

Reviewer 1

The manuscript Peano et al. compared results from seven land surface models (LSMs) to remote sensing estimated leaf area index (LAI) at global scale regarding the seasonal timing of LAI trough (lowest value) and peak (highest value) and amplitude (max. – min.). The results indicate that all studied LSMs tended to show delayed timing of LAI trough and peak, in particular for the Northern Hemisphere while the modeled amplitude is smaller than satellite estimates. While overall the work is pretty straightforward, I do have a few major comments.

We thank the reviewer for reading the manuscript and providing her/his comments.

1. Model description of phenology scheme (L85-125) and related discussions do not provide enough details for readers who are not familiar with these models or with the different phenology schemes or both to understand what might be contributing to the current variations in model simulated LAI and its seasonal dynamics. While most models tend to show delayed timing of LAI trough and peak, there's one model, LPG-GUESS did a better job capturing this (L170, L180). Why so? Similarly, when model differences, model-data discrepancy across different biomes/regions all present, it would be interesting to discuss the potential causes why we see all these differences. While the discussion tried to talk about the different sources of variability (L280) and potential observation biases (L295), it's quite thin and not well organized to address the observed spatiotemporal data-model discrepancy. I think expanding discussion and organize it in a similar way as how the results are organized, and maybe also add a little more detail on different phenology schemes in each LSM can be very helpful. It would also be nice to see some suggestions for future model development or priority areas for data collections for better model parameterization or better observations for model benchmarking.

We thank the reviewer for this comment. The companion paper (Peano et al., 2021) provides in Table 1 a description of the main differences in the phenology scheme used by the seven LSMs. The mentioned table is split into two sections: the first provides the main vegetation features of the LSMs, and the latter focuses on describing the drivers of the start and end of the growing season. We add the following table in this manuscript which is similar to Table 1 in Peano et al (2021) but includes references on the computation of LAI, which we hope addresses the comment

“2.2 Land Surface Model

The seven LSM utilised in the CRESCENDO project are evaluated in this study. A summary of their main features is provided below and listed in Table 1. Further details on the LSMs' phenology schemes are provided in Peano et al. (2021).

LSM	Original Resolution	PFT	CFT	Phenology schemes	LAI driver	LAI Reference
CLM4.5	1.25° x 0.9375°	15	1 C3	Evergreen; Seasonal Deciduous; Stress Deciduous	Leaf Carbon Specific Leaf Area	Thornton and Zimmermann (2007)
CLM5.0	0.5° x 0.5°	15	2 C3	Evergreen; Seasonal Deciduous; Stress Deciduous	Leaf Carbon Specific Leaf Area	Thornton and Zimmermann (2007)
JULES-ES	1.875° x 1.25°	13	1 C3, 1 C4	Deciduous Trees	Balanced LAI Temperature	Cox (2001)
JSBACH	1.9° x 1.9°	12	1 C3, 1 C4	Evergreen; summergreen; raingreen; grasses; tropical crops; extratropical crops	Maximum LAI Temperature Soil Moisture Net Primary Productivity	Dalmonech et al. (2015) Böttcher et al. (2016)
LPJ-GUESS	0.5° x 0.5°	25	3 C3, 2 C4	Evergreen; Seasonal Deciduous; Stress Deciduous	Specific Leaf Area Leaf Biomass	Reich et al. (1992) Smith et al. (2014)
ORCHIDEE	0.5° x 0.5°	15	1 C3, 1 C4	Deciduous; dry and semi-arid; grasses and crops	Temperature	Polcher (1994) Krinner et al. (2005)
ISBA-CTRIP	1° x 1°	16	1 C3, 1 C4	Leaf biomass	Leaf Biomass Specific Leaf Area Leaf Nitrogen	Delire et al. (2020) Gibelin et al. (2006)

Table 1: Grid spatial resolution used for each land surface model (LSM) and references for their principal features about Phenology and Leaf Area Index (LAI) computations. PFT stands for plant functional type, and CFT stands for crop functional type.”

Following the reviewer's suggestion, we increase the discussion on the differences between LSMs. We expand the discussion in the results section as follow:

“3.1.2 Modelled estimates

[...]

