Land surface models significantly underestimate the impact of land-use changes on global evapotranspiration

Qilin Wang1,2,*, Yingping Wang1, Lu Zhang1, Shujing Qin1,2,3, Quan Zhang1,2,3, Pan Liu1,2,3, Liu Liu1,*, Kaijie Zou1,2,3, Shujie Cheng1,2,3 and Lei Cheng1,2,3

1 State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan 430072, People’s Republic of China
2 Hubei Provincial Collaborative Innovation Center for Water Resources Security, Wuhan 430072, People’s Republic of China
3 Hubei Provincial Key Lab of Water System Science for Sponge City Construction, Wuhan University, Wuhan, Hubei, People’s Republic of China
4 CSIRO Oceans and Atmosphere, Aspendale, VIC 3195 Victoria, Australia
5 CSIRO Land and Water, Black Mountain, Canberra ACT 2601, Australia
6 College of Water Resources and Civil Engineering, China Agricultural University, Beijing 100038, People’s Republic of China

* Author to whom any correspondence should be addressed.
E-mail: lei.cheng@whu.edu.cn

Keywords: land-use changes, terrestrial evapotranspiration, land surface models, Budyko framework, land-use change data interpretation, model structure

Supplementary material for this article is available online

Abstract

Despite numerous assessments of the impact of land-use change (LUC) on terrestrial evapotranspiration (ET) that have been conducted using land surface models (LSMs), no attempts have been made to evaluate their performance in this regard globally. Errors in simulating LUC impacts on ET largely stem from LUC data interpretation (LI, i.e. mapping of gridded LUC data into annual plant function types) and model structure (MS, i.e. parameterization of land-surface processes). The objective of this study was to benchmark ET estimates from four LSMs using the Zhang-curve, a prototype of the Budyko framework that has been validated against global hydrological observations and used widely to quantify the impacts of LUC on ET. A framework was further proposed to quantify and attribute errors in estimated ET changes induced by LI or MS. Results showed that all LSMs underestimated ET changes by about 55%–78%, and 37%–48% of the error was attributable to LI, but only 11%–32% of the error was attributable to MS across the four LSMs. From a hydrological perspective, our analysis provided insights about the errors in estimated impacts of LUC on ET by LSMs. The results demonstrated that LUC data interpretation accounted for a larger fraction of errors than LSM structure. Therefore, there is an urgent need for the defining and development of consistent protocols for interpreting global LUC data for future assessments.

1. Introduction

Land-use change (LUC) has significantly affected global water and energy cycles, climate change, and ecosystem sustainability through direct changes to the timing and magnitude of evapotranspiration (ET) (Foley et al 2005, Bonan 2008, Sterling et al 2013, Cheng et al 2017a, Wang et al 2021). The impacts of LUCs on global terrestrial ET have been commonly evaluated using land surface models (LSMs) (Piao et al 2007, Liu et al 2008, Sterling et al 2013) that can also be coupled with General Circulation Models for estimating LUC effects on climate change and possible feedbacks between them (Pitman et al 2009, Ghent et al 2011, de Noblet-durocoudre et al 2012, Brovkin et al 2013). A number of model comparison projects have been conducted to investigate LUC impacts on terrestrial ET, such as the Trends in Net Land–Atmosphere Exchange (TRENDY) (Le Quéré et al 2016), the Land–Use and Climate, Identification of robust impacts project (LUCID) (Lejeune et al 2017), the Multi-scale Synthesis and Terrestrial Model...
Intercomparison Project (MsTMIP) (Huntzinger et al 2013), and the Land Use Model Intercomparison Project (LUMIP) (Lawrence et al 2016). However, considerable discrepancies remain in their values of estimated ET and LUC-induced ET changes (de Noblet-ducoudre et al 2012, Kumar et al 2013, Mao et al 2015, Pan et al 2020). Therefore, it is crucial to benchmark and to attribute errors in estimated LUC-induced ET changes to better understand their impacts on the hydrological cycle, climate change, and sustainable development.

Quantifying errors in estimated LUC-induced ET changes by LSMs globally is very difficult, as direct large-scale ET measurements are not presently available. Although a few studies have attempted to evaluate the performance of LSMs for estimating the impacts of LUC on ET using field observations (Chen et al 2018, Cai et al 2019), our understanding of the causes of the large discrepancies in the estimated surface fluxes among LSMs at the global scale is still very limited. Previous studies based on catchment hydrological observations have demonstrated that forest coverage change is a key factor for quantifying the impact of LUC on water yield or ET at mean annual scale (Zhang et al 2001, Brown et al 2005, Zhou et al 2015, Ning et al 2020). Impacts of forest cover changes on mean annual ET are well understood based on paired catchment experiments. The hydrological impacts of forest cover changes on catchment water balance have been widely recognized, and it is generally accepted that afforestation will lead to increased ET and decreased streamflow (Zhang et al 2001, Cheng et al 2017b). The Budyko framework has been shown to provide accurate estimates of vegetation cover change impacts on ET (Li et al 2009, Yang et al 2009, Zhang et al 2012, Lei et al 2014, Jaramillo et al 2018, Cheng et al 2021). Therefore, the Budyko framework can be considered as an independent and robust framework to benchmark the impacts of LUC on regional water balance or ET simulated by LSMs.

