

基于转移熵的长江流域土壤湿度对降水反馈研究

娄巍, 刘攀*, 程磊, 邹铠杰, 夏倩

武汉大学水资源与水电工程科学国家重点实验室, 湖北 武汉
Email: *liupan@whu.edu.cn

收稿日期: 2021年1月2日; 录用日期: 2021年2月1日; 发布日期: 2021年2月22日

摘要

陆气耦合过程复杂, 研究土壤湿度对降水的影响机制, 对气候预测和天气预报具有重要意义。目前土壤湿度对降水的反馈机制尚不明确, 采用物理模型存在较大的不确定性, 而通过数据直接反映非线性统计相关性的转移熵, 是一种可解释水文系统的新范式, 为研究此类问题提供了可能。本文提出并验证了归一化转移熵, 使量纲不同的耦合强度具备可比性; 使用长江流域2002~2018年168个气象站点观测的降水和蒸散发数据、全球陆面参数数据LPDR V2.0中的土壤湿度数据、以及MODIS中的植被指数NDVI数据, 采用显著滞时、不同滞时的相对预测度和归一化信息熵3个信息指标识别土壤湿度对降水的反馈特征, 并应用偏相关方法验证了长江流域土壤湿度对降水正反馈的合理性。结果表明: 1) 土壤湿度对降水的平均显著滞时为4.3 d, 与降水对土壤湿度的1.8 d相比滞时更长, 说明土壤湿度对降水的反馈过程存在滞时, 即土壤湿度响应前期降水后, 延迟较长时段影响降水过程; 2) 降水对土壤湿度的归一化转移熵为0.51, 与土壤湿度对降水的0.13相比耦合强度更高, 说明土壤湿度对降水的反馈显著弱于降水对土壤湿度的影响; 3) 金沙江下游和洞庭湖、鄱阳湖、太湖水系等近湖区域由于水量交换频繁, 土壤湿度对降水反馈较快且强度更高。

关键词

土壤湿度对降水反馈, 长江流域, 转移熵

Soil Moisture Feedbacks on Precipitation in the Yangtze River Catchment Based on Transfer Entropy

Wei Lou, Pan Liu*, Lei Cheng, Kaijie Zou, Qian Xia

State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan Hubei
Email: *liupan@whu.edu.cn

Received: Jan. 2nd, 2021; accepted: Feb. 1st, 2021; published: Feb. 22nd, 2021

作者简介: 刘攀, 教授、博导, 研究方向: 水文水资源。
*通讯作者。

文章引用: 娄巍, 刘攀, 程磊, 邹铠杰, 夏倩. 基于转移熵的长江流域土壤湿度对降水反馈研究[J]. 水资源研究, 2021, 10(1): 21-32. DOI: 10.12677/jwrr.2021.101003

Abstract

The land surface and atmosphere interact as a complexly linked system. The soil moisture-precipitation feedback mechanism is of great significance to weather forecast and climate prediction; however, it is uncertain in physical-based numerical simulation. Transfer entropy, which reflects non-linear statistical correlation directly through data, has been considered as a new paradigm to explain hydrological system and provides possibility to study the land-atmosphere coupling. Normalized transfer entropy was proposed in this study, and was verified to measure comparable coupling strength. Observation data on 168 weather stations during 2002~2018, MODIS-NDVI (MOD13A2) and soil moisture data from the global land parameters data record LPDR V2.0 were used. Characteristics of the feedback were explored by three information indexes. A positive soil moisture-precipitation feedback was also valid by the partial correlation. The results showed: 1) the average significant lag time of soil moisture-precipitation coupling was 4.3 d, which was longer than 1.8 d of precipitation-soil moisture coupling. It indicated that there was a delay effect during the feedback process. The soil moisture in response to precipitation affected the precipitation process after a longer period of time; 2) the normalized transfer entropy of precipitation-soil moisture coupling was 0.51, which was stronger than 0.13 of soil moisture-precipitation coupling; and 3) due to the frequent water exchange in the lower reaches of the Jinsha River, the Dongting Lake, the Poyang Lake and the Taihu Lake systems, soil moisture had a faster and stronger feedback on precipitation.

Keywords

Soil Moisture-Precipitation Feedback, The Yangtze River, Transfer Entropy

Copyright © 2021 by author(s) and Wuhan University.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

复杂的陆气反馈过程中，土壤湿度对大气降水的反馈机制是核心问题之一。一方面，土壤湿度作为陆面水文过程的重要状态变量，直接被大气降水驱动；另一方面，土壤湿度通过向大气补充可降水汽的直接方式，或改变大气条件的间接方式，从而触发降水事件发生。识别土壤湿度对降水的反馈关系特征，对搭建灵活易用的陆气耦合模型、提高水文模拟和气象预报精度有重要意义[1]。

土壤湿度对降水反馈在机制上存在较多不确定性，指标较多且同质[2] [3]，即使在同一研究区域，研究方法、数据尺度和评价指标不同，也会导致争议性结论。如伊利诺伊州，Findell 和 Eltahir [4]基于插值的伊利诺伊气象站网观测数据，以滞后相关性统计方法识别出春末夏初存在土壤湿度对降水的反馈作用；而 Salvucci、Saleem 和 Kaufmann [5]认为插值存在人为正向偏差，并使用格兰杰因果检验证实伊利诺伊州土壤湿度对降水的反馈作用并不显著。

长江流域位于陆-气-海系统复杂耦合的东亚季风区，有 58% 的降水水汽来源于流域内陆地蒸散发，相较于海面，陆地土壤湿度为流域降水输送了更多的气[6]。全球陆气耦合试验 GLACE 也表明，东亚季风区是土壤湿度对降水反馈的热点区域之一[7]。K. C. Chow [8]通过区域气候模式(RCM)做敏感性试验，证明青藏高原春季土壤湿度的增加会导致长江流域夏季降水量的增加；而王林、王磊等[9]用天气预测模型(WRF)进行数值模拟得到相反结论，春季高原土壤湿度偏湿，长江中下游地区初夏降水量大范围减小。

数值模拟和统计分析是研究土壤湿度对降水反馈的两种常用手段，数值模拟模型具有较大的不确定性且需庞大的建模工作，而统计分析中的指标难以度量非线性关系[10]。基于香农熵的信息论可检验非线性因果，近年被认为是一种可解释水文系统的新范式[11]。如用互信息确定水文气候变量的时空变异性[12]、时间信息划分网络构建生态水文网络去推断系统的响应行为[13] [14]、相对熵分析响应行为变化的趋势性[15]、特定互信息表示降水对径流的预测程度[16]等。转移熵(TE)由 Schreiber 和 Thomas [17]提出，能够很好地分析两个变量的耦合强度和不对称驱动响应关系，但未应用于陆气耦合的研究。此外，TE 仅以信息流大小衡量耦合强度，忽略了自信息量差异。

本文通过归一化的方式改进转移熵，统一量纲后使耦合强度具有可比性；基于长江流域实测气象站数据以及卫星反演的长序列土壤湿度数据，采用已有和改进的信息指标，探索从信息论的视角分析土壤湿度对降水的反馈特征及空间分布，为陆气耦合研究提供新途径。

