A comprehensive evaluation of hydrological processes in a second-generation dynamic vegetation model

Zhou, Hao; Tang, Jing; Olin, Stefan; Miller, Paul A.

Published in:
Hydrological Processes

DOI:
10.1002/hyp.15152

Publication date:
2024

Document version
Publisher's PDF, also known as Version of record

Document license:
CC BY

Citation for published version (APA):
A comprehensive evaluation of hydrological processes in a second-generation dynamic vegetation model

Hao Zhou1, Jing Tang1,2, Stefan Olin1, Paul A. Miller1

1Department of Physical Geography and Ecosystem Science, Lund University, Lund, Sweden
2Center for Volatile Interactions, Department of Biology, University of Copenhagen, Copenhagen, Denmark

Correspondence
Paul A. Miller, Department of Physical Geography and Ecosystem Science, Lund University, Lund SE22362, Sweden.
Email: paul.miller@natuko.lu.se
Jing Tang, VOLT – Center for Volatile Interactions, Universitetsparken 15, DK-2100 Copenhagen, Denmark.
Email: jing.tang@bio.ku.dk

Funding information
Svenska Forskningsrådet Formas, Grant/Award Number: 2016-01580; China Scholarship Council, Villum Young Investigator, Grant/Award Number: VILS3048; Danish National Research Foundation within the Center for Volatile Interactions, Grant/Award Number: DNRF168

Abstract
The global water and carbon cycles are greatly influenced by terrestrial vegetation, making trustworthy representations of dynamic biosphere–hydrosphere interactions a crucial component of both ecosystem and climate models. This paper comprehensively evaluates the hydrological performance of a leading dynamic global vegetation model Lund-Potsdam-Jena General Ecosystem Simulator (LPJ-GUESS), using a broad range of the latest available global observation-based gridded datasets that cover the main components of the hydrological cycle. Overall, we find that the hydrological components modelled by LPJ-GUESS agree well with global gridded datasets of runoff, evapotranspiration and surface soil moisture, though there are discrepancies in some regions and periods. Furthermore, LPJ-GUESS accurately captures both inter- and intra-annual variations of runoff in most regions and catchment areas, including the Danube, Murray, Yangtze, Yenisei and Nile basins. Total evapotranspiration modelled by LPJ-GUESS agrees closely with the evapotranspiration estimates of the Global Land Evaporation Amsterdam Model and PML-V2 datasets, but with some disagreement in the individual components, especially for evaporation. The surface soil moisture simulated by LPJ-GUESS aligns with ESA-CCI (v5.3) surface soil moisture datasets in most regions, with greatest discrepancies in subarctic areas. We attribute these discrepancies to two main sources: (1) absent or poor representation of processes such as river routing, storage and supply of water bodies, and cropland irrigation; and (2) uncertainties in both reference datasets and input to the model, including precipitation, soil texture, and land use.

KEYWORDS
dynamic global vegetation model, evapotranspiration, hydrological processes evaluation, observation-based global gridded datasets, runoff, surface soil moisture

1 | INTRODUCTION

Monitoring and modelling hydrological processes and their responses to global environmental change are necessary to quantify global water availability, to forecast agricultural water supply, and to perform analyses of flood and drought risks (Gerten et al., 2005; Murray et al., 2011; Rost et al., 2008). As a result of climate change, the global hydrological cycle is expected to undergo substantial changes, with high interannual, seasonal and regional variability (Allan et al., 2020; Douville et al., 2021). Vegetation and hydrological processes interact at many different scales, from the molecular to the global (Fatichi et al., 2016). On the one hand, the plant-related water fluxes and root
zone soil moisture greatly influence many biophysical and biogeochemical processes, such as vegetation water stress, nutrient leaching, litter decomposition, dissolved organic and inorganic carbon transfer, fire disturbance, vegetation growth and turnover (Horvath et al., 2021; Li et al., 2011; Sitch et al., 2003). On the other, key hydrological fluxes and states such as evapotranspiration (ET), soil moisture content and runoff generation are critically dependent on the structures and dynamics of plant communities (Brown et al., 2010; Gerten et al., 2004). For example, runoff is significantly affected by plant-related processes and factors such as canopy interception (Tao et al., 2020), rooting strategy (Collins & Bras, 2007), leaf area (Zhai & Tao, 2021), phenology (Chen et al., 2022), and transpiration (Li et al., 2020). Likewise, afforestation and the recovery of forests from disturbances could reduce runoff by increasing ET, whereas deforestation often increases runoff (Sun et al., 2006). Thus, it is necessary to simultaneously simulate the interactions between vegetation dynamics and hydrological processes, especially as the climate changes.

Dynamic global vegetation models (DGVMs), as a key component of Earth System Models (ESMs), are widely applied for their ability to simulate vegetation distributions and ecosystem carbon, nutrient, and water fluxes (Bierkens, 2015; Murray et al., 2012). Previous studies have shown that DGVMs perform well in comparison to stand-alone hydrological models when simulating runoff and ET across large areas (Gerten et al., 2004). In coupling the carbon and water cycles, DGVMs explicitly consider the complex feedback mechanisms between vegetation growth, carbon uptake, and water use, thereby including a potentially more accurate representation of the impact of vegetation on hydrological processes (Best et al., 2011; Schaphoff et al., 2018). By simulating the growth, mortality, turnover, phenology and establishment of vegetation as represented by diverse plant functional types (PFTs), DGVMs can also capture the effects of vegetation cover and biomass changes on water fluxes (e.g., canopy interception and plant transpiration, runoff) (Horvath et al., 2021). Furthermore, the root-water uptake dynamics in DGVMs is a critical component of the terrestrial water balance, affecting soil moisture and groundwater recharge (Murray et al., 2011). In addition, DGVMs are increasingly being developed into more sophisticated tools by incorporating more land surface processes (e.g., nutrient cycling and their feedbacks), and more detailed structural representations (e.g., sub-grid topographical heterogeneity), which adds greater parameter and structural uncertainty (Martín Belda et al., 2022). A better understanding of such uncertainties requires us to assess overall model performance, and as much as possible to confront individual modelled processes with a varied array of observational datasets (Gerten et al., 2004; Rost et al., 2008).

More specifically in regard to hydrological processes, the availability of large-scale observational or observation-derived hydrology-related datasets has significantly increased over the last few years, providing valuable resources for improving DGVMs. These range from machine learning-derived datasets, reanalysis data and remote sensing-based datasets. For instance, Ghiggi et al. (2019) utilized station-measured discharge datasets from 35,002 sites and a machine learning method to create a global gridded runoff dataset valid from 1902 to 2014, and Martens et al. (2017) derived a global terrestrial evaporation and root-zone soil moisture dataset valid from 1980 to 2015 based on satellite and climate data.

