

Reply to Editor

Dear Editor,

Thank you for handling our manuscript titled “Evaluating Dynamic Global Vegetation Models in China: Challenges in capturing trends in Leaf Area and Gross Primary Productivity” to *Earth System Dynamics*. We are grateful for the constructive feedback received from the reviewers, facilitated by your editorial guidance. Their comments and your editorial input have been instrumental in strengthening the manuscript. We have carefully considered all points raised during the review. We believe the revisions we have made comprehensively address the concerns highlighted by the reviewers and the editorial team. We hope the revised manuscript now meets the high standards for publication in *Earth System Dynamics*.

Below, we provide detailed point-by-point responses to all reviewer and editorial comments. Thank you once again for the time and expertise in reviewing our manuscript.

Sincerely,

Anzhou Zhao

Editor revisions:

Overview

Thank you for your detailed replies to the Reviewers’ comments and the edits that you have implemented in the manuscript. Before your submission may be accepted for publication in *Earth System Dynamics*, I would encourage you to consider the minor points below.

General comment (1)

In one of their comments, Reviewer #1 suggests to: “explain the broader impacts of the study and how this is beneficial to future regional-scale studies using DGVMs or regional studies in general.” In response to this, you added the following text: “This work aims to identify priority pathways for DGVM structural improvements, advance mechanistic understanding of China’s unique carbon cycle dynamics. This study aims to identify key pathways for improving DGVM structure and parametrization, enhance mechanistic understanding of China’s unique carbon cycle dynamics, and provide insights into quantifying regional carbon sinks, thus supporting climate-related policy development and guiding future regional-scale ecosystem modelling studies.”

I would suggest avoiding the repetition “This work” – “This study”. Perhaps for the second reference to your work you could use “Specifically, our analysis aims to” or something similar. You may also consider highlighting the applicability of your results to DGVM studies beyond China, as ESD has a focus on studies whose relevance extends beyond a specific region.

Reply:

We appreciate the editor's constructive suggestions. We have made the following improvements to address the comments.

Changes in the paper:

Line 110-114

“This work aims to identify priority pathways for DGVM structural improvements and enhance understanding of carbon cycle applicable to China and other regions sharing similar ecological characteristics. Specifically, our analysis seeks to identify key pathways for improving DGVM structure and parametrization, advance mechanistic understanding of China's unique carbon cycle dynamics, and provide insights into quantifying regional carbon sinks, thus supporting climate-related policy development and guiding future regional-scale ecosystem modelling studies.”

General comment (2)

In their General Comment #1, the second Reviewer asks: "DGVMs simulate dynamically the terrestrial biogeochemical cycles, accounting for carbon pools and fluxes. However, these pools need to be spun up. For this reason, spin-up simulations are run. Are the considered DGVMs applying the same spin-up protocol? Do they show relevant differences at the beginning of the simulations?" In your reply, you provide a clear answer to the first question of the Reviewer, but never address the second question, namely whether the DGVMs show relevant differences at the beginning of the simulations.

Reply:

We thank the Reviewer and Editor for this insightful question. All DGVMs considered in our study followed a harmonized spin-up protocol, as defined by the TRENDY simulation experiment design. Specifically, each model was spun up using repeated cycles of early 20th-century climate (e.g., 1901–1920 CRU-NCEP data), with fixed atmospheric CO₂ concentrations (e.g., 276.59 ppm), and under potential natural vegetation without land-use changes. The spin-up duration was sufficiently long (typically 500–1500 years) to allow major ecosystem carbon and water pools to reach a near-equilibrium state.

Despite this standardized protocol, relevant differences do exist at the beginning of the transient simulations. These discrepancies arise due to inherent differences in model structure, such as the treatment of soil and litter carbon pools, vegetation dynamics and turnover rates, phenology schemes, and PFT parameterizations and spin-up convergence criteria. As a result, the initial values of variables such as LAI and GPP can differ among DGVMs, even when they have followed the same spin-up forcing and duration. These differences can propagate into early transient simulations and may influence model–data comparison in the initial decades.

Changes in the paper:

Lines 144-147

“Within the TRENDY project, all DGVMs employ uniform spin-up durations. Initialization of simulations utilizes the 1700 baseline conditions with model-specific plant functional types (PFTs). Pre-industrial climate forcing data (1901–1920) is recycled through model-dependent cycling or stochastic sampling until a carbon equilibrium state is achieved (Sitch et al., 2024). Despite uniform TRENDY spin-up protocols, inherent structural differences among DGVMs (e.g., carbon cycling, vegetation dynamics) generate divergent initial states (LAI, GPP) that propagate into early transient simulations. Although such discrepancies may impact model-data comparisons in the first few decades, their effects are minimal during our study period (2003–2019).”