2. 研究方法

先采用偏相关排除混杂因素影响，初步验证相互间的关系，再以信息论为基础的特征指标，分析得出变量关系的详细时空变化特征。

2.1. 偏相关分析

当变量 y 同时与变量 x_1 、 x_2 相关，可采用偏相关法剔除变量 x_1 的影响，只分析 y 和 x_2 的相关关系，被剔除的变量为控制变量。在包含 3 个自变量 x_1 、 x_2 、 x_3 和 1 个因变量 y 的系统中，偏相关系数的计算公式如下：

$$r_{x_1y} = \frac{\sum_{i=1}^n [(x_i - \bar{x})(y_i - \bar{y})]}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (1)$$

$$r_{x_1y \cdot x_2} = \frac{r_{x_1y} - r_{x_1x_2} r_{x_2y}}{\sqrt{(1 - r_{x_1x_2}^2)(1 - r_{x_2y}^2)}} \quad (2)$$

$$r_{x_1y \cdot x_2x_3} = \frac{r_{x_1y \cdot x_2} - r_{x_1x_3 \cdot x_2} r_{y \cdot x_3 \cdot x_2}}{\sqrt{(1 - r_{x_1x_3 \cdot x_2}^2)(1 - r_{y \cdot x_3 \cdot x_2}^2)}} \quad (3)$$

式中： x_i 、 y_i ($i=1,2,3,\dots,n$) 是第 i 个数据中的 x 、 y 值， \bar{x} 、 \bar{y} 是 x_1, x_2, \dots, x_n 和 y_1, y_2, \dots, y_n 的均值； r 表示变量间的相关关系，如 r_{x_1y} 表示变量 y 和 x_1 的 Pearson 相关系数； $r_{x_1y \cdot x_2}$ 表示在控制变量 x_2 时，变量 y 和 x_1 的一阶偏相关系数； $r_{x_1y \cdot x_2x_3}$ 是控制变量为 x_2 、 x_3 时，变量 y 和 x_1 的二阶偏相关系数。

2.2. 转移熵方法

与模型模拟方法相比，信息论不受物理过程参数化和数据时空分辨率的影响；与现有统计方法相比，转移熵采用非参数估计的概率密度来描述，而不仅仅是数值的大小[18]，可不带假设条件地衡量传统相关性指标等不能反映的非线性关系，从信息论的视角认识水文过程[13] [14]。

2.2.1. 数据预处理

1) 归一化

每个时间序列变量 x 按下式将标准值映射至(0, 1):

$$x_{\text{Norm}} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (4)$$

式中： x_{\max} 和 x_{\min} 是样本数据的最大值和最小值， x 和 x_{Norm} 是标准值和映射结果值。

2) 零值效应

若变量 x 、 y 有大量零值，二维概率密度函数按下式计算：

$$\begin{cases} p(x_i, y_j) = p & \text{if } i = 1, j = 1 \\ p(x_i, y_j) = p_{x0} f_{x0}(x_i) h_i & \text{if } i > 1, j = 1 \\ p(x_i, y_j) = p_{0y} f_{0y}(y_j) h_j & \text{if } i = 1, j > 1 \\ p(x_i, y_j) = p_{xy} f_{xy}(x_i, y_j) h_i h_j & \text{if } i > 1, j > 1 \end{cases} \quad (5)$$

式中： $f_{x0}(x_i)$ 和 $f_{0y}(y_j)$ 是边缘概率密度函数， p_{xx} 是直接计数得到的各类数据点所占的比例。

3) 巴特沃斯高通滤波

对于变量 x ，巴特沃斯高通滤波可以保留短期波动以及忽略日或季节周期。

2.2.2. 信息熵和互信息

在信息论中，信息熵可定量描述不确定性，信息熵越大，表明不确定性越大。互信息是两个或多个变量共享的信息量，相关性越大，则互信息量越大。

$$H(x) = -\sum p(x_i) \log_2 p(x_i) \quad (6)$$

$$H(x|y) = -\sum \sum p_{x|y}(x_i|y_j) p_{x|y}(x_i|y_j) \log_2 p_{x|y}(x_i|y_j) dx \quad (7)$$

式中： $H(x)$ 是离散型随机变量 x 的信息熵， $H(x|y)$ 是随机变量 x 在得到变量 y 提供的信息后所剩余的不确定性，即条件互信息熵； x 的可能取值 x_i 的概率为 $p(x_i)$ ， x_i 关于 y_i 的条件分布函数为 $p_{x|y}(x_i|y_i)$ 。

互信息 $I(x; y)$ 为信息熵与条件互信息熵的差值。同理，偏互信息 $I(x; y|z)$ 定义为除去变量 z 的影响后， y 对 x 提供的信息量。

$$I(x; y) = H(x) - H(x|y) \quad (8)$$

$$I(x; y|z) = H(x|z) - H(x|y, z) \quad (9)$$

信息熵计算的关键是估计联合概率密度，可用非参数方法求得，包括直方图法、核函数法和平均直方图法等。

2.2.3. 特征指标

在过程网络中，多个变量通过信息流相互关联。根据信息分离理论[14]， U 指仅由一个源变量单独提供目标变量的信息； R 指多个源变量提供目标变量的重复信息源变量共同提供目标变量的信息。设每条信息流始于源变量 x ，终于目标变量 y ，本文用到以下特征指标：

1) 显著滞时

即 x 对 y 发生最大影响的时间 τ_0 ，于 $I_{\max}(\tau = \tau_0) = \max_{\tau=1}^t I(x(t-\tau); y(t))$ 取得，其中 τ 是滞时 ($\tau = 0, 1, \dots, t$)。

2) 不同滞时的相对预测度

将总信息量 $I_{tot, \tau}$ 按式(10)分解，滞时 1 个单位时间跟 τ 个单位时间 ($\tau = 2, 3, \dots, t$) 的 x 分别提供 y 的信息比例为 Γ 和 η ($\Gamma + \eta = 1$)，即 x 的长短滞时历史对 y 现在状态的相对预测度(式(11)和式(12)) [19]。根据分离出的信息种类 U 、 S 、 R ，可指示链接冗余度， η 值小则冗余度高。

$$I_{tot, \tau} = I(x(t-1), x(t-\tau); y(t)) = I(x(t-1); y(t)) + I(x(t-\tau); y(t)|x(t-1)) \quad (10)$$

$$\frac{I(x(t-1); x(t))}{I_{tot}} = U_{x(t-1)} + R_{x(t-1), x(t-\tau)} = \Gamma \quad (11)$$

$$\frac{I(x(t-\tau); y(t)|x(t-1))}{I_{tot}} = U_{x(t-\tau)} + S_{x(t-1), x(t-\tau)} = \eta \quad (12)$$

3) 归一化转移熵

当式(9)的 z 设置为 x 的滞时序列, 得到的偏互信息称转移熵 TE , 可指示信息转移强度, 但未考虑自信息差异。本文提出指标 TE/H , 以源变量的熵值为底进行归一化, 使不同自信息量条件下的耦合强度可比。

$$TE/H(x(t-\tau) \rightarrow y) = \frac{TE(x(t-\tau) \rightarrow y)}{H(x(t-\tau))} = \frac{I(x(t-\tau); y | y(t-\tau))}{-\sum p(x(t-\tau)) \log_2 p(x(t-\tau))} \quad (13)$$