This study adopts this approach and aims to evaluate the hydrological cycle components simulated by a leading DGVM—the Lund-Potsdam-Jena General Ecosystem Simulator model (LPJ-GUESS) (Smith et al., 2014)—through simultaneous cross-comparison with a wide array of large-scale observational datasets. Modelled monthly and annual runoff, ET, and surface soil moisture (SSM) were evaluated using observation-based datasets at the basin, ecoregion and global scale, and from seasonal to multidecadal timescales, in order to better assess model performance at the process level, and to identify model processes in need of improvement.

2 | METHODS

2.1 | LPJ-GUESS description

LPJ-GUESS is a climate-driven dynamic ecosystem model, with explicit processes describing vegetation dynamics, land-atmosphere carbon exchange, water exchanges, soil biogeochemistry, land use, land management and wetlands. Here, only a brief overview will be provided as it is described in detail elsewhere (Sitch et al., 2003; Smith et al., 2014). The code is available at https://zenodo.org/record/8070308. In LPJ-GUESS, vegetation establishment, mortality and disturbance are implemented as stochastic processes on replicates patches (default 15 patches) representing samples from a landscape. This leads to different dynamics in different patches, which stimulate more realistic resource competition for light, interactions between the 20 plant functional types (PFTs) (Appendix Tables A1, A2 and A3), included to capture the diversity of plant structure and function globally (Åhström et al., 2012; Smith et al., 2014). Temperature and water stress thresholds based on Smith et al. (2014) determine daily the leaf phenology of summer green and rain green PFTs, which shed their leaves in winter or during the dry season, respectively. Remaining tree PFTs do not shed their leaves and have an evergreen phenology that has leaf longevity (life span) of more than a year. Net primary production is determined by subtracting autotrophic respiration from gross primary production calculated using a coupled photosynthesis–water balance scheme operating with a daily timestep, accounting for nitrogen availability (Smith et al., 2014). Annual NPP is allocated to three pools for the production of new tissues: leaves, sapwood, and fine roots. The litter contains carbon and nitrogen from decaying leaves and roots while the decomposition of soil organic matter and litter is driven by temperature and moisture level in the upper 50 cm of soil (Smith et al., 2014). At the end of each year, a PFT-specific mortality rate is derived based on temperature, growth efficiency, carbon deficit, light competition, or a breach of bioclimatic limits referring to Smith et al. (2014).

LPJ-GUESS performance has been evaluated in numerous studies and the model has been shown to capture well estimates of global.
terrestrial carbon and nitrogen pools and fluxes, as well as vegetation dynamics (Olin, Schurgers, et al., 2015; Sitch et al., 2003; Smith et al., 2014).

2.1.1 Hydrological processes

Since the previous hydrology module description by Gerten et al. (2004), there have been some updates to the hydrological processes in LPJ-GUESS and here we will provide a brief overview. The schematic representation of the water balance computed for each grid cell in LPJ-GUESS is shown in Appendix Figure B1. One of the biggest differences in the new version is the greater vertical resolution of the soil scheme. The model now has 15 soil layers with each layer being 0.1 m thick rather than the previous version with two layers of thickness of 0.5 and 1.0 m (Gerten et al., 2004). The availability of water for plant growth is based on storage and flow within the 15 layers of the soil profile. The precipitation input is interpolated daily as rain or snow depending on the temperature; when temperatures are below 0°C, precipitation falls as snow. When the rain encounters the plants, interception loss occurs (Appendix Figure B2) and the remaining rain falls through vegetation as input to the soil layers. Transpiration of vegetation is influenced by available water above a wilting point determined by the soil properties, PFT-specific canopy conductance and CO2 concentrations. Uptake by plants is partitioned according to the PFT-specific fraction of roots situated in each layer. Trees and shrubs are usually assumed to have a larger proportion of their roots in the lower soil layers, while herbaceous plants are shallow-rooting. The PFT-specific root fractions for each 10 cm soil layer, parameterized following Jackson et al. (1996) are shown in Appendix Figure A4. This feature affects the relative performance of woody and herbaceous PFTs in dry environments differing in the seasonality of rainfall. On the fraction of the grid cell not covered by vegetation, daily evaporation occurs from bare soil. The two soil layers nearest the surface representing the top 0.2 m of soil are distinguished to calculate soil evaporation. Actual ET in every patch is calculated by adding up interception loss, transpiration for all PFTs, and evaporation from bare soil. Furthermore, irrigation for agriculture is calculated according to the water deficiency of cropland vegetation in the model (Olin, Lindeskog, et al., 2015; Smith et al., 2014). Water input (from rain or snowmelt) to a saturated soil layer is lost as runoff. Additional depletion of soil water may occur through percolation beyond the soil layers (0–1.5 m) and out of reach by plant roots. The water content of soil layers is updated daily, considering the processes of interception, throughfall, snowpack change, root water uptake, evaporation, percolation, and runoff. To confirm that water balance is maintained in the model, we calculated whether the incoming water volume is equal to the outgoing water volume in each gridcell that is, whether precipitation and irrigation is equal to the sum of runoff, ET, soil moisture change, and snow change. The water balance check results are shown in Appendices B3 and B4. For the water balance check of reference datasets, we compare the precipitation and the sum of runoff and ET in Appendices B5 and B6.

2.1.2 Model input

In this study, we ran LPJ-GUESS for the period 1901–2015, following a 1000-year spinup period to reach an equilibrium between climate and carbon pools (Sitch et al., 2003) in 1901. Climate data including gridded monthly air temperature, humidity, precipitation rate, incoming solar radiation, surface pressure, and surface winds (0.5° × 0.5°) from the CRU-NCEP Version 7 database are utilized to drive the simulations (Viovy, 2018). Although CRU-NCEP has a reliance on meteorological station data quality and density, it includes multiple climate variables with comprehensive global land surface coverage. Beck et al. (2017) found that there are some precipitation datasets that consistently outperform others in specific regions, but no one dataset performs best in all regions. Thus, even though a comparison of different precipitation input datasets is not the key point in our study, here we wish to highlight the differences by directly comparing spatial differences of the precipitation datasets in Appendices D2, E1 and F2. The weather generator GWGEN (Global Weather GEnerator) (Sommer & Kaplan, 2017) was introduced to offer downscaled daily weather data derived from monthly sums or averages for various parameters, including precipitation, maximum and minimum temperature, net downward shortwave solar radiation, and 10 m horizontal wind, utilizing a Markov-Chain approach. The soil properties (sand, silt, clay fractions) are based on the top layer of the WISE30min 3.0 (Wide-field Infrared Survey Explorer, version 3) dataset (Batjes, 2005) and determine the hydrological properties and thermal diffusivity of the soil, as described by Olin, Schurgers, et al. (2015).

