, 其中 V_{max} 出现峰值, 在 $50 \mu\text{mol} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$ 附近; w_{cfield} 峰值在 $0.16 \sim 0.23 \text{cm}^3 \cdot \text{cm}^{-3}$ 。其他选定的五个参数, 饱和导水率 k_{wsat} 、萎蔫含水量 w_{wilt} 最大、最小比叶面积 SLA_{max} 、 SLA_{min} 以及光合计算的参数 Am_{coef} 不敏感。说明各参数的敏感性不同, 分两类: 敏感性参数与非敏感性参数。

(3) 发现参数等效性现象, 即几组不同参数取值的组合对应的似然函数值相同, 而单一参数相同取值的不同参数组合对应的似然函数值存在明显差异, 这说明模型参数组合决定了模型的模拟效果, 而非单个参数。

(4) 模型预报不确定性分析的结果中, 生物量、叶面积指数以及土壤含水量等变量的观测值几乎全部位于置信度为 95% 的置信区间内, 证明了 VIP 模型在这些变量方面有较好的模拟能力。

本研究结果为掌握 VIP 模型结构和参数提供了参考, 有助于该模型的发展应用。

参考文献:

- [1] Beven K, Binley A. The future of distributed model calibration and uncertainty prediction [J]. *Hydrological Processes*, 1992, 6(6): 279-298
- [2] Blason R S, Madsen H, Rosbjerg D. Uncertainty assessment of integrated distributed hydrological models using GLUE with Markov chain Monte Carlo sampling [J]. *Journal of Hydrology*, 2008, 353: 18-32
- [3] Romanowicz R J, Beven K J, Tawn J. Bayesian calibration of flood inundation models [A]. *Flood plain processes* [C]. Chester Anderson and D E Walling 1996: 333-360
- [4] Christensen S. A synthetic groundwater modeling study of the accuracy of GLUE uncertainty intervals [J]. *Nordic Hydrology*, 2004, 35(1): 45-59
- [5] Mo X, Pappenberger F, Beven K, et al. Parameter conditioning and prediction uncertainties of the LISFLOOD-WB distributed hydrological model [J]. *Hydrological Sciences*, 2006, 51(1): 45-65
- [6] Mo X, Beven K. Multi-objective parameter conditioning of a three-source wheat canopy model [J]. *Agricultural and Forest Meteorology*, 2004, 122: 39-63
- [7] Trnblay M, Wallach D. Comparison of parameter estimation methods for crop models [J]. *Agronomy*, 2004, 24: 351-365
- [8] Blason R S, Vrugt A, Madsen H, et al. Generalized likelihood uncertainty estimation (GLUE) using adaptive Markov Chain Monte Carlo sampling [J]. *Advances in Water Resources*, 2008, 31: 630-648
- [9] Franks SW, Beven K J, Gash H C. Multi-objective conditioning of a simple SVAT model [J]. *Hydrology and Earth System Sciences*, 1999, 3(4): 477-489
- [10] Vrugt J A, Braak C J F, Gupta H V, et al. Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modeling? [J]. *Stochastic Environmental Research and Risk Assessment*, 2009, 23: 1011-1026
- [11] Jin X, Xu C Y, Zhang Q, et al. Parameter and modeling uncertainty simulated by GLUE and a formal Bayesian method for a conceptual hydrological model [J]. *Journal of Hydrology*, 2010, 383: 147-155
- [12] Montanari A. Large sample behaviors of the generalized likelihood uncertainty estimation (GLUE) in assessing the uncertainty of rainfall runoff simulations [J]. *Water Resources Research*, 2005, 41, W08406, doi: 10.1029/2004WR003826
- [13] Mantovan P, Todini E. Hydrological forecasting uncertainty assessment: incoherence of the GLUE methodology [J]. *Journal of Hydrology*, 2006, 330(1-2): 368-381
- [14] Mantovan P, Todini E, Martina M L V. Reply to comment by Keith Beven, Paul Smith and Jim Freer on "Hydrological forecasting uncertainty assessment: incoherence of the GLUE methodology" [J]. *Journal of Hydrology*, 2007, 338(3-4): 319-324
- [15] Stedinger J R, Vogel R M, Lee S U, et al. Appraisal of the generalized likelihood uncertainty estimation (GLUE) method [J]. *Water Resources Research*, 2008, 44, W00B06, doi: 10.1029/2008WR006822
- [16] Varela H, Guerif M, Buis S. Global sensitivity analysis measures the quality of parameter estimation: the case of soil parameters and a crop model [J]. *Environmental Modelling & Software*, 2010, 25: 310-319
- [17] He J, Jones J W, Graham W D, et al. Influence of likelihood function choice for estimating crop model parameters using the generalized likelihood uncertainty estimation method [J]. *Agricultural Systems*, 2010, 103: 256-264
- [18] 王昆, 莫兴国, 林忠辉, 等. 植被界面过程 (VIP) 模型的改进与验证 [J]. *生态学杂志*, 2010, 29(2): 387-394
- [19] Mo X, Liu S, Lin Z, et al. Prediction of crop yield, water consumption and water use efficiency with a SVAT-crop growth model using remotely sensed data on the North China Plain [J]. *Ecological Modelling*, 2005, 183: 301-322
- [20] Willmott C J. On the validation of models [J]. *Physical Geography*, 1981, 2(2): 184-194
- [21] Thompson W A, Wheeler A M. Photosynthesis by mature needles of field-grown *Pinus radiata* [J]. *Forest Ecology and Management*, 1992, 52(1-4): 225-242
- [22] 舒畅, 刘苏峡, 莫兴国, 等. 新安江模型参数的不确定性研究 [J]. *地理研究*, 2008, 27(2): 343-352