以土壤湿度 - 降水耦合为例, 降水时间序列为 $P(t)$, 土壤湿度时间序列为 $SM(t) (t=1,2,3,\dots,n)$, 设置滞时状态为 $t-\tau (\tau=0,1,\dots,t)$ 。当 $I(P(t-\tau); SM(t))$ 取最大值, τ_{01} 代表前期降水对土壤湿度发生显著影响的滞时, 同理, 当 $I(P(t); SM(t-\tau))$ 取最大值, τ_{02} 代表前期土壤湿度对降水发生显著影响的滞时; τ 量化了仅通过了解滞时 1 个单位时间的土壤湿度 $SM(t-1)$ 的信息而减少 $P(t)$ 信息量的比例, 包含 $SM(t-1)$ 单独提供的 U 信息, 以及 $SM(t-1)$ 和 $SM(t-\tau)$ 提供的 R 信息。 η 则表示 $SM(t-\tau)$ 提供的 U 信息以及 $SM(t-1)$ 和 $SM(t-\tau)$ 共同提供的 S 信息所减少 $P(t)$ 信息量的比例, 该比例量化了较长滞时 $SM(t-\tau)$ 对 $P(t)$ 影响的相对重要程度。因此, $\eta=0$, 代表 $P(t)$ 的不确定性通过了解 $SM(t-1)$ 获得了最大程度的降低, $SM(t-\tau)$ 提供的信息必定是冗余的。相反地, $\eta=1$, 代表了解 $SM(t-1)$ 不能降低 $P(t)$ 任何不确定性, 而要通过更长滞时 $SM(t-\tau)$ 的额外信息来降低; 归一化转移熵指标 $TE(P(t-\tau)-SM(t))/H(P(t-\tau))$, 即降水显著滞时下的历史序列 $P(t-\tau)$ 向当前土壤湿度状态 $SM(t)$ 转移的信息强度; 而指标 $TE(SM(t-\tau)-P(t))/H(SM(t-\tau))$, 即土壤湿度显著滞时下的历史序列 $SM(t-\tau)$ 向当前降水状态 $P(t)$ 转移的信息强度。

3. 试验验证

为验证指标适用性, 采用蒙特卡罗试验模拟两变量 x_1 和 x_2 的互馈作用, 随机产生 1000 个样本。模型 1 中, 两变量以 $\tau=3$ 和 4 相互约束; 模型 2 中, 随机生成 x_1 再以 $\tau=3$ 约束 x_2 。设置 $f(X_i) = aX_i(1-X_i)$, 令 $a=4$ 计算指标。此外, 设定变化耦合系数 a 的 9 个情景, 测试 TE/H 的强度指示作用。

$$\text{模型 1 } x_1(t) = f(x_2(t-4)), x_2(t) = f(x_1(t-3)) \quad (14)$$

$$\text{模型 2 } x_1(t) = z \sim U(0,1), x_2(t) = f(x_1(t-3)) \quad (15)$$

试验结果如图 1, 可得出以下结论: 1) 改变变量的初始值, 识别的显著滞时接近于理论值; 2) η 值可区分变量作用类型, $\eta=1$ 指示了 x_1 对 x_2 的单向作用, 而 $\eta < 1$ 指示了双向作用; 3) TE 、 H 曲线非单调, 而 TE/H 随耦合系数 a 的增加而上升, 验证其对耦合强度的指示作用。通过归一化, 也解决了不同自信息量条件下的耦合强度相比时量纲不统一的问题。

4. 研究数据

4.1. 数据来源

长江流域(90°33'~122°25'E, 24°30'~35°45'N)总面积 180 万 km^2 , 自西向东依次横跨 19 个省、市和自治区。流域流经青藏高原、四川盆地和长江中下游平原三大阶梯地貌, 地形复杂, 受西太平洋副高季风系统影响, 降水量年内分配不均, 空间分布差异明显, 形成复杂的水文过程。

气象资料为长江流域 2002~2018 年 168 个气象站降水和蒸散发的逐日观测数据, 取自中国气象数据网 (<http://data.cma.cn>) 的中国地面气候资料日值数据集(V3.0), 站点及水系分布如图 2。

土壤湿度实测数据缺乏, 微波遥感是实现土壤湿度大范围长时间连续观测的主要手段, 不受大气影响且对地表具有一定穿透能力。先进微波扫描辐射计(AMSR-E)和第二代先进微波辐射成像仪(AMSR2)是搭乘在近年对地观测卫星 EOS-PM1 和 GCOM-W1 的传感器。Du、Kimball、Jones、Kim 和 Glassy [20] AMSR-E 和 AMSR2, 进行了相似的校正和多频亮度温度(T_b)反演, 得到新的全球陆地参数数据集(LPDR V2.0) (http://files.ntsg.umd.edu/data/LPDR_v2),