Changes in ET resulting from LUCs can be estimated using LSMs. However, large differences exist in the estimates of LUC impacts among LSMs. These differences can be attributed to differences in (a) interpretation of LUC data (hereafter denoted as LI) and (b) model structure (hereafter denoted as MS). They are sourced from two modeling steps (figure 1). One of the first steps in estimating the impact of LUC on ET in LSMs is to convert the input data of LUC (i.e. land-use types, LUTs) into the broadly defined plant function types (PFTs) that are commonly used in LSMs. However, it is known that the definition and number of PFTs vary among different LSMs (Poulter et al 2011). For example, the number of PFTs varies from 5 to 25 among the 12 LSMs used in CMIP5 (Li et al 2018). This means that a given input LUT can be represented differently in terms of PFTs by LSMs. Furthermore, LUC are represented as changes in fractional PFTs in the LSMs based on transition rules that differ among the models (de Noblet-ducoudre et al 2012, Brovkin et al 2013, Hartley et al 2017, Peng et al 2017, Di Vittorio et al 2018). For instance, cropland expansion would be represented by first reducing fractional grassland PFTs and then forest PFTs in some models (e.g. MOSES2), while other models would proportionally reduce all natural PFTs (e.g. ORCHIDEE). Different transition rules would lead to differences in the estimated LUC impacts on ET. Model structure can also affect estimated impacts of LUC on ET. LSMs simulate exchange of water, carbon, and energy between the surface and the atmosphere using models with different structural complexity. Hence, it is expected that differences in LSM structure will lead to differences in their estimates of LUC impacts on ET. In this study, model structure included process representation, parameterization, and spatial implementation of the model. No attempts were made to examine more specific aspects of process representation, for example, methods for modeling different components of ET (e.g. the Priestley–Taylor method used in LPJ-wsl and the Modified Penman–Monteith method used in ORCHIDEE for transpiration), and different parameterization of PFT attributes (e.g. rooting depth, maximum assimilation rate).

Previous studies have focused only on either LI or MS, and have shown that both can lead to large discrepancies in simulated ET among different LSMs (e.g. de Noblet-ducoudre et al 2012, Chen et al 2018, Di Vittorio et al 2018). However, to our knowledge, no studies have attempted to separate the contributions of LI and MS with respect to estimated discrepancies. To better evaluate the simulated effects of LUC on ET and their errors by LSMs, a framework based on the Zhang-curve (one prototype of the Budyko framework) was proposed to benchmark estimated ET changes by four LSMs of TRENDY. The primary objectives of the present study were to: (a) determine whether LSMs can capture LUC-induced ET changes estimated by the Zhang-curve (hereafter ΔET₂); (b) quantify the total relative errors in estimated ET changes by LSMs (hereafter ΔETₘ); (c) attribute the relative errors to either LI or MS.

2. Proposed benchmarking framework

2.1. The Zhang-curve for estimating ET

Based on paired catchment observations and the Budyko theoretical framework, Zhang et al (2001) proposed a semi-empirical function for quantifying the impacts of LUCs (e.g. forest to grassland) on ET:

\[ ET_Z = f \left( \frac{ET_{Z,F} + (1-f) ET_{Z,G}}{P} \right) \]

\[ = \left[ \frac{1 + \frac{w_{ET_f}}{P}}{1 + \frac{w_{ET_f}}{P} + \frac{P}{ET_0}} \right] \left[ \frac{1 + \frac{w_{ET_g}}{P}}{1 + \frac{w_{ET_g}}{P} + \frac{P}{ET_0}} \right] \]

\[ = \left[ \frac{1 + \frac{w_{ET_f}}{P}}{1 + \frac{w_{ET_f}}{P} + \frac{P}{ET_0}} \right] \left( 1 - \frac{1}{1} \right) \]

\[ = \left[ \frac{1}{1 + \frac{w_{ET_f}}{P}} \right] + \left( 1 - \frac{1}{1 + \frac{w_{ET_f}}{P}} \right) \]

\[ \text{(1)} \]
Figure 1. Schematic representation of two modeling steps and the sources of errors in the estimated LUC-induced surface flux changes of LSMs in the case of deforestation/afforestation. Source 1 is land-use change data interpretation (LI), including definition and prescribed number of PFTs and transition rules. Source 2 is model structure (MS), including process representation, parameterization, and spatial implementation.

where ET is mean annual evapotranspiration (mm); ET,F and ET,G are ET for forest and grassland, respectively; p is the forested land-use fraction; p is mean annual precipitation (mm yr⁻¹); ET,F and ET,G and w₁ and w₂ are mean annual potential evapotranspiration (mm yr⁻¹) and empirical plant-available water coefficients of forest (grassland), respectively. Based on observations from over 200 forest and grassland catchments across a wide range of climatic zones (see appendix A in Zhang (1999)), ET,F and ET,G and w₁ and w₂ were determined as 1410 (1100) mm yr⁻¹ and 2.0 (0.5),
respectively. Parameters are spatially invariant, and the Zhang-curve has been widely validated and applied globally (Brown et al. 2005, McVicar et al. 2007, Zhang et al. 2011, Ponce-Campos et al. 2013). In this study, changes in mean annual ET of LUC cells estimated by the Zhang-curve were regarded as reference values for benchmarking the estimated changes in ET from LUC by LSMs.