Annual CO2 concentrations derived from ice-core measurements and atmospheric observations, are provided by McGuire et al. (2001) and the TRENDS project (http://cdiac.esd.ornl.gov/). Monthly N-deposition rates, varying by decade and across the global 0.5-degree grid were taken from Lamarque et al. (2013). Global land use data between 1901 and 2006 are obtained from Hurr et al. (2020). The fire regime is derived from the SIMFIRE-BLAZE fire model (Knorr et al., 2014; Rabin et al., 2022).

2.2 Evaluation data and method

In this study, multi-source observation datasets are selected to evaluate the different hydrological components of LPJ-GUESS including runoff, ET, and soil moisture. The selected datasets are widely-used and often have suitable temporal coverage (1985–2014) and resolution (daily to monthly), as well as adequate spatial coverage (globally) and resolution (≤0.5°). These datasets are observation-based, either based on site-level data or remote-sensing images. Here, we provide a brief description of all the datasets used, which are listed in Table 1.

2.2.1 Runoff

A global gridded dataset of monthly runoff for evaluation is taken from the Global Runoff Reconstruction (GRUN) dataset (Ghiggi
TABLE 1  Datasets used in this study for model evaluation.

<table>
<thead>
<tr>
<th>Type/variable</th>
<th>Dataset</th>
<th>Source</th>
<th>Period</th>
<th>Resolution (degrees)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Runoff</td>
<td>GRUN</td>
<td>Derived from multiple sources</td>
<td>1985-2014</td>
<td>0.5 × 0.5</td>
<td>Ghiggı et al. (2019)</td>
</tr>
<tr>
<td></td>
<td>LandFlux-EVAL (only yearly mean)</td>
<td>Merging from multiple sources</td>
<td>1989-2005</td>
<td>0.5 × 0.5</td>
<td>Mueller et al. (2013a)</td>
</tr>
<tr>
<td></td>
<td>GLEAM (monthly)</td>
<td>Derived with remote sensing data</td>
<td>1985-2014</td>
<td>0.25 × 0.25</td>
<td>Martens et al. (2017)</td>
</tr>
<tr>
<td></td>
<td>PML-V2 (daily)</td>
<td>Derived with remote sensing data</td>
<td>2003-2014</td>
<td>0.05 × 0.05</td>
<td>Zhang (2019)</td>
</tr>
<tr>
<td>SSM</td>
<td>ESA-CCI (v5.3) (monthly)</td>
<td>Merging from satellite products</td>
<td>1985-2014</td>
<td>0.5 × 0.5</td>
<td>Dorigo et al. (2017)</td>
</tr>
<tr>
<td>Others</td>
<td>Ecoregions map</td>
<td>Distribution of distinct assemblages of species and communities</td>
<td>-</td>
<td>-</td>
<td>Dinerstein et al. (2017)</td>
</tr>
<tr>
<td></td>
<td>STN-30p</td>
<td>Basin scope</td>
<td>-</td>
<td>0.5 × 0.5</td>
<td>Fekete et al. (2001)</td>
</tr>
<tr>
<td></td>
<td>Global DEM data of 5 km</td>
<td>Elevation data derived from multiple sources</td>
<td>-</td>
<td>5 km</td>
<td>Amatullı et al. (2018)</td>
</tr>
<tr>
<td></td>
<td>GLIMS glacier map</td>
<td>Glacier extent based on satellite measurement</td>
<td>-</td>
<td>150 m</td>
<td>GLIMS (2018)</td>
</tr>
<tr>
<td></td>
<td>GLWD wetland map</td>
<td>Wetland extent merged from multiple sources</td>
<td>-</td>
<td>1 km</td>
<td>Lehner and Döll (2004)</td>
</tr>
<tr>
<td></td>
<td>GFSA crop mask</td>
<td>Irrigated and rainfed cropland cover</td>
<td>-</td>
<td>-</td>
<td>Thenkabail et al. (2016)</td>
</tr>
<tr>
<td>Primary drivers map of forest cover loss</td>
<td>Utilizing the remote sensing images</td>
<td>-</td>
<td>0.5</td>
<td>Curtis et al. (2018)</td>
<td></td>
</tr>
</tbody>
</table>

To evaluate LPJ-GUESS ET output, the LandFlux-EVAL benchmark products are used (Mueller et al., 2013). LandFlux-EVAL provides annual mean ET synthesis products for the period 1989-2005, merging three categories of data sources: (1) satellite and in situ observations, (2) modelled ETs from land-surface models (LSMs) driven with observation-based forcing, and (3) outputs from atmospheric reanalysis (i.e., computed with LSMs within a global model assimilating most atmospheric observations). The merged annual mean ET synthesis products are based on a total of 14 data sets in them. The main component of ET of the Global Land Evaporation Amsterdam Model (GLEAM) land cover model is the MERAA-land model. The specific ET product available from the Global Land Evaporation Amsterdam Model (GLEAM) product includes 5 LSMs: VIC, ET-ORCH, CRU-ORCH, GL-NOAHFR, and GL-NOAHFF. The GLEAM product assimilates mostly atmospheric observations, although the uncertainty analysis and the importance of precipitation and land cover types (Mueller et al., 2013). To evaluate the individual components of ET (i.e., soil evaporation, interception loss, and plant transpiration), we have used two additional ET products. The latitudinally-averaged differences in the additional ET products are then compared with that of LPJ-GUESS. The GRUN dataset, through the validation and cross-comparison with global hydrological models (GHMs) and observational data, exhibits its reliability and accuracy for global and basin-scale hydrological analyses with the relative bias of 0.047, the Nash-Sutcliffe efficiency (NSE) of 0.525, the squared correlation coefficient (R²) for globe scales, compared to many global hydrological models (Ghiggi et al., 2019). The GRUN dataset, through the validation and cross-comparison with global hydrological models (GHMs), still performs satisfactorily at scales from grid cell to basin scales, compared to many global hydrological models (Ghiggi et al., 2019).
and a daily time step with a spatial resolution of 0.25°, based on satellite-observed soil moisture, vegetation optical depth and snow-water equivalent, reanalysis air temperature and radiation, and a multi-source precipitation product. The approach aims to optimize the utilization of satellite observations of climatic and environmental variables to enhance the accuracy of ET estimation. Uncertainties in the GLEAM dataset’s ET estimates have been quantified by Martens et al. (2017), who report an unbiased root mean square difference (ubRMSE) of 0.71 mm day\(^{-1}\) and correlation coefficients averaging 0.79 against in situ measurements over the validation period 2011–2015, indicating the model’s reliable performance. Penman-Monteith-Leuning Evapotranspiration V2 (PML-V2) products include ET and its three components at 500 m and 8-day resolution during 2002–2017 and with a spatial range from \(-60°\)S to \(90°\)N (Zhang, 2019; Zhang et al., 2019). The products provide estimates of gross primary production (GPP), and three components of ET: transpiration from vegetation, evaporation and evaporation of intercepted rainfall from vegetation. PML-V2 is based on MODIS data (leaf area index, albedo, and emissivity) together with GLDAS meteorological forcing data as model inputs (Zhang, 2019; Zhang et al., 2019). The PML-V2 dataset achieves a root mean square error (RMSE) of 0.73 mm day\(^{-1}\) and a bias of \(-3\%\), with a coefficient of determination (\(R^2\)) of 0.69, across the 95 flux tower sites used for validation, demonstrating its robustness and accuracy across various metrics (Zhang et al., 2019). GLEAM and PML-V2 are both remote sensing based datasets and have been used to evaluate hydrological models in other studies (He et al., 2022; Jin & Jin, 2020). GLEAM ET has longer temporal coverage and finer temporal resolution while PML-V2 has better spatial resolution (500 m). GLEAM has strengths in modelling land surface processes but may incorporate model assumptions that introduce uncertainty. PML-V2 may be more trusted in applications requiring high spatial detail and frequent monitoring, such as agriculture, forestry, or urban studies. The check of observation based datasets shows that the GSWP3-GRUN-PML-V2 group has good water balance consistency and GSWP3-GRUN-GLEAM has comparably worse but acceptable water balance (Appendices B5 and B6).