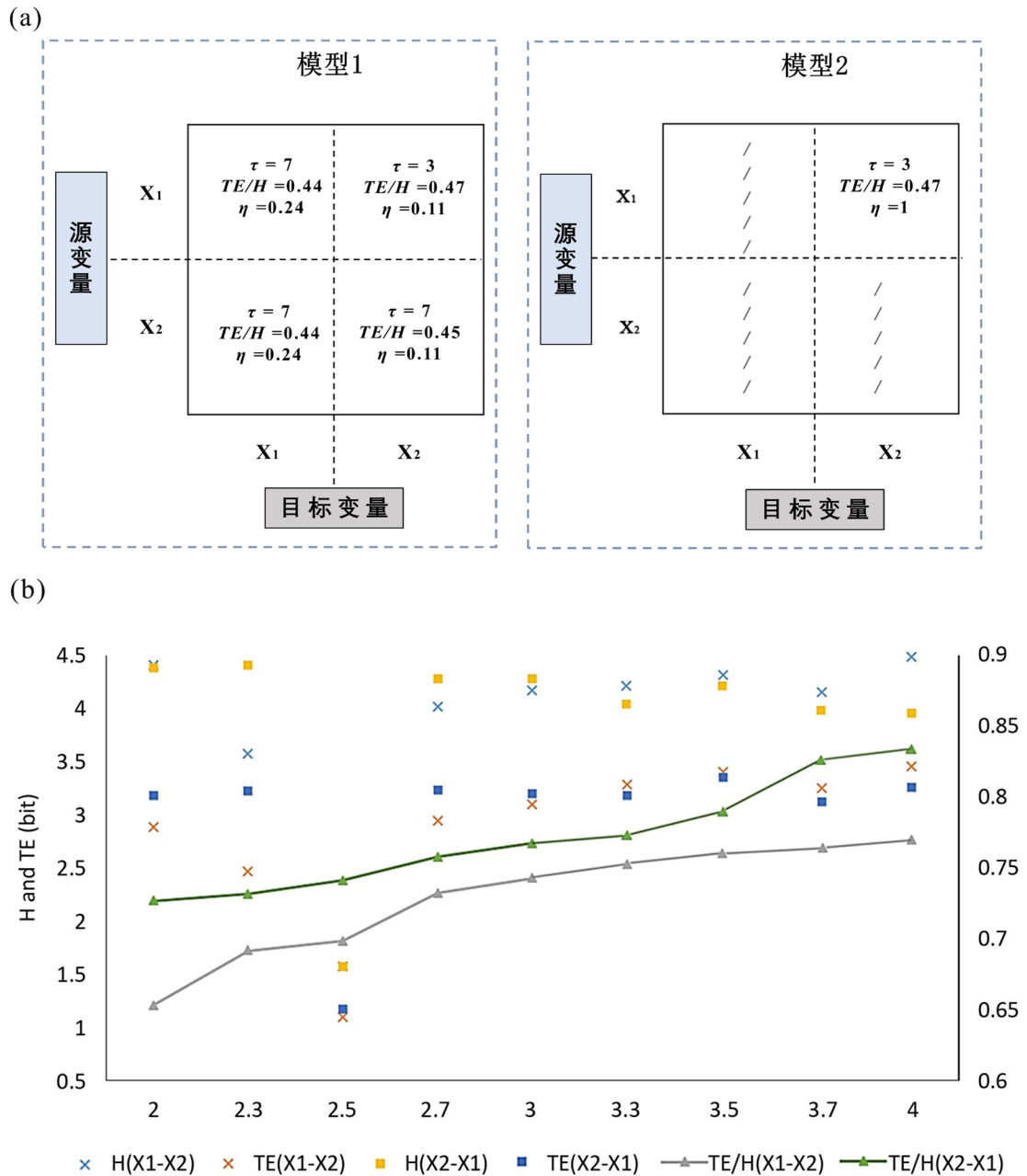


Figure 1. Results of Monte Carlo test: (a) τ , TE/H , η values of two models under different source and target variables; (b) Bidirectional H , TE , TE/H values of model 1 under the changing coupling coefficient

图 1. 蒙特卡罗试验结果：(a) 模型 1 跟 2 在不同源变量和目标变量下对应的 τ 、 TE/H 、 η 值；(b) 模型 1 在变化耦合系数下双向的 H 、 TE 、 TE/H 值

空间分辨率 0.25° 。本研究调整坐标系为 WGS-1984，并按气象站点位置，提取该数据集在点上的逐日土壤体积含水量。

为提供混杂因素，还使用了美国航天局 NASA 发布的 MODIS NDVI (ID = MODIS/006/MOD13A2) 数据，时间分辨率 16 日 (<https://ladsweb.modaps.eosdis.nasa.gov/>)。数据在谷歌地球数据云平台(Google Earth Engine)中处理，利用数据集的质量控制字段(SummaryQA)，选取所有最优观测值(SummaryQA 等于 0)的最大值作为月 NDVI。

4.2. 数据预处理

由于降水和土壤湿度在量级和变化幅度上存在较大差异, 需要根据数据特征进行预处理, 以还原它们在原始数据中所反映的相关性。本文采用以下预处理方式: 1) 将所有数据按最值归一化至区间 0~1; 2) 连续的降雨数据中有大量缺测值, 使概率密度偏移并影响信息熵计算, 因此考虑了零值效应, 以提升概率密度估算; 3) 对单个变量, 采用高通巴特沃斯滤波过滤日循环和季节循环, 仅关注短期波动来挖掘降水和土壤湿度微小变化中的反馈机制[13]。

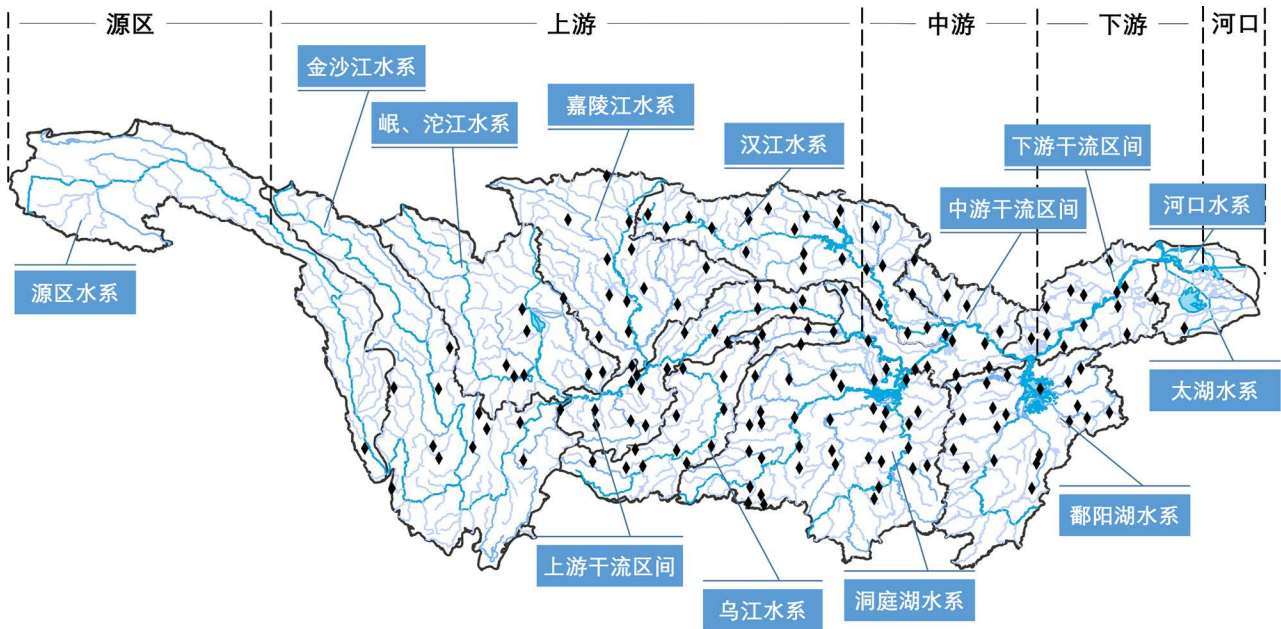


Figure 2. Distribution of river system network and 168 weather stations in the Yangtze River catchment

图 2. 长江流域水系及 168 个气象站分布

5. 结果与讨论

5.1. 基于偏相关的数据验证

常用的相关性方法可用来识别土壤湿度对降水的反馈, 但两者的相互作用过程中常混杂其他因素, 利用偏相关关系可排除干扰, 更真实地反映变量间的相关性。

蒸散发和土壤湿度相互制约, 且蒸散发与降水可能是正相关关系, 也可能是负相关关系, 使土壤湿度 - 降水机制具有较大不确定性; 在植被覆盖的土地上, 植被覆盖状况直接影响植被的蒸腾, 从而影响土壤湿度的大小。因此选取降水、土壤湿度、蒸散发、NDVI 四个变量, 并将蒸散发及 NDVI 设置为控制变量。受 NDVI 数据尺度控制, 所有变量数据均为 2002~2018 年的月尺度系列。