Based on these results, the high vegetation heterogeneity represented by LPJ-GUESS (due to both the number of PFTs and CFTs, and original resolution, Table 1) provides a better agreement with CGLS compared to the other LSMs, especially in trough timings, where also ISBA-CTRIP (the second LSM in number of PFTs, Table 1) shows high agreement with CGLS. On the

contrary, the heterogeneity of phenology schemes may improve the ability to capture the correct timings, as done by JSBACH in peak timings (Figure 2 and Table 1). “

“3.2.2 Modelled estimates

[...]

Among the LSMs, ISBA-CTRIP exhibits the best agreement and the lowest error compared to CGLS in seasonal LAI amplitude. This behaviour mainly derives from the ability of ISBA-CTRIP to capture the minimum LAI values (Figure S3h), combined with low biases in maximum LAI (Figure S4h). This ability could originate from a better inclusion of the nitrogen cycle within the LAI computation (Table 1).”

We expand also the discussion section to contain suggestions for future development

“4.2 Comparison with onset and offset evaluation in Peano et al. (2021)

[...]

In general, LSMs should improve their ability to capture vegetation heterogeneity in both plant traits, such as a higher number of PFTs, and phenology features, for example, by increasing the variety of phenology schemes to enhance their ability to represent vegetation responses to various stresses.”

Finally the discussion section on land surface model has been extended as follow:

“4.3 Sources of variability between land surface models

The LSMs involved in this study and its companion (Peano et al., 2021) use state-of-the-art boundary conditions (i.e. atmospheric and land use forcings). The LUH2 land-use evolution (Hurtt et al., 2020) employed in the CMIP6 effort is used in this study. It is noteworthy that each LSM implements the same land-use boundary conditions differently due to dissimilarities in the original resolution, number of PFTs, and land-use scheme implemented in each model (Table 1). Nonetheless, the implementation of a common land-use dataset allows all LSM to reproduce the same spatial land coverage evolution, leaving only differences in plant growth and seasonality among them. Besides, the same set of atmospheric variables has been used to force all the LSMs following the TRENDY protocol (Sitch et al., 2015; Zhao et al., 2016). However, the atmospheric conditions (e.g. temperature and water availability) strongly influence the vegetation growth, and a different source of atmospheric forcing may partially affect the biases obtained in this study. For example, the ISBA-CTRIP model forced by WFDEI forcing (Weedon et al., 2014) exhibits smaller biases compared to the present results (Dewaele et al., 2017).

Despite using the same boundary conditions (atmospheric and land use forcings), the variability between the LSM may derive from various sources, such as differences in vegetation parameterization, crop and plant functional type population, soil characterization, and initial spatial resolution, as already noted in Peano et al. (2021). In particular, the results of this study highlight the relevance for LSMs to capture the high vegetation heterogeneity on both PFTs and CFTs populations, that is the case for LPJ-GUESS, and phenology schemes, as for JSBACH .

Moreover, vegetation parameterizations used in the LSM are based on data from localised areas, typically located in the Northern Hemisphere (e.g. Thornton and Zimmermann, 2007), which may lead to a possible misrepresentation of South Hemisphere features. Consequently, the parameters used in LSMs need to be calibrated against more recent and widespread observations. The results presented in this study highlight a prevalent delay in plant active season compared to observations, despite each LSM implementing different parametrization and processes, emphasising the need for further investigation of the representation of the processes involved in the start of the growing season within the LSMs.

In addition, the case of the Community Land Model provides an example of the impact of model structure versus parameterization on simulated LAI. In the present study, CLM is evaluated in two versions, namely CLM4.5 and CLM5.0. The latter contains various changes in the representation of soil, plant hydrology and carbon and nitrogen cycles (i.e. model structure Lawrence et al., 2019), influencing the simulated vegetation quantity (see differences in seasonal amplitude in Figure 6). On the other hand, CLM5.0 applies limited modifications to the phenology parameterization, resulting in minimal differences in the biases of growing season timings compared to its earlier version (Figures 2b,c and 3b,c), as also shown in Li et al. (2022b). The comparison between CLM versions, then, stresses the separate influence of model structure and phenology parameterization on the simulated LAI features. Moreover, both CLM4.5 and CLM5.0 implement the same LUH2-derived PFT distribution, avoiding the influence of mismatched vegetation type in this version comparison. In general, it highlights the complexity of modelling land surface and vegetation processes and the need for further model development and evaluation.”