2.2.3 | Soil moisture

Global gridded surface soil moisture (SSM) reference datasets are provided by the Soil Moisture CCI project (ESA-CCI SM v5.3) (Dorigo et al., 2017; Lee & Veizer, 2003). The Soil Moisture CCI project is part of the ESA Programme on the Climate Change Initiative (CCI), initiated in 2010 and produces an updated soil moisture product every year. Based on remote sensing, CCI Soil Moisture produces an annual global climate data record of soil moisture spanning over 40 years with 3 separate soil moisture products derived from active, passive and combined (active-passive) sensors. The ESA Soil Moisture dataset has employed triple collocation analysis and a weighted average approach based on signal-to-noise ratios (SNR) to reduce random errors (lower than 0.02) (Dorigo et al., 2017). Because the penetration properties limitation of remote sensing light waves, the satellite product can only capture the surface soil moisture (usually top 5 cm). In this study, the combined products are only used to evaluate the surface soil moisture of LPJ-GUESS (i.e. the uppermost 0.1 m layer).

2.2.4 | Other ancillary data

Ecoregions are related to the biome, climate, elevation and even the precipitation, each of which is key to the hydrological process and influences the hydrological components (Dinerstein et al., 2017). The ecoregions2017 map (Dinerstein et al., 2017) includes 846 ecoregions globally which can be aggregated into 14 broader categories or terrestrial biomes (three ecoregions are neglected as lack of observation data or too small). The remaining 11 categories are Tropical Subtropical Moist Broadleaf Forests (TSMBF), Tropical Subtropical Coniferous Forests (TSCF), Tropical & Subtropical Grasslands, Savannas & Shrublands (TSGSS), Tropical & Subtropical Dry Broadleaf Forests (TSDBF), Temperate Grasslands, Savannas & Shrublands (TSGSS), Temperate Conifer Forests (TFC), Temperate Broadleaf & Mixed Forests (TBM), Montane Grasslands & Shrublands (MGS), Mediterranean Forests, Woodlands & Scrub (MFWS), Boreal Forests Taiga (BFT), Tundra (TU) (Appendix Figure C1). We evaluated the monthly and yearly variability of LPJ-GUESS runoff, ET, and SSM with reference data in each category. We also used basin boundaries from the global Simulated Topological Network (STN-30p) (Fekete et al., 2001). This basin boundary data was used to mask the global gridded runoff into 10 basins for analysis: Amazon, Danube, MacKenzie, Mississippi, Murray, Yenisei, Congo, Nile, Ganges, and Yangtze.

For investigating the model performance impacted by factors such as topography or missing processes, the following data are used: The Global Digital Elevation Model(DEM) data of 5 km resolution is from Amatulli et al. (2018). The glacier map is from the Global Land Ice Measurements from Space (GLIMS) using data from satellite observations (GLIMS, 2018). The wetland map is from the Global Lakes and Wetlands Database (GLWD), which includes large lakes and reservoirs (Lehner & Döll, 2004). The Cropland map is from the Global Food Security Support Analysis Data (GFSAD) Crop Mask Global 1-kilometre dataset which provides the spatial distribution of irrigated crops and rained crops (Thenkabail et al., 2016). Primary drivers of forest cover loss for the period 2001–2015 are provided by Curtis et al. (2018).

2.2.5 | Accuracy estimation

Modelled annual and monthly hydrological components were compared with the observation-based data to evaluate LPJ-GUESS performance. The ratio of annual mean hydrological fluxes of LPJ-GUESS and observed datasets for the period 1985–2014 to quantify the variations are shown in Appendices D3, D4 and D5. The R-square value and regression line slope and intercept were computed to evaluate the linear relationship between modelled and observed hydrological components. Furthermore, to derive the statistical goodness of fit of
the estimates of the hydrological components, we employed Taylor diagrams to illustrate the standard deviation, Pearson correlation coefficient (PCC) and the root mean square deviation (RMSD) of the model output. All standard deviations of different regions are normalized (i.e. standard deviation of the model over a region divided by that of the observations). The smaller the normalized standard deviations are, the smaller the variation of the model data over the region. The closer a point on the Taylor diagram is to the observation point (where PCC = 1 and RMSD = 0), the better it matches the reference dataset in terms of standard deviation, correlation and RMSD.

3 | RESULTS

3.1 | Runoff

The evaluations of LPJ-GUESS modelled runoff have been conducted at global, ecoregions and river basin levels.

3.1.1 | Global evaluation

The mean values of annual runoff in LPJ-GUESS agree well with observation-based datasets GRUN for the period 1985–2014 in most regions (Figure 1b and Appendix D3). Figure 1c illustrates the latitude-averaged differences in runoff which range from $-1548.62$ to $133.50$ mm year$^{-1}$, revealing that LPJ-GUESS and GRUN are in good agreement, especially in the northern hemisphere. However, LPJ-GUESS tends to overestimate the runoff between $10^\circ$N and $20^\circ$N, while GRUN exhibits larger averaged runoff near the equator ($10^\circ$S–$10^\circ$N) and notably at the most southern part of South America ($60^\circ$S) (Figure 1c). The large underestimations only occur at very few grid cells with extremely high runoff in GRUN as the latitude-summed differences (Figure 1d) show not extreme value ranging from $-104.94$ to $65.98$ km$^3$ year$^{-1}$ (multiplying with the grid-cell area).