分前期降水 - 土壤湿度(P-SM)和土壤湿度 - 后期降水(SM-P)两组、设置 0 d, 5 d, 10 d, 15 d, 20 d 五种滞时开展研究, 同时也检测控制变量与两个分析项的相关性, 验证控制变量设置的合理性。结果见图 3, 可得以下结论:

1) 排除蒸散发及 NDVI, P-SM 在 0~10 d 滞时范围内存在显著($\alpha = 0.01$)偏相关关系, SM-P 在 1~8 d 滞时范围也存在显著($\alpha = 0.01$)偏相关关系;

2) P-SM 在 0~20 d 滞时范围内是正相关, SM-P 在 15~20 d 滞时范围内是负相关;

3) 经检验, 蒸散发及 NDVI 是强控制变量, P-SM 和 SM-P 下均通过 $\alpha = 0.01$ 的验证。

综上，降水 0~10 d 内显著作用于土壤湿度，土壤湿度增多后又于 0~8 d 内显著作用于降水，均显示为正相关关系。即表明，排除混杂因素蒸散发及 NDVI 后，本研究所用数据序列存在土壤湿度 - 降水的正反馈。蒸散发及 NDVI 是强控制变量，可归因于存在于东亚季风区土壤湿度 - 降水的局地正反馈作用：较大的土壤湿度带来较大的蒸散发，输送水汽的同时，地表感热减少、潜热增加，波恩比变小使大气边界层变薄，增加了低层大气的湿熵，加大了对流不稳定性，水汽叠加边界条件改变导致降水发生。

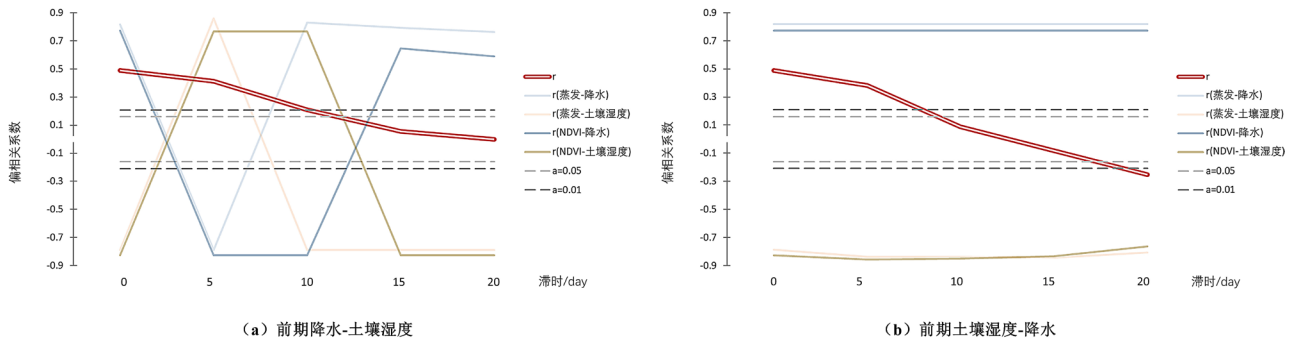


Figure 3. Results of partial correlation between soil moisture and precipitation in the Yangtze River catchment. Red hollow broken lines represent correlation between early precipitation-soil moisture and soil moisture-later precipitation. The horizontal axis refers to five lag time situations. Other broken lines represent evapotranspiration-precipitation, evapotranspiration-soil moisture, NDVI-precipitation, NDVI-soil moisture (see legend on the right). The dark and light dashed lines respectively represent the correlation thresholds at a significance level of $a = 0.05$ and 0.01

图 3. 长江流域土壤湿度对降水反馈的偏相关关系结果。红色空心折线表示前期降水 - 土壤湿度和土壤湿度 - 后期降水的偏相关关系，横坐标指代五种滞时情况；其他折线表示蒸散发 - 降水、蒸散发 - 土壤湿度、NDVI - 降水、NDVI - 土壤湿度的偏相关关系(见右侧图例)；深色和浅色虚线分别代表显著性水平 $a = 0.05$ 和 0.01 下的相关性阈值

5.2. 基于转移熵的反馈特征

降水和土壤湿度的日时间序列 $P(t)$ 、 $SM(t)$ 可视为动态网络的节点，将序列按四季分为四组，设置滞时范围 $0, 1, \dots, 20d$ ，基于转移熵识别土壤湿度对降水反馈的时空特征。本文采用直方图法估计概率密度函数，代入可计算 3 个特征指标，分析如下。

图 4 是长江流域土壤湿度对降水的反馈空间特征。P-SM 的显著滞时 τ_{01} 较小，空间差异性也较小， τ_{01} 较长的区域集中于流域中部的汉江和乌江水系；SM-P 的显著滞时 τ_{02} 较大，空间差异性也较大， τ_{02} 较长的区域集中于流域西北部的金沙江、岷江、沱江和嘉陵江水系。整体上，P-SM 的长滞时相对预测度 η 值大于 SM-P，即降水滞时历史 $P(t-\tau)$ 对土壤湿度当前状态 $SM(t)$ 提供的信息较不冗余，提供的信息中，U 信息与 S 信息比例居多，P-SM 不冗余的区域位于金沙江下游和流域东部的洞庭湖、鄱阳湖、太湖水系，冗余的区域位于流域中部的嘉陵江水系、汉江水系、上游干流区间和乌江水系。而土壤湿度滞时历史 $SM(t-\tau)$ 对降水当前状态 $P(t)$ 提供的信息较冗余，即 R 信息比例居多，SM-P 冗余的区域位于流域南部的金沙江下游和流域东部的洞庭湖、鄱阳湖、太湖水系，不冗余的区域位于流域北部的汉江水系、中下游干流区间。金沙江下游和流域东部的洞庭湖、鄱阳湖、太湖水系在 P-SM 和 SM-P 都较冗余，因此该区域内土壤湿度 - 降水反馈更受蒸散发、NDVI 的混杂影响。

图 4(a)包含指标 $TE(P(t-\tau) - SM(t))/H(P(t-\tau))$ ，图 4(b)包含指标 $TE(SM(t-\tau) - P(t))/H(SM(t-\tau))$ 。整体上，P-SM 的归一化转移熵 TE/H 明显高于 SM-P，苏辉[21]以格兰杰因果关系和新型因果关系识别同深度下降水与土壤湿度的因果关系，得到相近结论。降水对土壤湿度具有直接的因果关系，因果影响较强，而土壤湿度对降水的因果混杂了间接关系，因果影响较弱，从因果机理上可验证信息熵识别耦合强度的适用性，可据此比较不同地区的耦合强弱。P-SM 的信息转移强度在长江流域的干流区域较高，在其余支流区域上较低，推测是由于降水导致干流的河道水位上升，从而补充了周边土壤湿度，使 P-SM 信息转移强度较高。SM-P 的信息转移强度与干支流关系不明

显，强度更高的区域集中在流域东南部，尤其是洞庭湖、鄱阳湖地区，个别分布于流域西部的金沙江地区。

图 5 是长江流域土壤湿度对降水的反馈季节特征。SM-P 的 τ_{02} 长度在秋季最长， τ_{02} 长度较长的区域位于流域南部的乌江和洞庭湖水系；冬季的 τ_{02} 长度次之，集中区域向北部偏移至乌江水系和长江干流；春秋的 τ_{02} 长度相对短，分布也较为平均。整体上，春冬季的 η 值低于夏秋季；空间维度上，夏秋冬季的 η 值在中部的上游干流区间和嘉陵江水系较高，冬季的鄱阳湖和太湖具有较低的 η 值，而春季则在西部的金沙江下游及岷江水系较高，洞庭湖附近 η 较低。