### 2.2. Benchmarking and model error decomposition

Based on the Zhang-curve, a benchmarking framework was proposed to quantify total relative error in LUC-induced ET changes as estimated by LSMs, and to attribute the total relative error to either LUC data interpretation errors (i.e., LI) or to model structure errors (i.e., MS). The proposed benchmarking framework is illustrated schematically in figure 2.

For two cells with the same climate (i.e., rainfall and potential evaporation) but different forest cover, the difference in mean annual ET can be estimated from equation (2) as:

$$\Delta ET_Z = (f_1 - f_2)(ET_{Z,F} - ET_{Z,G})$$

where $f_1$ and $f_2$ are the fractions of forested LUTs in cell 1 and cell 2, respectively.

In the LSMs, the difference in ET between these two cells is estimated based on PFTs rather than LUTs. ET values are calculated as:

$$ET_{Z,m,1} = p_1 ET_{m,F} + (1 - p_1) ET_{m,G}$$

$$ET_{Z,m,2} = p_2 ET_{m,F} + (1 - p_2) ET_{m,G}$$

where $ET_{Z,m,1}$ and $ET_{Z,m,2}$ are estimated ET values for the two cells, respectively; $p_1$ and $p_2$ are fractions of all forest PFTs in cell 1 and cell 2, and $ET_{m,F}$ and $ET_{m,G}$ are estimated ET values for forest and grassland PFTs, respectively.

The difference in estimated ET between the two cells can be calculated as:

$$\Delta ET_m = ET_{m,1} - ET_{m,2} = (p_1 - p_2) (ET_{m,F} - ET_{m,G}).$$

The difference between estimated ET change using the Zhang-curve and that of the LSM can be expressed as:

$$\Delta ET_m - \Delta ET_Z = (p_1 - p_2) (ET_{m,F} - ET_{m,G}) - (f_1 - f_2)(ET_{Z,F} - ET_{Z,G})$$

or

$$\epsilon_{TOT} = \Delta ET_m - \Delta ET_Z = \Delta p \delta_m - \Delta f \delta_Z.$$

Equation (7) can be expressed as:

$$\epsilon_{TOT} = (\Delta p - \Delta f) \delta + (\delta_m - \delta_Z) \Delta p = \epsilon_{LI} + \epsilon_{MS}$$

where $\Delta p \delta$ is the estimated ET change between forest and grassland PFTs using the Zhang-curve and is denoted as $\Delta ET'_Z$. $\epsilon_{LI}$ represents the difference in estimated changes in ET due to differences in mapping of LUTs to PFTs. $\epsilon_{MS}$ represents the difference in estimated changes in ET due to differences in model structure.

In this study, the linear slope between $\epsilon_{TOT}$ and $\Delta ET_Z$ of all study cells was considered to be the total relative error (i.e., $RE_{TOT}$). Total relative error (%) can be estimated as:

$$RE_{TOT} = \frac{\Delta ET_m - \Delta ET_Z}{\Delta ET_Z} = \frac{\epsilon_{TOT}}{\Delta ET_Z}.$$
It is expected that $\Delta ET$ is greater than zero (Zhang et al 2001). If the value of $\Delta ET_{TOT}$ is negative (positive), it means that LSMs underestimate (overestimate) the LUC-induced ET change. Similarly, if the value of $\Delta ET_{TOT} (\Delta ET_{MS})$ is negative (positive), it means that LSMs underestimate (overestimate) the LI (MS)-induced ET changes.

3. Data and application

3.1. LSM output

Annual outputs of four LSMs from TRENDY project v5 (Le Quéré et al 2016) were collected and used in this study. These four LSMs were CABLE (Zhang et al 2013), ISAM (Jain et al 2013), LPI-GUESS (Smith et al 2014), and ORCHIDEE (Krinner et al 2005). These LSMs were selected because both fractional PFTs and grid-level ET are available from these models with a spatial resolution of 0.5°. LSMs in the TRENDY v5 project are driven by the same climate forcing (i.e. Climate Research Union (CRU) datasets) and the same LUC data product (i.e. LUH-HYDE3.1).