General comment (3)

In their General Comment #2, the second Reviewer asks: “Land-use changes in DGVMs are based on reference global products such as LUH2, as stated in the manuscript. Do all DGVMs use the same land-use change forcing dataset? Does the dataset used in the DGVMs present relevant differences compared to observations?” In your reply, you provide a clear answer to the first question of the Reviewer, but never address the second question, namely to what extent the datasets used in the DGVMs adhere to observations.

Reply:

We thank the reviewer for this important question regarding the land-use change (LUC) forcing used in the DGVM simulations. All participating DGVMs in our study adopted the same standardized land-use forcing dataset, LUH2, as required by the TRENDY protocol. While LUH2 is not a direct observational product, it is a harmonized and downscaled dataset that structurally adheres to observations by integrating three empirical sources: HYDE 3.3 historical land reconstructions, FAO wood harvest statistics, and LUH2-GCB transition matrices. Its core design explicitly separates anthropogenic drivers from natural processes, ensuring methodological alignment with observational principles.

Although absolute agreement with observations at all spatial scales is unattainable for any global product, LUH2’s CMIP-optimized architecture maximizes consistency with the constraints required by DGVMs. This design facilitates inter-model comparability and long-term scenario analysis. That said, LUH2 may still differ from high-resolution regional remote sensing or national inventory datasets in aspects such as spatial resolution (~0.25°), treatment of shifting cultivation or pasture expansion, and sub-grid land-cover representation. Such discrepancies are particularly notable in regions with complex or rapidly changing land-use histories.

Changes in the paper

Lines 150-151

“Concurrently, all participating DGVMs adopt LUH2, leveraging its CMIP-optimized architecture to standardize forcing and enhance DGVM compatibility as the unified land-use change dataset.”

Lines 481-486

“In TRENDY project, most of DGVMs use the HYDE and LUH2 databases for LULCC data input variables (Klein Goldewijk et al., 2017; Hurtt et al., 2020; Sitch et al., 2024). However, absolute agreement with all observations remains unattainable for any global product. While these LULCC datasets are common across models, their implementation varies according to the land-use processes and functionalities inherent to individual DGVMs. In this study, we employ Table 1 to describe differences in the performance of individual DGVMs, which were adapted from Sitch et al. (2024).”

Line 493

Table 1 DGVMs and their main processes relevant to LULCC emissions

Model	Fire	Wood Harvest	Shifting Subgrid transitions	Cultivation/ Crop harvest
CABLE	N	Y	Y	Y
CLASSIC	Y	N	N	Y

CLM5.0	Y	Y	Y	Y
DLEM	N	Y	N	Y
IBIS	Y	Y	N	Y
ISAM	N	Y	N	Y
ISBA	Y	Y	Y	Y
JULES	Y	N	N	Y
LPJ-GUESS	Y	Y	Y	Y
LPX	Y	Y	N	Y
OCN	N	Y	N	Y
ORCHIDEEv3	N	Y	N	Y
SDGVM	Y	N	N	Y
VISIT	Y	Y	Y	Y

Specific comment

You have implemented all of the technical corrections suggested by the Reviewers, but there are still a number of typographical errors and formulations that should be corrected. I list some examples below, but this is by no means a comprehensive list. I would recommend a thorough proof-reading of the text.

Technical corrections(1):

l. 100 “site-scale(Han et al” missing space

Changes in the paper:

Line 110

“Current evaluations of DGVM applicability in China have predominantly relied on site-scale (Han et al., 2025; Zhu and Zeng, 2024)”

Technical corrections(2):

Fig. 1 caption: “error line” --> “error bar”

Changes in the paper:

Fig.1 and the Supplementary figure captions

“The error bar for MME is the standard deviation of the 14 simulated trends.”

Technical corrections(3):

Supplementary figure captions: “The dot indicated the significant trend” --> “Dots indicate significant trends”.

Changes in the paper:

Supplementary figure captions

“Dots indicate significant trends ($p < 0.05$)”

Technical corrections(4):

Fig. 9 caption: “The values over the red line indicated” --> “The values over the red line indicate”

Figs 1, 4, 6, 9, 10, 11, and the Supplementary figure captions

Revise tense truthfulness in sentences.

Reply:

We sincerely appreciate the meticulous review of our manuscript. All typographical errors and suboptimal formulations identified in the editorial comments have been systematically corrected in the revised manuscript. Additionally, we have conducted a full linguistic audit of the text, with all technical corrections implemented and tracked via highlighted text.