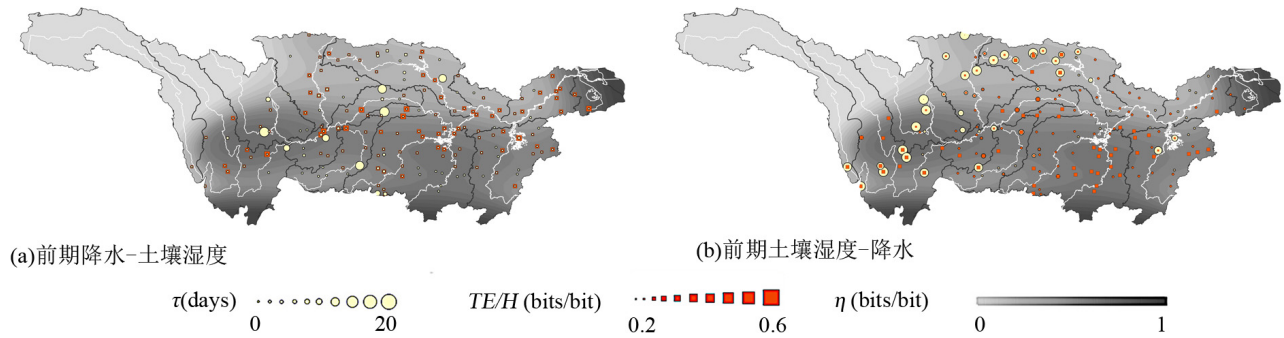


Figure 4. Spatial characteristics of soil moisture-precipitation feedback in the Yangtze River catchment. η is represented by gray background, and the color depth corresponds to its value; TE/H is represented by orange squares, and the side length corresponds to its value; τ is represented by yellow circles, and the radius corresponds to its value

图 4. 长江流域土壤湿度 - 降水反馈空间特征。 η 以灰底色表示，颜色深度对应指标值； TE/H 以橘色方块表示，方块边长对应指标值；显著滞时 τ 以黄色圆表示，圆半径对应指标值

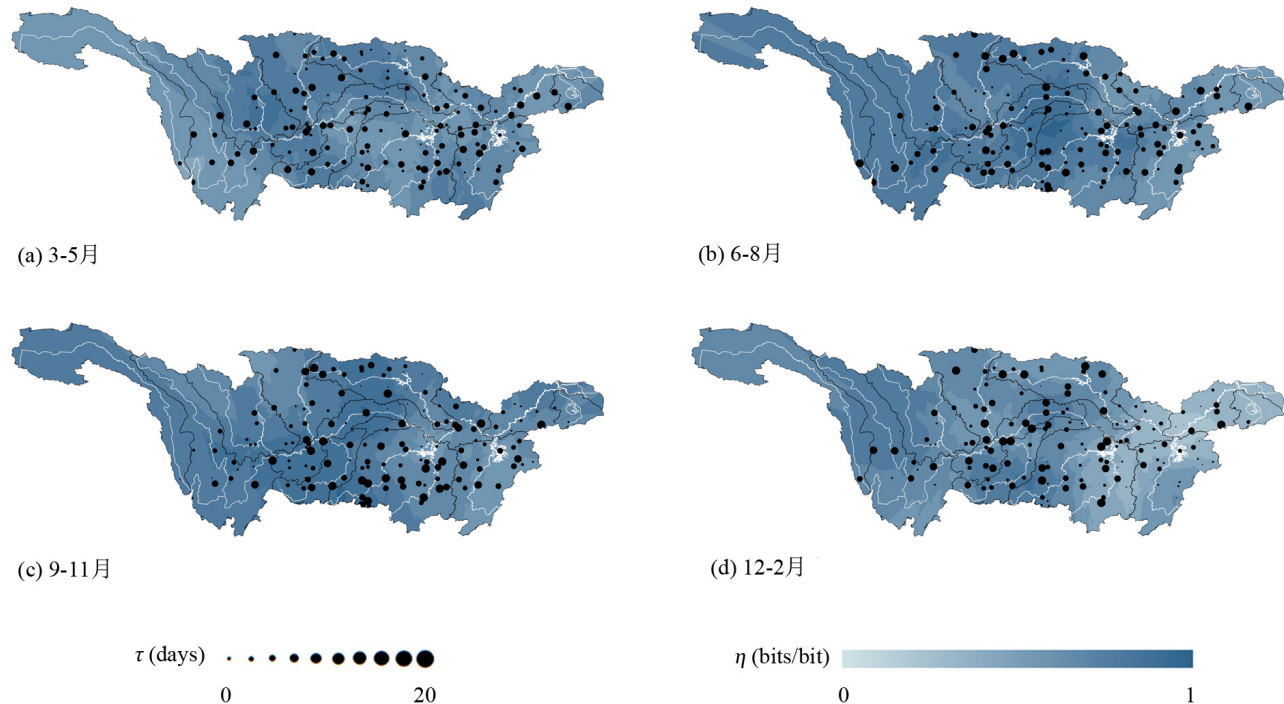


Figure 5. Seasonal characteristics of soil moisture-precipitation feedback in the Yangtze River catchment. Figures from the upper left corner are March to May (spring), June to August (summer), September to November (autumn) and December to February (winter); η is represented by blue background, and the color depth corresponds to its value; τ is represented by black circles, and the radius corresponds to its value

图 5. 长江流域土壤湿度 - 降水反馈季节特征。 η 以蓝底色表示，颜色深度对应指标值；显著滞时 τ 以黑色圆表示，圆半径对应指标值

图 6 和表 1 分别是 168 个站点土壤湿度对降水反馈显著滞时的频数分布图和统计值。可知, P-SM 的区域平均显著滞时 1.8 d, 集中于 1~2 d 滞时范围; SM-P 的区域平均显著滞时 4.3 d, 大部分站点显著滞时 1 d, 其余站点零散分布于 2~20 d 滞时范围。

图 7 和表 2 分别是 168 个站点土壤湿度对降水反馈归一化转移熵的频数分布图和统计值。可知, P-SM 的区域平均归一化转移熵 0.51, 明显高于 SM-P 的 0.13, 且 P-SM 归一化转移熵的空间差异性略高。

Table 1. Statistical values of significant lag time for soil moisture-precipitation feedback in the Yangtze River catchment

表 1. 长江流域土壤湿度对降水反馈显著滞时统计值

显著滞时/days	最小值	最大值	均值	标准差	零值数
前期降水 - 土壤湿度	1	20	1.85	3.29	0
前期土壤湿度 - 降水	1	20	4.3	5.94	0