For two neighboring land cells with different land-use cover (i.e. forest and grassland), climatic conditions can be assumed to be the same or to be very similar. Hence, the simulated ET change between these two cells can be attributed to LUC. In this study, two neighboring land cells (one under forest cover and the other under grassland) were considered as paired-cells. $\Delta ET_{m}$ between the two cells was calculated as the ET of the forested cell minus that of the grassland cell over the period of 1996–2005 based on the modeling outputs of scenario 3 of the TRENDY project with changing CO$_2$, climate, and LUC data.

3.2. LUC and climate data

LUC data used in this study were the same as those used in the TRENDY simulations (i.e. LUH-HYDE3.1) (Hurt et al 2011, Klein Goldewijk et al 2011). They were annual, half-degree, fractional data of five LUTs including cropland, pasture, urban, primary land, and secondary land for the historical period from 1500 to 2005 (see table S1 available online at stacks.iop.org/ERL/16/124047/mmedia for a description of some LUTs). The LUH-HYDE3.1 dataset also provided mean aboveground biomass density (kg C m$^{-2}$) of secondary land and the location of potential forest (figure S1(a)). Potential forest is defined as the area with a potential aboveground biomass density larger than 2.0 kg C m$^{-2}$. In this study, annual forest coverage of LUTs in each individual cell was the sum of primary land located in the potential forest area and secondary land located in the potential forest area with mean aboveground biomass density over 2 kg C m$^{-2}$ and it was used to calculate $\Delta p$ between two neighboring cells. Derived mean annual global forested area in terms of LUTs during 1996–2005 is shown in figure S1(b). forested areas for the four LSMs are the sum of fractional forest PFTs (table S2, figures S1(c)–(f)), and were used to calculate $\Delta p$ between two neighboring cells.

The precipitation data for equation (2) was the same as the climate forcing for the TRENDY LSMs simulations (i.e. the CRU dataset) at a resolution of 0.5° averaged over the period of 1996–2005.

3.3. Identification of paired cells and application

Selection of paired cells was based on LUTs (LUH-HYDE3.1). Paired cells were identified using two criteria to minimize the difference in hydro-climatic conditions. One criterion was the window size of 2° × 2°, and the other was a maximum elevation difference of 100 m. The GMTED2010 global digital elevation model dataset was used to implement the elevation criterion (Danielson and Gesch 2011). Cells with a threshold of 80% forested (grassland) LUTs over the period of 1996–2005 were considered as fully forested (grassland) cells. Identified pairs of cells that satisfied the above criterion are shown in figure 3 (n = 1133). Most paired cells were in the transition zones of different climate types (supplementary figure S2). The uncertainty caused by different thresholds for identification of paired cells were further assessed, including threshold of forest (non-forest) LUTs coverage (85%, 80%, 75%, and 70%), window size (2.5° × 2.5°, 2° × 2°, 1.5° × 1.5°), and elevation differences (±100, ±200, ±300, ±400, ±500 m). Our benchmarked results and conclusions were robust and relatively insensitive to choices of the thresholds used for selecting the paired cells (see figures S3–S5).

4. Results

To demonstrate the application of the proposed benchmarking framework, the results of one paired-cell example in North America are presented in figure 3. These results include the estimated LUC-induced ET change (mm yr$^{-1}$) by the Zhang-curve ($\Delta ET_{z}$) and the ET changes estimated by the four LSMs ($\Delta ET_{m}$). Also presented are the relative errors (%) due to land-use change data interpretation (RE$_{LI}$) and model structure (RE$_{MS}$). For this paired-cell example, all LSMs estimated an increase in ET as expected due to LUC changing from grassland to forest. However, the ET changes estimated by the LSMs were much smaller than the changes estimated by the Zhang-curve (262 mm) with underestimations ranging from 32% to 71%. A closer examination shows that LI was responsible for 18%–53% of the underestimation, while MS accounted for 3%–18% of the underestimation.

Figure 4 shows the relationship between errors in estimated ET change and the reference ET change (i.e. $\Delta ET_{2}$) for all of the paired cells. The negative linear
Figure 3. Spatial distribution of the identified paired cells (forest cells with green circles and grassland cells with blue crosses) and the decomposed relative errors contributed by land use data interpretation (RE_{LI}) and model structure (RE_{MS}) from four LSMs for an example paired-cell location in North America. $\Delta ET_z$ is the ET difference between forest cell and grassland cell that estimated using Zhang-curve and forested land use change. $\Delta ET_m$ is the ET difference estimated by LSMs. $\Delta ET'_Z$ is the ET difference estimated using Zhang-curve with forest PFTs changes.

slopes in figure 4(a) (i.e. total relative error, RE_{TOT}) were $−0.78$, $−0.55$, $−0.58$, and $−0.59$ for CABLE, ISAM, LPJ-GUESS, and ORCHIDEE, respectively, indicating that all four LSMs underestimated LUC-induced ET change. Three of the four LSMs showed a total relative error of about $−60\%$, while CABLE had the largest relative error ($−78\%$). Clearly, all LSMs considerable underestimated LUC-induced ET change. Figures 4(b) and (c) show the relationship between $\varepsilon_{LI}$ and $\Delta ET_Z$ (figure 4(b)) and $\varepsilon_{MS}$ and $\Delta ET_Z$ (figure 4(c)). The linear slopes shown in the subpanels are the relative error contributions attributed to LI and MS (i.e. RE_{LI}, RE_{MS}). Results showed that LI accounted for about 46%, 37%, 40%, and 48% of the errors for CABLE, ISAM, LPJ-GUESS, and ORCHIDEE, respectively, while MS accounted for about 32%, 18%, 18%, and 11% of the respective errors (see figure 4(d)).