The model performance in estimating annual mean runoff (Figure 1b; Appendix Figures D1 and D3) can be roughly categorized into five zones: (1) The model outputs demonstrate good agreement with observations for the temperate regions and large areas of central Asia and northern America. (2) Overestimations are observed in semi-arid and arid regions, particularly in northern Africa, and parts of southern America. (3) Overestimations occur in crop-intensified areas such as Eastern Europe, eastern America and southeastern China, and India (Appendix Figure J4). (4) Underestimations are frequent in sub-arctic regions such as northwestern Canada and Scandinavia. (5) Underestimations occur in areas of deforestation and shifting agriculture such as the Congo basins and the Amazon basin (approximately 66.24% and 44.85% of the land area in Congo and Amazon

![Figure 1](image-url)
basin experienced forest loss driven by the above two factors from 2001 to 2015) (Appendix Figure J5).

3.1.2 | River basin evaluation

The model performance in simulating runoff varies across different river basins. LPJ-GUESS captures both interannual and seasonal variations of runoff for the Danube, Murray, Yangtze, Yenisei and Nile basins (Figure 2; Figure 3), but underestimates annual mean runoff in the Amazon (−17.1%), Congo (−34.6%) and MacKenzie (−40.4%) basins, and overestimates the runoff in the Ganges (17.9%) and Mississippi (24.7%) basins. The underestimations of monthly mean runoff occur almost throughout the year in deforestation and shifting agriculture basins (Amazon [from March to November] and Congo [whole year]). The modelled runoff of subarctic basins (the MacKenzie and Yenisei) at January and February is close to 0, slightly lower than the observed value, while the modelled runoff is much lower than the observed in the autumn/winter months. For cropland-dominated basins (Ganges and Nile), the overestimations mainly occur at high precipitation periods. For the Yangtze basin, the model underestimates runoff in its upstream area (western part) and overestimates runoff in the downstream area (eastern part) (Appendix Figure D1).

Figure 2 also shows the link between monthly mean runoff and ET from 2003 to 2024. Using PML-V2 as the ET reference for water balance consistency, the results show that the overestimated runoff can be partly explained by the underestimated ET in April and June in the Yenisei catchment and April to June in the Ganges catchment. The underestimated runoff can be partly linked with the overestimated ET during April to June in the Danube catchment, from October to May in the Congo, from September to December in the Ganges, and from July to December in the Yangtze catchment.

As shown in Figure 3, the modelled runoff captures the significant (p < 0.05) trends of three basins: Amazon (modelled trend: 5.29 km³ year⁻¹, observed trend: 3.45 km³ year⁻¹), Ganges (modelled: −1.66 km³ year⁻¹, observed: −0.79 km³ year⁻¹), Yenisei (modelled: 0.4 km³ year⁻¹, observed: 0.53 km³ year⁻¹). The modelled runoff has high correlation index (>0.50) with the observed runoff at most catchments such as Murray (0.90), Mississippi (0.88) and Danube (0.85). There are no significant trends in modelled runoff and observed runoff in MacKenzie, Murray, Danube, and Mississippi basins. In Congo and Yangtze basins, the modelled runoff does not capture the increasing trend in the observed runoff. In the Nile basin, the model runoff has an opposite runoff trend (0.96 km³ year⁻¹), compared to the observed trend (−0.44 km³ year⁻¹). The Appendix Figure E1 further shows the both significant precipitation trends of
CRU-NCEP (LPJ-GUESS used) and GSWP3 (GRUN used) in Amazon, Mackenzie, Ganges, Mississippi, Yenisei. The CRU-NCEP precipitation has significant trends in Amazon (5.62 km$^3$ year$^{-2}$), Ganges (1.65 km$^3$ year$^{-2}$), Yenisei (0.78 km$^3$ year$^{-2}$) basins, which are in line with the simulated runoff trends. The comparison of results of the two precipitation datasets (Appendices D2 and E1) shows that the annual mean differences between CRU-NCEP and GSWP3 range from 125 to 125 mm year$^{-1}$ between 1985 and 2010 in most regions and their time series have a high correlation index (>0.67) in all selected basins (km$^3$ year$^{-2}$).

3.1.3 | Ecoregions evaluation

The annual mean runoff ratio (modelled compared with the observed) varies in different ecoregions (Appendix Figure G1). The modelled runoff seems to have larger monthly variability in almost all ecoregions when compared with the observed (see the runoff ratio in Appendix Figure H1). LPJ-GUESS seems to underestimate annual mean runoff in some subarctic regions such as Boreal Forests & Taiga (BFT) and Tundra (TU) areas (Appendix Figure G1) as these areas have quite lower (ratio <0.3) runoff estimates than the observed in winter time from December to February (Appendix Figure H1, Figure 2). Similarly, for Montane Grasslands & Shrublands (MGS) located mostly in the Tibetan Plateau and the Andes (Appendix Figure C1), the model underestimates the runoff from September to March (ratio <0.5), and the model underestimates the runoff for the almost the whole year except for March and April for Temperate Conifer Forests (TCF) areas (Appendix Figure H1).

LPJ-GUESS also underestimates runoff in Tropical Subtropical Coniferous Forests (TSCF) area (ratio <0.5) from November to April but overestimates runoff from May to July. This combined effect results in a similar magnitude in the annual mean runoff of TSCF with observations. Modelled annual and monthly mean runoff in Tropical Subtropical Moist Broadleaf Forests (TSMBF) area is similar to observations, but it is evident from the spatial distribution of TSMBF that the model underestimates runoff in the Amazon and Congo basins and overestimates runoff in south China and southeastern Asia.

In addition, LPJ-GUESS overestimates annual mean runoff at Tropical & Subtropical Grasslands, Savannas & Shrublands (TSGSS), Tropical & Subtropical Dry Broadleaf Forests (TDBBF), Temperate Grasslands, Savannas & Shrublands (TGSS) area. The main overestimations occur separately for TSGSS (July–August), TDBBF (June–August), and TGSS (March–April).

To summarize modelled runoff evaluation results, we compared global and ecoregional annual mean runoff with observations using a Taylor diagram (Figure 4a). Compared with the GRUN runoff observation point (normalized deviation for globe and all ecoregions), the model (global and across ecoregions) shows comparable normalized standard deviation (0.5–1.5), PCC (>0.6), and RMSD (<1). Across all regions, Temperate Conifer Forests (TCF) have the lowest RMSD and highest PCC, which is consistent with the annual mean runoff spatial
pattern in Figure 1b. Compared to TCF, the global and MFWS also have lower RMSD (<0.5) and higher PCC (>0.8). Temperate Grasslands, Savannas & Shrublands (TGSS) have the highest RMSD and lowest PCC (<0.7), making this ecoregion with the poorest model performance. Except for the Tropical Subtropical Coniferous Forests (TSCF), Tropical & Subtropical Grasslands, Savannas & Shrublands (TSGSS), Tropical & Subtropical Dry Broadleaf Forests (TSDBF), the remaining ecoregions have smaller standard deviation (<1) than the GRUN data, which shows the model has lower spatial variability than observations in these regions.