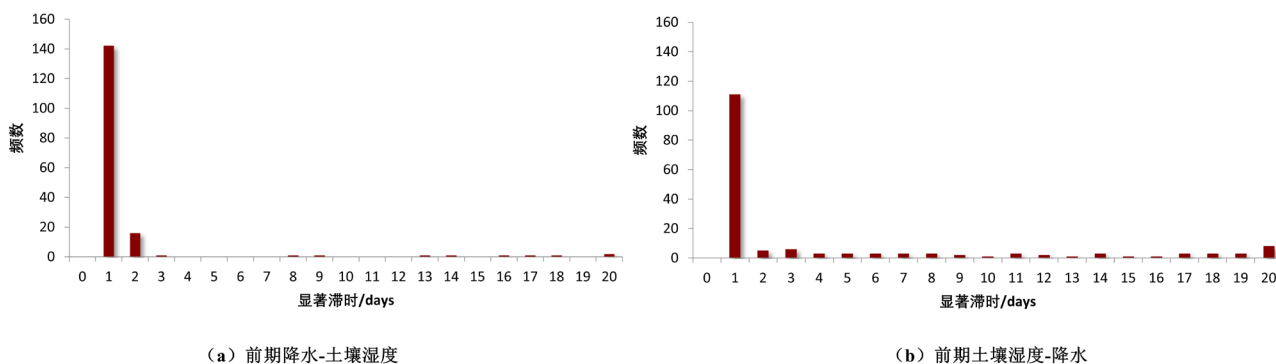


Figure 6. Frequency distribution of significant lag time for soil moisture-precipitation feedback in the Yangtze River catchment

图 6. 长江流域土壤湿度对降水反馈显著滞时频数分布图

Table 2. Statistical values of TE/H for soil moisture-precipitation feedback in the Yangtze River catchment

表 2. 长江流域土壤湿度对降水反馈归一化转移熵统计值

TE/H	最小值	最大值	均值	标准差	零值数
前期降水 - 土壤湿度	0.22	0.61	0.51	0.056	0
前期土壤湿度 - 降水	0.046	0.22	0.13	0.03	0

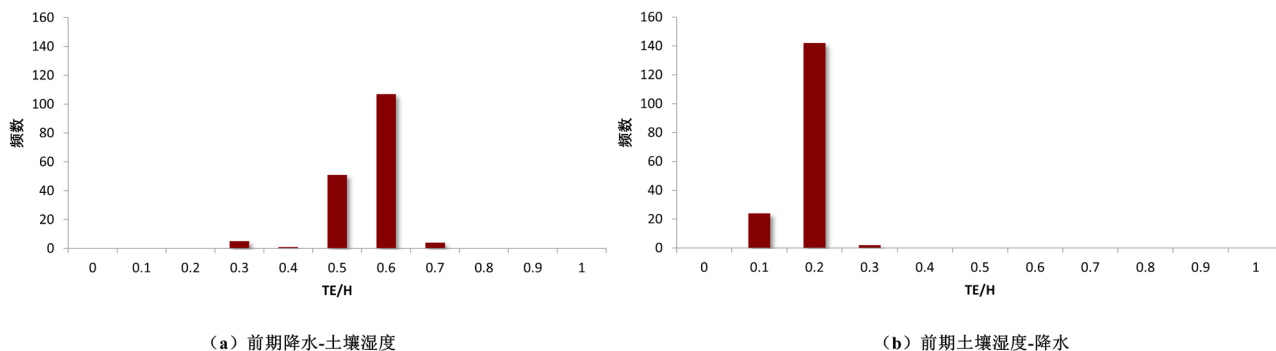


Figure 7. Frequency distribution of TE/H for soil moisture-precipitation feedback in the Yangtze River catchment

图 7. 长江流域土壤湿度对降水反馈归一化转移熵频数分布图

综上, P-SM 和 SM-P 均存在滞时。一方面, 土壤湿度的记忆性使土壤湿度对降水补偿的响应存在滞后; 另一方面, 土壤湿度若要直接补充大气水汽, 要以蒸散发为中介变量, 若要通过间接反馈, 则要改变成云致雨的

大气条件, 滞时较前者更长, 并且受陆面状况影响很大, 因此空间变异性较大。信息指标可反映序列间的时间依赖程度, 据此分析得到 P-SM 和 SM-P 的滞时集中在 1~2 天。Lam、Bierkens 和 van den Hurk [22] 和李泽君 [1] 以广义线性模型得出前一天的土壤湿度会影响次日的降水概率, 与信息指标的结果相近, 且根据下渗率和土壤深度, 距地表 2~5 cm 深度的土壤湿度应在较短的时间内对降水产生响应, 因此从机理上看, 信息指标的滞时识别更具合理性。显著滞时有明显的地区差异, 因此可据此推断各地区陆气过程的滞时效应。

金沙江下游和洞庭湖、鄱阳湖、太湖水系等近湖区域土壤湿度对降水反馈较快且强度较大。基于陆气耦合模式 LIS-WRF、CAM-CLM、GLACE 等试验表明, 位于大陆内部, 远离海洋的干湿过渡地带是土壤湿度对降水反馈的热点地区 [7], 如北美大平原、美国中西部地区 [23], 同样为近湖区域, 位于美国五大湖西部及南部。邹海波 [24] 指出鄱阳湖跟美国五大湖一样具有湖效应降水, 同理, 洞庭湖水域面积更大, 识别出较大的反馈强度。近湖区域土壤湿度对降水可能的反馈机制可解释为: 湖陆温差激发湖陆风环流, 低层气流向湖中心辐合, 抬升湖体上空的暖湿空气, 进而触发对流形成湖效应降水, 水量交换过程频繁, 可能导致较低的 η 值, 较高的 r 值和 TE/H 值, 即近湖区域土壤湿度在短时间反馈于降雨且反馈强度较高。

6. 结语

本文基于长江流域 2002~2018 年的气象观测及陆面遥感反演数据, 以显著滞时、不同滞时的相对预测度和归一化信息熵 3 个信息指标识别土壤湿度对降水的反馈特征, 并以偏相关方法验证了数据系列中土壤湿度对降水的反馈, 主要结论如下:

1) 长江流域前期降水对土壤湿度影响的平均显著滞时 1.8 d, 空间差异性小、归一化转移熵 0.51, 影响强度较大; 土壤湿度对后期降水影响的平均显著滞时 4.3 d, 空间差异性大、归一化转移熵 0.13, 影响强度较小;

2) 金沙江下游和洞庭湖、鄱阳湖、太湖水系等近湖区域, 由于水量交换频繁且存在湖效应降水, 识别出土壤湿度对降水反馈较快且强度较大;

3) 排除蒸散发及 NDVI 后, 偏相关分析得到降水 0~10 d 内作用于土壤湿度, 土壤湿度增多后又在 0~8 d 内作用于降水, 均显示为显著的正相关关系, 可归因于土壤湿度对降水的局地正反馈, 相比于信息论方法, 偏相关识别出的滞时长度不合理。

应当指出, 受限于高维概率密度函数估计, 基于转移熵的因果推断未能排除多维变量影响; 同时, 今后可开展陆气耦合模式不同集合下的数值模拟试验, 以期对前述的反馈特征予以验证。