The overall total and relative errors are summarized in figure 4(d). For all four LSMs, the relative error caused by LI (ranging from $−37\%$ to $−48\%$) accounted for about 60% of total error on average, while the relative error caused by MS (ranging from $−11\%$ to $−32\%$) accounted for about 40% of the total error on average. Therefore, LI was responsible for a much greater fraction of the total error than MS.

5. Discussion

5.1. Contributions of land-use change data interpretation

Results suggest that LI resulted in larger errors than MS in quantifying LUC-induced ET changes using LSMs (RE_{LI}, the green bars in figure 4(d)). LI-induced errors were largely caused by underestimation of the forested area when mapping LUT-based LUC into PFT-based LUC. The vegetated fraction interpreted by LSMs was obviously different from the original land-use data (LUH-HYDE3.1), although the general pattern and LUC of LUH-HYDE3.1 were captured by the four LSMs. For the 1131 paired-cells used in the study, mapping of the LUH-HYDE3.1 data into PFT-based LUC. The vegetated fraction interpreted by LSMs was obviously different from the original land-use data (LUH-HYDE3.1), although the general pattern and LUC of LUH-HYDE3.1 were captured by the four LSMs. For the 1131 paired-cells used in the study, mapping of the LUH-HYDE3.1 data into PFTs by the LSMs resulted in lower forested fraction and higher non-forested fraction compared with the initial input LUTs (figure S6). The change in forested area interpreted by the four LSMs in terms of PFTs was 0.9–1.3 $\times 10^6$ km$^2$, significantly lower than the change values derived from LUH-HYDE3.1 data in terms of LUTs (1.6 $\times 10^6$ km$^2$). Because forest cover change is the most sensitive and important characteristic of LUC affecting ET, underestimation of LUC-induced ET change is likely to be caused by lower forested areas represented by the different modelling
groups. The four LSMs had different LI values representing land-cover change of the LUH-HYDE3.1 (Le Quéré et al. 2016). The number of PFTs in the four LSMs varied from 10 to 24 (see table S1). Generally, LSMs with greater numbers of PFTs (e.g. LPJ-GUESS) had relatively lower RE_{LI}. However, for some specific LSMs, paired cells with fewer numbers of PFTs had relatively lower RE_{LI} (figure S7). For the identified paired-cells in different regions, locations classified as arid savannah climate were clearly different (e.g. African savannas in figures S2, S7, and S8) than other climate classifications, possibly because vegetation dynamics of savannah ecosystems are poorly represented by LSMs (Whitley et al. 2017). In situ observation and remote sensing data can be used to improve the accuracy of PFT classification and parameterization of LSMs in these regions (Wullschleger et al. 2014, Baudena et al. 2015, Whitley et al. 2017). For the Qinghai-Tibet plateau, all LSMs interpreted identified cells in this region as fully covered with non-forest (figures S1 (b)–(f)). Thus, different LI resulted in a mismatch between PFT-based and LUT-based LUC, resulting in errors in estimated ET changes.

A previous study of LUCID experiments also showed that the impacts of deforestation on ET differed considerably among LSMs (de Noblet-ducoudre et al. 2012) due to differences in LI. Peng et al. (2017) and Di Vittorio et al. (2018) found that different transition rules resulted in differences in forest areas, and thus different carbon emission changes. How LUC data are interpreted still varies greatly among different modeling groups. This study suggested that mapping of LUC data in terms of PFTs by the four LSMs introduced large errors in estimated LUC-induced ET fluxes. Recently, a few numerical experiments (e.g. the deforestation experiment of LUMIP) have conducted to evaluate the LUC impacts on climate with more consistent deforestation experiment setup (Lawrence et al. 2016). Ma et al. (2020) highlighted the need for consistent transition rules for LUC data interpretation, and recommended a set of transition rules based on the new version of LUH-HYDE3.1 data. Therefore, efforts should be made to consistently represent LUC data in LSMs in the future for assessment of vegetation water use and its changes.