3.2 Evapotranspiration

The mean annual ET modelled by LPJ-GUESS agrees well with observation-based datasets GLEAM for the period 1985–2014 in most regions, such as the north of Asia and Europe, as well as Eastern Canada and America (Figure 1f and Appendix D4). Compared to the GLEAM datasets, the model overestimates annual mean ET in southeastern China and northern India as well as the southeastern part of South America but underestimates ET in central South America (Amazon basin), south of Africa (Congo basin), western Canada, northeastern Asia and Tibet area.
The latitude-averaged ET modelled by LPJ-GUESS agrees closely with that of GLEAM and PML-V2 during 2003–2018 (Figure 5). PML-V2 has higher ET between 35°S and 90°N compared to the other two estimates while LPJ-GUESS overestimates ET between 35°S and 60°S compared to GLEAM. The GLEAM dataset seems to have the lowest ET averaged by latitude. The modelled individual components of ET that is, transpiration, interception, and evaporation—have different degrees of agreement with observation-based datasets. Compared to GLEAM and PML-V2, LPJ-GUESS overestimates transpiration and underestimates interception at almost all latitudes, but LPJ-GUESS captures the latitudinal variability of the two components of GLEAM and PML-V2. In terms of evaporation, LPJ-GUESS and GLEAM have similar estimated magnitudes, although LPJ-GUESS has higher estimates than GLEAM between two ranges 10°S–10°N and 45°N–90°N. However, PML-V2 has much higher evaporation estimates than LPJ-GUESS and GLEAM.

As shown in Appendix Figure H2, the monthly ET ratio (LPJ-GUESS/GLEAM) has a large monthly variability in some ecoregions such as Tundra (TU), Temperate Conifer Forests (TCF), and Boreal Forests Taiga (BFT). From Appendix Figure G2, LPJ-GUESS has underestimated the annual mean ET in Tundra (TU) and Temperate Conifer Forests (TCF), with a ratio <0.6 from November to January. Modelled ET in Boreal Forests Taiga (BFT) is equal to the reference data in the annual mean since the model underestimates ET from November to January while it considerably overestimates ET (ratio >1.8) in March. LPJ-GUESS underestimates ET in Montane Grasslands and Shrublands (MGS) for most of the year and Tropical Subtropical Coniferous Forests (TSCF) from November to April.

In contrast, LPJ-GUESS overestimates annual mean ET in Mediterranean Forests, Woodlands & Scrub (MFWS) (January–April) and Tropical and Subtropical Dry Broadleaf Forests (TDBF) (May–October).

LPJ-GUESS ET matches the reference data in the Temperate Broadleaf & Mixed Forests (TBMF), Temperate Grasslands, Savannas & Shrublands (TGSS), Tropical Subtropical Moist Broadleaf Forests (TSMBF) and Tropical & Subtropical Grasslands, Savannas & Shrublands (TSGSS) ecoregions.

To summarize modelled ET evaluation results, we compared global and ecoregional annual mean ET with the observation-based datasets GLEAM and PML-V2 using a Taylor diagram (Figure 4b,c). Compared with the GLEAM observation point (normalized deviation for the globe and all ecoregions), the model shows comparable
normalized standard deviation (0.75–1.5), PCC (>0.65), and RMSD (<0.8). Among all regions, the globe and Tropical & Subtropical Grasslands, Savannas & Shrublands (TSGSS) have lower RMSD and higher PCC than other regions, which agree mostly with reference data. Temperate Grasslands, Savannas & Shrublands (TGG) also have low RMSD (<0.5) and high PCC (>0.9). Tundra (TU) has the highest RMSD (0.79) and lowest PCC (<0.7), making this region the poorest match with observations. Temperate Conifer Forests (TCF), Tropical Subtropical Moist Broadleaf Forests (TSMBF), Mediterranean Forests, Woodlands & Scrub (MFWS) have a smaller standard deviation (<1) than the reference data GLEAM, which shows that the model has lower spatial variability than the reference data in these regions. Despite this variance, modelled ET across various regions exhibits a reasonable degree of agreement with GLEAM data.

Compared with the PMLV2 observations (normalized deviation for the globe and all ecoregions), the model shows comparable normalized standard deviation (<1.0), PCC (>0.70), and RMSD (<0.75), except for the Boreal Forests Taiga (BFT) and Temperate Broadleaf & Mixed Forests (TBMF) ecoregions. In particular, LPJ-GUESS has the largest RMSD (>2) and lowest PCC (<0.1) in the BFT area, which mostly differs from reference data. Across all regions, the Temperate Conifer Forests (TCF) have the lowest RMSD and highest PCC.

3.3 | Surface soil moisture

For gridded surface soil moisture evaluation (Figure 1) and Appendix D5, LPJ-GUESS overestimates annual mean SSM at surface 0.1 m depth in a large part of the arctic and subarctic regions but underestimates SSM in eastern America, southern Africa and in the west and north of Europe, where the runoff is also underestimated. However, the model overestimates SSM at surface 0.1 m depth in the Amazon and Congo basin areas, which is contrary to its underestimation of runoff.

As shown in Appendix Figure H3 the monthly SSM ratio (LPJ-GUESS/ESA) has a small monthly variability in most ecoregions. The model underestimates monthly mean SSM in the Temperate Broadleaf & Mixed Forests (TBMF), Tropical and Subtropical Dry Broadleaf Forests (TSDBF), Tropical and Subtropical Grasslands, Savannas and Shrublands (TSGSS), Tropical Subtropical Coniferous Forests (TSCF) and Tropical Subtropical Moist Broadleaf Forests (TSMBF) ecoregions for most of the year.

In contrast, LPJ-GUESS overestimates monthly mean SSM in the Tundra (TU) and Temperate Grasslands, Savannas & Shrublands (TGG) ecoregions throughout the year.

The model could capture well the observed annual mean SSM in Boreal Forests Taiga (BFT), Mediterranean Forests, Woodlands and Scrub (MFWS), Montane Grasslands and Shrublands (MGS) as well as Temperate Conifer Forests (TCF). However, the model has a large discrepancy in monthly mean SSM in Boreal Forests Taiga (BFT).

Compared with the ESA observations (normalized deviation for the globe and all ecoregions), the model shows low PCC (<0.7) and high RMSD (0.8–2.0) (Figure 4d). Across all regions, Tropical & Subtropical Grasslands and Savannas & Shrublands (TSGSS) has lower RMSD (<1) and higher PCC (>0.6) than other regions, which agree mostly with reference data. The globe and Mediterranean Forests, Woodlands & Scrub (MFWS) also have low RMSD (<1.2) and high PCC (>0.9). Boreal Forests Taiga (BFT), Tundra (TU) and Tropical Subtropical Moist Broadleaf Forests (TSMBF) have very low correlation (PCC <0.1) with reference data. In particular, Boreal Forests Taiga (BFT) has the highest RMSD (1.9) and lowest PCC (<0.1), which has the worst agreement with reference data. Except for Tropical Subtropical Coniferous Forests (TSCF), and Tropical Subtropical Moist Broadleaf Forests (TSMBF), most regions have a higher standard deviation (>1) than the reference data ESA, which shows the model has larger spatial variability than reference data in these regions.