基金项目

国家自然科学基金项目(51861125102)和清江水电开发有限责任公司(ENQJ-TD1-FW-2019185)资助。

参考文献

- [1] 李泽君. 陆-气反馈关系的统计特征识别及建模研究[D]: [博士学位论文]. 武汉: 武汉大学, 2019.
LI Zejun. Statistical relationships identifying and modeling for land-atmosphere feedback. Ph.D. Thesis, Wuhan: Wuhan University, 2019. (in Chinese)
- [2] 刘源. 中国境内土壤湿度-降水耦合强度研究[D]: [硕士学位论文]. 兰州: 兰州大学, 2016.
LIU Yuan. A study of soil moisture-precipitation coupling strength in China. Master's Thesis, Lanzhou: Lanzhou University, 2016. (in Chinese)
- [3] 王健. 土壤湿度变化对全球陆-气耦合热点地区近地层温度影响的研究[D]: [硕士学位论文]. 兰州: 兰州大学, 2018.
WANG Jian. Impact of soil moisture variations on near-surface temperature of land-atmosphere coupling hot spot regions. Master's Thesis, Lanzhou: Lanzhou University, 2018. (in Chinese)
- [4] FINDELL, K. L., ELTAHIR, E. A. B. An analysis of the soil moisture-rainfall feedback, based on direct observations from Illinois. *Water Resources Research*, 1997, 33(4): 725-735. <https://doi.org/10.1029/96WR03756>
- [5] SALVUCCI, G. D., SALEEM, J. A. and KAUFMANN, R. Investigating soil moisture feedbacks on precipitation with tests of

- Granger causality. *Advances in Water Resources*, 2003, 25(8): 1305-1312. [https://doi.org/10.1016/S0309-1708\(02\)00057-X](https://doi.org/10.1016/S0309-1708(02)00057-X)
- [6] FREMME, A., SODEMANN, H. The role of land and ocean evaporation on the variability of precipitation in the Yangtze River valley. *Hydrology and Earth System Sciences*, 2019, 23(6): 2525-2540. <https://doi.org/10.5194/hess-23-2525-2019>
- [7] KOSTER, R. D., DIRMEYER, P. A., GUO, Z. C., *et al.* Regions of strong coupling between soil moisture and precipitation. *Science*, 2004, 305(5687): 1138-1140. <https://doi.org/10.1126/science.1100217>
- [8] CHOW, K. C. Time-lagged effects of spring Tibetan Plateau soil moisture on the monsoon over China in early summer. *International Journal of Climatology*, 2007, 28(1): 55-67. <https://doi.org/10.1002/joc.1511>
- [9] 王林, 王磊, 等. 青藏高原春季土壤湿度对长江中下游地区初夏短期气候影响的数值模拟[J]. 成都信息工程大学学报, 2017(2): 81-88.
WANG Lin, WANG Lei, *et al.* An numerical simulation on the effect of spring soil moisture in Tibetan Plateau on early summer short-term climate over middle and lower reaches of Yangtze River. *Journal of Chengdu University of Information Technology*, 2017(2): 81-88. (in Chinese)
- [10] Boé, J. Modulation of soil moisture-precipitation interactions over France by large scale circulation. *Climate Dynamics*, 2012, 40(3-4): 875-892. <https://doi.org/10.1007/s00382-012-1380-6>
- [11] KUMAR, P., GUPTA, H. V. Debates—Does information theory provide a new paradigm for earth science? *Water Resources Research*, 2020, 56(2): 1-13. <https://doi.org/10.1029/2019WR026398>
- [12] BRUNSELL, N. A. A multiscale information theory approach to assess spatial-temporal variability of daily precipitation. *Journal of Hydrology*, 2010, 385(1): 165-72(8). <https://doi.org/10.1016/j.jhydrol.2010.02.016>
- [13] GOODWELL, A. E., KUMAR, P. Temporal information partitioning networks (TIPNets): A process network approach to infer ecohydrologic shifts. *Water Resources Research*, 2017, 53(7): 5899-5919. <https://doi.org/10.1002/2016WR020218>
- [14] GOODWELL, A. E., KUMAR, P. Temporal information partitioning: Characterizing synergy, uniqueness, and redundancy in interacting environmental variables. *Water Resources Research*, 2017, 53(7): 5920-5942. <https://doi.org/10.1002/2016WR020216>
- [15] SILVA, V. D. P. R. D., FILHO, A. F. B., ALMEIDA, R. S. R., *et al.* Shannon information entropy for assessing space-time variability of rainfall and streamflow in semiarid region. *Science of the Total Environment*, 2016, 544: 330-338. <https://doi.org/10.1016/j.scitotenv.2015.11.082>
- [16] FRANZEN, S. E., FARAHANI, M. A. and GOODWELL, A. E. Information flows: Characterizing precipitation-stream flow dependencies in the Colorado Headwaters with an information theory approach. *Water Resources Research*, 2020, 56(10). <https://doi.org/10.1029/2019WR026133>
- [17] SCHREIBER, T. Measuring information transfer. *Physical Review Letters*, 2000, 85(2): 461-464. <https://doi.org/10.1103/PhysRevLett.85.461>
- [18] ANDREW, B., BART, N., GENGXIN, O., *et al.* Quantifying process connectivity with transfer entropy in hydrologic models. *Water Resources Research*, 2019, 55(6): 4613-4629. <https://doi.org/10.1029/2018WR024555>
- [19] GOODWELL, A. E., KUMAR, P. A changing climatology of precipitation persistence across the United States using information-based measures. *Journal of Hydrometeorology*, 2019, 20(8): 1649-1666. <https://doi.org/10.1175/JHM-D-19-0013.1>
- [20] DU, J. Y., KIMBALL, *et al.* A global satellite environmental data record derived from AMSR-E and AMSR2 microwave earth observations. *Earth System Science Data*, 2017, 9(2): 791-808. <https://doi.org/10.5194/essd-9-791-2017>
- [21] 苏辉. 降水与土壤湿度因果关系分析及预测研究[D]: [硕士学位论文]. 杭州: 杭州电子科技大学, 2018.
SU Hui. Causality analysis between precipitation and soil moisture and precipitation prediction. Master's Thesis, Hangzhou: Hangzhou Dianzi University, 2018. (in Chinese)
- [22] LAM, A., BIERKENS, M. F. P. and VAN DEN HURK, B. J. J. M. Global patterns of relations between soil moisture and rainfall occurrence in ERA-40. *Journal of Geophysical Research: Atmospheres*, 2007, 112(D17): D17116. <https://doi.org/10.1029/2006JD008222>
- [23] MEI, R., WANG, G. Summer land-atmosphere coupling strength in the United States: Comparison among observations, reanalysis data, and numerical models. *Journal of Hydrometeorology*, 2012, 13(3): 1010-1022. <https://doi.org/10.1175/JHM-D-11-075.1>
- [24] 邹海波. 鄱阳湖湖效应降水的统计分析 with 个例研究[D]: [博士学位论文]. 兰州: 兰州大学, 2020.
ZOU Haibo. Statistical analysis and cases studies of lake-effect precipitation over Poyang Lake. Ph.D. Thesis, Lanzhou: Lanzhou University, 2020. (in Chinese)