5.2. Contributions of model structure

This study also showed that model structure induced all four LSMs to significantly underestimate LUC-induced ET changes during afforestation (figure 4). In this study, model structure included process representation (e.g. equations for estimating ET), parameterization, and spatial implementation of the model.
No attempts were made to examine more specific components of process representation. Basically, underestimations caused by MS were due to inadequate parameterization of key hydrological processes. Hydrological processes in the LSMs are usually simplified as one-dimensional and free-draining flow, and are therefore unable to capture lateral water exchange (Clark et al. 2015). The consequences of this unrealistic representation can lead to failure in accurately capturing increases in continental water storage and residence time (Fan et al. 2019). Based on a 1D free-draining model, Miguez-Macho and Fan (2012) showed that ET and headwater streamflow in the Amazon were zero during the dry season, but this result was not consistent with observed data from a flux tower and hydrological gauge stations in the region. The differences in the parameterization of key processes pertinent to ET (such as LAI dynamics, soil moisture parameterizations, and rooting depth) may also induce underestimation of LUC-induced ET differences. For example, Decker (2015) showed that there was a positive ET bias in the default CABLE LSM over the PFTs with short canopies, and the bias was significantly reduced by using improved soil moisture parameterizations. Many LSMs have an unrealistically shallow rooting depth (generally less than 2 m) that can profoundly impact the accuracy of the hydrological simulation (Lee et al. 2005), and can result in lower plant available water for forest. Thereby it may underestimate the LUC-induced ET difference. However, it remains a grand challenge to identify the specific drivers or sources of errors caused by differences in model structures among different LSMs.

The proposed diagnostic framework is established based on catchment water and energy balance theory (i.e. Zhang-curve), which is one of top-down approaches and can achieve parsimony of model parameters with satisfactory performance worldwide. It can be applied to other modeling experiments as well. However, this benchmarking framework is limited to total ET and mean annual time scale and cannot provide more information about errors at processes level (i.e. different ET components) and at finer time scale (e.g. monthly or annually). More physical process of ET can be incorporated into the current framework for deeper understanding in the future.

6. Conclusions

Many modeling studies have been conducted to quantify the impact of LUC on global ET and the water cycle. However, few studies have attempted to benchmark apparent inconsistent results and to quantify the relative contributions of LUC data interpretation (LI) and model structure (MS) of LSMs. By making use of the Budyko framework that has been widely validated and used to quantify the impacts of forest cover changes on ET, estimated LUC-induced ET changes by four LSMs from the TRENDY project were benchmarked and a framework was proposed to attribute the error to either LI and MS.

Our results showed that all four LSMs captured the direction of ET change with LUC from non-forest to forest, but all four LSMs considerably underestimated the change (55%–78%). This finding was consistent with previous studies using observations from paired flux sites (Chen et al. 2018, Cai et al. 2019). More importantly, we found that both LI and MS contributed to underestimates of ET changes, and LI was the major cause for this underestimation. The results of this study demonstrate a new way to benchmark and quantify errors in estimated LUC-induced ET changes by LSMs from a hydrological perspective. The results demonstrate that there is an urgent need for consistent protocols to interpret LUC data for future assessments of LUC impacts on ET.

Data availability statement

All data, models, or code generated or used during the study are available from the corresponding author by request.

The data that support the findings of this study are available upon reasonable request from the authors.

Acknowledgments

This study was supported by the National Natural Science Foundation of China (51879193, 41890822, 51961145104); the National Key Research and Development Program of China (2018YFC0407202); and the Overseas Expertise Introduction Project for Discipline Innovation (111 Project) funded by Ministry of Education and State Administration of Foreign Experts Affairs P. R. China (B18037). Sincere gratitude is extended to the editor and anonymous reviewers for their professional comments and corrections, which greatly improved the quality of the paper.

Conflict of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

ORCID iDs

Qilin Wang @ https://orcid.org/0000-0002-2724-1790
Quan Zhang @ https://orcid.org/0000-0003-1127-5969
Pan Liu @ https://orcid.org/0000-0002-3777-6561
Liu Liu @ https://orcid.org/0000-0002-4915-206X
References

Baudena M et al 2015 Forests, savannas, and grasslands: bridging the knowledge gap between ecology and dynamic global vegetation models Biogeosciences 12 1833–48

Bonan G B 2008 Forests and climate change: forcings, feedbacks, and the climate benefits of forests Science 320 1444–9

Brovkin V et al 2013 Effect of anthropogenic land-use and land-cover changes on climate and land carbon storage in CMIP5 projections for the twenty-first century J. Clim. 26 6859–81

Brown A E, Zhang L, McMahon T A, Western A W and Vertessy R A 2005 A review of paired catchment studies for determining changes in water yield resulting from alterations in vegetation J. Hydrol. 310 28–61

Cai X, Riley W J, Zhu Q, Tang J, Zeng Z, Bishl G and Randerson J T 2019 Improving representation of deforestation effects on evapotranspiration in the ESM land model J. Adv. Model. Earth Syst. 11 2412–27

Chen L, Dirmeyer P A, Guo Z and Schulz N M 2018 Pairing FLUXNET sites to validate model representations of land-use/land-cover change Hydrol. Earth Syst. Sci. 22 111–25