3.4 | Precipitation partitioning for 10 large catchments

In Figure 6, we show the partitioning of the individual components as a percent of the precipitation, compared to the reference data for 10 river basins. The underestimation of runoff and overestimation of transpiration occur in the Amazon, MacKenzie, Murray and Yangtze basins. The overestimation of runoff and the underestimation of transpiration in the Mississippi and Yenisei basins, while the interception is underestimated in almost all basins. Except for the Murray basin, there is an overestimation of evaporation in other basins compared to GLEAM.

4 | DISCUSSION

DGVMs are among our most useful tools to study biosphere–hydrosphere interactions (Gerten et al., 2004; Gerten et al., 2005; Murray et al., 2011; Rost et al., 2008). Simulated features such as vegetation dynamics, root water take-up dynamics, PFT properties and land use management play a critical role in the hydrological cycle by affecting the timing, variability and magnitude of runoff, ET and soil moisture, showing the importance of these tools (Chadburn et al., 2022; Drüke et al., 2019; Horvath et al., 2021). Overall, we find the hydrological performance of the DGVM used in this study, LPJ-GUESS—assessed using a broad range of recently-released independent observation-based datasets of runoff, ET and SSM—to be good in terms of magnitude and variability across a range of temporal and spatial scales. In particular, we find that the model captures: (1) both inter- and intra-annual variations of runoff in most regions, such as the Danube, Murray, Yangtze, Yenisei and Nile basins; (2) global total ET, but with some disagreement in the individual components, especially for evaporation; and (3) surface soil moisture simulated in most regions, though there are discrepancies, particularly in subarctic areas. Furthermore, the significant (p < 0.05) runoff trends seen in the observation data are captured by the model in the Amazon, Ganges and Yenisei basins, and this is in line with the precipitation dataset from CRU-NCEP (Appendix Figure E1).
Compared to Gerten et al. (2004), this version of LPJ-GUESS has added many new modules directly and indirectly related to hydrological processes. For instance, the current model with cropland use improves the underestimation of the runoff at some crop regions such as Western Europe (Appendix Figures K1 and K2), although the irrigation is not fully represented now. The LPJ-GUESS version with soil water freezing has higher runoff in the Yenisei basin, which is closer to the observed values (Appendix Figure K3). The regional and temporal model-data disparities, across multiple components of the terrestrial water cycle, help to identify the processes that are not adequately represented in the model and provide valuable insights for its improvement.

4.1 Model insights in semi-arid and arid regions

LPJ-GUESS appears to overestimate runoff in semi-arid and arid regions for northern Africa and eastern South America as shown in Figure 1c, although the model’s monthly runoff reflects the seasonal pattern of precipitation in the Nile basins (Figure 2). The annual mean precipitation data (CRU-NCEP) used to force LPJ-GUESS seems to be higher in some regions such as northwestern Africa and eastern South America, compared to the GSWP3 dataset used to generate GRUN runoff (Appendix Figure D2), which can partly explain the overestimation. However, the relationship between basin-wide precipitation and runoff in these regions is also complex. In LPJ-GUESS, the lack of explicit consideration of surface energy fluxes and surface temperature may cause underestimations in transpiration (Martín Belda et al., 2022). The comparison with the multiple source ensemble Land-FluxEVAL ALL from 1989 to 2005 (Appendix Figure F1a) also shows that LPJ-GUESS underestimates ET in these areas. Also, the downstream runoff of the Nile basin is probably affected by factors such as evaporation from lakes, reservoirs, wetlands, non-perennial ponds, and river channels, as well as flood plain-channel interactions, groundwater seepage, inter-basin transfers, and human water withdrawal.

FIGURE 6 The observed and modelled decomposition into separate components, shown as a percentage of precipitation, for 10 river basins.
(Clark et al., 2015; Fan et al., 2019, Gerten et al., 2004). However, these factors are not yet explicitly considered in LPJ-GUESS. Furthermore, for enhancing the hydrological accuracy of LPJ-GUESS in semi-arid and arid regions, it may be beneficial to refine the model’s parameterization of vegetation responses to water stress and water use efficiency (Peters et al., 2018), in addition to measures such as implementing a PFT-specific supply approach (Zhang et al., 2013) and dynamically representing root structure, distribution, depth and function (Warren et al., 2015).

4.2 Seasonal events in glacial and subarctic regions

The notable runoff underestimations of LPJ-GUESS at very few grid cells at the most southern part of South America (60°S) (Figure 1c,d) compared to the extremely high runoff values in the GRUN dataset are likely due to an absent model process for short-term glacier melt. The glaciers map (Appendix Figure J1) shows that this region of South America is partly covered by the glaciers over the mountains, which may be affected by processes such as seasonal glacier melt events. The melting of glaciers in the tropical Andes, including the Desert and Central Andes, has been found to contribute as much as 15% to the annual runoff and 27% to the runoff during dry seasons (Soruco et al., 2015).

Additionally, LPJ-GUESS exhibits a high-magnitude spring peak and zero runoff tendency in autumn and winter months for subarctic regions like the MacKenzie and Yenisei basins (Figure 2), which may be linked to the model’s weather generation error of average daily air temperature (Gerten et al., 2004) and the simplified representation of snow formation and melt (Choudhury & DiGirolamo, 1998; Wang et al., 2016; Wever et al., 2014). For example, if the model underestimates short-term snowmelt events in winter months, this will result in an overestimation of both the accumulated snowpack (Pongracz et al., 2021) and associated runoff in spring when it melts.