Cheng L, Zhang L, Chiew F H S, Canadell J G, Zhao F, Wang Y, Hu X and Lin K 2017 Quantifying the impacts of vegetation changes on catchment storage-discharge dynamics using paired-catchment data Water Resour. Res. 53 5963–79

Cheng L, Zhang L, Wang Y, Canadell J G, Chiew F H S, Beringer J, Li L, Miralles D G, Piao S and Zhang Y 2017a Recent increases in terrestrial carbon uptake at little cost to the water cycle Nat. Commun. 8 1–10

Cheng S, Cheng L, Liu P, Qin S, Zhang L, Xu C-Y, Xiong L, Liu L and Xia J 2021 An analytical baseflow coefficient curve for depicting the spatial variability of mean annual catchment baseflow Water Resour. Res. 57 e2020WR029529

Clark M P et al 2015 Improving the representation of hydrologic processes in Earth System Models Water Resour. Res. 51 5929–56

Danielson J J and Gesch D B 2011 Global multi-resolution terrain elevation data 2010 (GMTED2010): U.S. Geological Survey Open-File Report 2011–1073 p 26

de Noblet-ducoudre N et al 2012 Determining robust impacts of land-use-induced land cover changes on surface climate over North America and Eurasia: results from the first set of LUCID experiments J. Clim. 25 3261–81

Decker M 2015 Development and evaluation of a new soil moisture and runoff parameterization for the CABLE LSM including subgrid-scale processes J. Adv. Model. Earth Syst. 7 1788–809

Di Vittorio A V, Mao J, Shi X, Chini L, Hurtt G and Collins W D 2018 Quantifying the effects of historical land cover conversion uncertainty on global carbon and climate estimates Geophys. Res. Lett. 45 974–82

Fan Y et al 2019 Hillslope hydrology in global change research and earth system modeling Water Resour. Res. 55 13757–72

Foley J A et al 2005 Global consequences of land use Science 309 570–4

Gerten D, Kaduk J, Remedios J and Balter H 2011 Data assimilation into land surface models: the implications for climate feedbacks Int. J. Remote Sens. 32 617–32

Hartley A J, MacBean N, Georgievski G and Bontemps S 2017 Uncertainty in plant functional type distributions and its impact on land surface models Remote Sens. Environ. 203 71–89

Huntzinger D N et al 2013 The North American carbon program multi-scale synthesis and terrestrial model intercomparison project—part 1: overview and experimental design Geosci. Model Dev. 6 2121–33

Hurtt G C et al 2011 Harmonization of land-use scenarios for the period 1500–2100: 600 years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands Clim. Change 109 117–61

Jain A K, Meiyappan P, Song Y and House J I 2013 CO₂ emissions from land-use change affected more by nitrogen cycle, than by the choice of land-cover data Glob. Change Biol. 19 2893–906

Jaramillo F, Cory N, Arheimer B, Laudon H, van der Velde Y, Hasper T B, Teutschbein C and Uddling J 2018 Dominant effect of increasing forest biomass on evapotranspiration: interpretations of movement in Budyko space Hydrol. Earth Syst. Sci. 22 567–80

Klein Goldewijk K, Beusen A, Van Drecht G and De Vo M 2011 The HYDE 3.1 spatially explicit database of human-induced global land-use change over the past 12 000 years Glob. Ecol. Biogeogr. 20 73–86

Kramer G, Viowy N, de Noblet-ducoudre N, Ogée J, Polcher J, Friedlingstein P, Ciais P, Sitch S and Prentice I C 2005 A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system Glob. Biogeochem. Cycles 19 GB1015

Kumar S, Dirmeyer P A, Merwade V, DeSole T, Adams J M and Niyogi D 2013 Land use/cover change impacts in CMIP5 climate simulations: a new methodology and 21st century challenges J. Geophys. Res. 118 6337–53

Lawrence D M et al 2016 The Land Use Model Intercomparison Project (LUMIP) contribution to CMIP6: rationale, and experimental design Geosci. Model Dev. 9 2973–98

Le Quéré C et al 2016 Global carbon budget 2016 Earth Syst. Sci. Data 8 605–49

Lee J, Oliveira R S, Dawson T E and Fung I 2005 Root functioning modifies seasonal climate Proc. Natl Acad. Sci. USA 102 17576–81

Lei H, Yang D and Huang M 2014 Impacts of climate change and vegetation dynamics on runoff in the mountainous region of the Haihe River basin in the past five decades J. Hydrol. 511 786–99

Lejeune Q, Seneviratne S I and Davin E L 2017 Historical land-cover change impacts on climate: comparative assessment of LUCID and CMIP5 multimodel experiments J. Clim. 30 1439–59

Li L et al 2018 Evaluating global land surface models in CMIP5: analysis of ecosystem water- and light-use efficiencies and rainfall partitioning J. Clim. 31 2995–3008

Li Z, Liu W, Zhang X and Zheng F 2009 Impacts of land use change and climate variability on hydrology in an agricultural catchment on the Loess Plateau of China J. Hydrol. 377 35–42

Liu M, Tian H, Chen G, Ren W, Zhang C and Liu J 2008 Effects of land-use and land-cover change on evapotranspiration and water yield in China during 1900–2000 J. Am. Water Resour. Assoc. 44 1193–207

Ma L et al 2020 vol 13 pp 3232–20

Mao J et al 2015 Disentangling climatic and anthropogenic controls on global terrestrial evapotranspiration trends Environ. Res. Lett. 10 094008

McVicar T R et al 2007 Developing a decision support tool for China’s re-vegetation program: simulating regional impacts of afforestation on average annual streamflow in the Loess Plateau Forest Ecol. Manage. 251 65–81

Miguez-Macho G and Fan Y 2012 The role of groundwater in the Amazon water cycle. 2. Influence on seasonal soil moisture and evapotranspiration J. Geophys. Res. Earth Syst. Sci. 117 D15114

Ning T, Li Z, Peng Q, Chen W and Li Z 2020 Effects of forest cover change on catchment evapotranspiration variation in China Hydrol. Process. 34 2219–28

Pan S et al 2002 Evaluation of global terrestrial evapotranspiration using state-of-the-art approaches in remote sensing, machine learning and land surface modeling Hydrol. Earth Syst. Sci. 4 1485–509

Peng S, Ciais P, Maignan F, Li W, Chang J, Wang T and Yue C 2017 Sensitivity of land use change emission estimates to historical land use and land cover mapping Glob. Biogeochem. Cycles 31 626–43

Piao S, Friedlingstein P, Ciais P, Noblet-Ducoudre N D, Labat D and Zaehle S 2007 Changes in climate and land use have a
larger direct impact than rising CO\textsubscript{2} on global river runoff trends. \textit{Proc. Natl Acad. Sci.} 104 15242

Pitman A J \textit{et al} 2009 Uncertainties in climate responses to past land cover change: first results from the LUCID intercomparison study. \textit{Geophys. Res. Lett.} 36 171–83

Ponce-Campos G E \textit{et al} 2013 Ecosystem resilience despite large-scale altered hydroclimatic conditions. \textit{Nature} 494 349–52

Poulter B, Ciais P, Hodson E, Lischke H, Maignan F, Plummer S and Zimmermann N E 2011 Plant functional type mapping for earth system models. \textit{Geosci. Model Dev.} 4 993–1010

Smith B, Wårlind D, Arneth A, Hickler T, Leadley P, Siltberg J and Zachle S 2014 Implications of incorporating N cycling and N limitations on primary production in an individual-based dynamic vegetation model. \textit{Biogeosciences} 11 2027–54

Sterling S M, Ducharme A and Polcher J 2013 The impact of global land-cover change on the terrestrial water cycle. \textit{Nat. Clim. Change} 3 385–90

Wang Q, Cheng L, Zhang L, Liu P, Qin S, Liu L and Jing Z 2021 Quantifying the impacts of land-cover changes on global evapotranspiration based on the continuous remote sensing observations during 1982–2016. \textit{J. Hydrol.} 598 126231

Whitley R \textit{et al} 2017 Challenges and opportunities in land surface modelling of savanna ecosystems. \textit{Biogeosciences} 14 4711–32

Wullschleger S D, Epstein H E, Box E O, Euskirchen E S, Goswami S, Iversen C M, Kattge J, Norby R J, van Bodegom P M and Xu X 2014 Plant functional types in Earth system models: past experiences and future directions for application of dynamic vegetation models in high-latitude ecosystems. \textit{Ann. Bot.} 114 1–16

Yang D, Shao W, Yeh P J-F, Yang H, Kanae S and Oki T 2009 Impact of vegetation coverage on regional water balance in the nonhumid regions of China. \textit{Water Resour. Res.} 45 W06A14

Zhang H, Pak B, Wang Y P, Zhou X, Zhang Y and Zhang I 2013 Evaluating surface water cycle simulated by the Australian community land surface model (CABLE) across different spatial and temporal domains. \textit{J. Hydrometeorol.} 14 1119–38

Zhang L \textit{et al} 1999 Predicting the effect of vegetation changes on catchment average water balance. \textit{Catchment Hydro.} 12 1–35

Zhang L, Dawes W R and Walker G R 2001 Response of mean annual evapotranspiration to vegetation changes at catchment scale. \textit{Water Resour. Res.} 37 701–8

Zhang L, Zhao F, Chen Y and Dixon R N M 2011 Estimating effects of plantation expansion and climate variability on streamflow for catchments in Australia. \textit{Water Resour. Res.} 47 W12539

Zhang Y, Leuning R, Chiew F H S, Wang E, Zhang L, Liu C, Sun F, Peel M C, Shen Y and Jung M 2012 Decadal trends in evaporation from global energy and water balances. \textit{J. Hydrometeorol.} 13 379–91

Zhou G \textit{et al} 2015 Global pattern for the effect of climate and land cover on water yield. \textit{Nat. Commun.} 6 1–9