4.3 Topographic heterogeneity effects

In mountainous areas of southeastern South America (Appendix Figure J2), the inadequate representation of precipitation distribution could be major contributors to the model’s overestimation of runoff in this region (Appendix Figure D2). Several potential solutions exist to address this issue, including correcting for errors in precipitation measurements (Zhao et al., 2021), either on a global scale or on a basin level, or adjusting the calculated runoff with machine learning methods (Bozorg-Haddad et al., 2016). Additionally, the model underestimated the runoff in the Amazon basin while it overestimated it in the mountainous areas southeast of the Amazon basin (upper stream of Amazon) (as shown in Appendix Figure D1). One reason is likely to be the topographic effect on the lateral water transfer, which should drive the water from the mountainous areas with high elevation to the lower elevations in the Amazon basin (Nie et al., 2005). However, these effects of topographic heterogeneity and lateral water transfer along hillslopes are not considered in this version of LPJ-GUESS. Topographic effects at the large basin scale can introduce significant biases in hydrological modelling, especially in mountainous or heterogeneous terrains. These biases stem from the variations in elevation, slope, and aspect that affect solar radiation, precipitation distribution, and subsequently, ET and runoff processes (Ruiz-Arias et al., 2011; Yang et al., 2016). For instance, the study by Wong et al. (2018) on Peninsular Malaysia highlights how topographic heterogeneity can impact the accuracy of coarse resolution grid-based runoff simulations. Kummu and Lauri (2014) further underscore the necessity of slope and river length corrections to improve model accuracy in the Mekong basin. To address the issue, incorporating subgrid topographic heterogeneity and river routing between and within grid cells is required for hydrology improvements in the model (Singh et al., 2015; Tang et al., 2015; Wong et al., 2018).

4.4 River routing

LPJ-GUESS tends to underestimate the runoff levels in the Danube basin between May and October and to overestimate them between January and April (Figure 2), as compared to the reference data derived from observed discharge records at the gauge.

This simulated earlier spring runoff peak in the Danube basin is primarily caused by the model’s lack of a river routing scheme. In order to dampen the peak discrepancy, it would be necessary to consider the time it takes for a wave to travel along the length of the Danube, and to consider the spatial differences in input from the runoff from each cell in the catchment. The simple addition of daily runoff from each grid cell within a river basin results in an unrealistic arrival of the total runoff at the river mouth on the same day (Nguyen-Quang et al., 2018; Oki et al., 1999). Other models that do not incorporate a river routing scheme may also exhibit similar problems (Tang et al., 2014; Ye et al., 2013). We expected the discrepancies of routing time however here we discuss about it for making it as a starting point for further model refinements, such as the introduction of lateral water flow. The higher, earlier peak, and the underestimation later in the year of monthly runoff (Figure 2) in the Nile and Ganges basins is also likely in part due to the absence of river routing and the neglect of water retention by wetlands and lakes (Appendix Figure J3), which are not accounted for in this version of LPJ-GUESS. The incorporation of a river routing scheme, the addition of water storage in lakes, reservoirs, and wetlands, and the improvement of snow storage and melt representation in LPJ-GUESS (Pongracz et al., 2021) would facilitate improved, direct comparisons between simulated runoff and observed discharge in similar catchments (Bloch et al., 2010; Nguyen-Quang et al., 2018; Rost et al., 2008). These enhancements are also expected to improve the seasonal runoff modelling of LPJ-GUESS.
4.5 | Land use change and management

LPJ-GUESS overestimates runoff in agricultural regions because the model’s irrigation scheme assumes that enough water is withdrawn from groundwater to avoid water deficit and stress, and does not consider irrigation from river water. As shown in Appendix Figure D1, the area of overestimated runoff in the Mississippi, Ganges, Nile and eastern part of Yangtze basins overlaps with the extent of irrigated croplands (Appendix Figure D4). Uncertainties in human land-use conversions and management, such as crop cultivation, can impact the model’s runoff estimates (Joannon et al., 2006; Sajikumar & Remya, 2015). Therefore, simulated runoff may be underestimated in regions where the model identifies a dominance of woody plant functional types, despite croplands dominating.

The model underestimates the Amazon’s runoff from April to November and in the Amazon and Congo basins for almost the whole year (Figures 1, 2). Combined with the ET compared with LandFluxEVAL all from 1989 to 2005 (Appendix Figure F1a), we find that the model also overestimates ET in these two basins, which partly explains the runoff underestimation. Forestation and land use conversion (shifting agriculture) in the two basins, which are not fully considered in the model configuration used here, may have impacts on ET/runoff partitioning. In comparison to MPL-V2, both LPJ-GUESS and GLEAM exhibit substantially lower evaporation levels. This potential evaporation underestimation may be attributed to LPJ-GUESS’s failure to account for evaporation from the soil on the grid cell fractions covered by vegetation, as well as the lack of consideration of evaporation from water surfaces such as wetlands (which GLEAM also lacks). Explicit representation of evaporation from both perennial and ephemeral water bodies, and from soil in vegetation-covered areas, could potentially lead to further enhancements in the model’s performance (Gerten et al., 2004). Furthermore, these two basins are also regions with intensive deforestation activities, which could also impact the hydrological cycle (Chaves et al., 2008; Sun et al., 2006).

In summary, the disagreement of these modelled hydrological variables with reference datasets can be attributed to two main causes: (1) the lack of representation of certain processes such as river routing, storage and supply of water bodies, and cropland irrigation derived from runoff; and, (2) uncertainties in input and reference datasets, including precipitation, soil data, and land use data.

5 | CONCLUSIONS

In addition to its demonstrated utility for global carbon cycling applications (Friedlingstein et al., 2022; Pugh et al., 2019; Tagesson et al., 2020), LPJ-GUESS can also capture the observed spatiotemporal variability, magnitude and trends of the main hydrological components such as runoff, ET and SSM. By confronting the model with such a range of hydroclimatological datasets, we could also identify clear discrepancies in some areas and in some months and seasons, which we then used to identify missing or poorly represented processes in the model, such as sub-grid scale topographic heterogeneity, water withdrawal for irrigation and a river routing scheme, which sets the agenda for future model improvements. Overall, we contend that only through standard, comprehensive assessments of the ability of DGVMs like LPJ-GUESS to shed light on and replicate observed interdependencies between terrestrial water cycling and global biogeochemistry, carbon cycling and biogeochemistry generally, can we have confidence in their utility for subsequent climate impact and adaptation studies.

ACKNOWLEDGEMENTS

Hao Zhou is financially supported by the China Scholarship Council and a stipend was created to finance the amount not covered by the CSC by the Department of Physical Geography and Ecosystem Science, Lund University. Jing Tang is supported by Villum Young Investigator (VIL5048), Swedish FORMAS (Forskningsråd för hållbar utveckling) mobility Grant (2016-01580) and Danish National Research Foundation within the Center for Volatile Interactions (VOLT, DNRF168). This study is a contribution to the strategic research areas MERGE (www.merge.lu.se) and BECC (www.becc.lu.se).

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT


ORCID

Hao Zhou https://orcid.org/0000-0001-6335-6950
Jing Tang https://orcid.org/0000-0001-7961-8214
Stefan Olin https://orcid.org/0000-0002-8621-3300
Paul A. Miller https://orcid.org/0000-0003-1105-2132

REFERENCES

of the water cycle to climate change. *Annals of the New York Academy of Sciences*, 1472(1), 49–75.


Batjes, N. (2005). ISRIC-WISE global data set of derived soil properties on a 0.5 by 0.5 degree grid (version 3.0). ISRIC-World Soil Information.