2. How much of the data-model discrepancy in LAI seasonality can be contributed to model structure and parameterization respectively? I know this is not the main focus of this study, but I think it's very important to know this before we conclude that we need better phenology models (specifically refer to model structure improvement). Related, it is not clear to me how simulated distribution of the different PFTs (evergreen vs deciduous: they vary in phenology in addition to many other features that can influence plant growth and mortality thus phenology) can influence the modeled LAI seasonality both spatially and temporally. Is it possible that the delayed timing of LAI trough and LAI peak is a mismatch between the observed and the simulated vegetation type that dominate a particular grid cell?

The comparison between the two versions of the Community Land Model, namely CLM4.5 and CLM5.0, supplies a “quality assessment” of the contribution of model structure versus model parameterization. The two LSMs apply the same phenology scheme (= model parameterization) but differ in representing other soil and vegetation processes (= model structure). In this case, it is possible to see differences in the simulated magnitude of the Leaf Area Index but limited differences in timings.

The difference in PFT distribution between models and observations is another potential source of discrepancies. This factor has a limited impact on the intra-models comparison since all the LSMs derive their PFT distribution from the same original dataset (i.e. Land Use Harmonization version 2, LUH2). However, differences in the amount of resolved PFTs can impact the final

result. Similarly, LUH2 should account for observed PFT distribution, mitigating this source of discrepancy.

We added this in the main text in section 4.3 as reported in the reply to point 1.

Specific comments

L20: add relevant citations.

We will add the reference to Running and Coughlan (1988), Bonan et al. (2012), Murray-Tortarolo et al. (2013), Fang et al. (2019).

L40: maybe can expand to add a little more detail here.

We thank the reviewer for the suggestion. We add a description of satellite datasets applications:

“[...] Therefore, the satellite-derived LAI data have been used in assessing model biases (e.g. Murray-Tortarolo et al., 2013; Peano et al., 2019, 2021), performing data assimilation (e.g. Ling et al., 2019), or evaluating the vegetation response to / influence on the ongoing climate change (e.g. Forzieri et al., 2017; Li et al., 2022). [...]”

L140: how the different domain resolution might influence model results? Also how coarse is the model resolution?

The original resolution used by each LSM will be added to the novel table 1 (see reply to point 1 above). The seven LSMs have an original resolution spanning from half-degree to almost two degrees in horizontal resolution. This discrepancy in the original resolution may influence the representation of vegetation heterogeneity and is part of the results' discussion.

L168: multi-model ensemble mean

We thank the reviewer for pointing at this typo.

References

- Bonan, G. B., Oleson, K. W., Fisher, R. A., Lasslop, G., and Reichstein, M.: Reconciling leaf physiological traits and canopy flux data: Use of the TRY and FLUXNET databases in the Community Land Model version 4, *J. Geophys. Res.*, 117, G02026, doi:10.1029/2011JG001913, 2012.
- Böttcher, K., Markkanen, T., Thum, T., Aalto, T., Aurela, M., Reick, C., Kolari, P., Arslan, A., and Pulliainen, J.: Evaluating biosphere model estimates of the start of the vegetation active season in boreal forests by satellite observations, *Remote Sensing*, 8, 580, <https://doi.org/10.3390/rs8070580>, 2016.
- Cox, P. M.: Description of the TRIFFID Dynamic Global Vegetation Model, Hadley Centre Technical Note 24, Hadley Centre, Met Office, Bracknell, UK, 2001.
- Dalmonech, D., Zaehle, S., Schürmann, G., Brovkin, V., Reick, C. H., and Schnur, R.: Separation of the effects of land and climate model errors on simulated contemporary land carbon cycle trends in the MPI Earth System Model version 1, *J. Climate*, 28, 272–291, <https://doi.org/10.1175/JCLI-D-13-00593.1>, 2015.
- Delire C., Séférian, R., Decharme, B., Alkama, R., Calvet, J.-C., Carrer, D., Gibelin, A.-L., Joetzjer, E., Morel, X., Rocher, M., and Tzanos, D.: The global land carbon cycle simulated with ISBA: improvements over the last decade, *J. Adv. Model. Earth Sy.*, 12, e2019MS001886, <https://doi.org/10.1029/2019MS001886>, 2020.
- Dewaele, H., Munier, S., Albergel, C., Planque, C., Laanaia, N., Carrer, D., and Calvet, J.-C.: Parameter optimisation for a better representation of drought by LSMs: inverse modelling vs. sequential data assimilation, *Hydrol. Earth Syst. Sci.*, 21, 4861–4878, <https://doi.org/10.5194/hess-21-4861-2017>, 2017.
- Fang, H., Baret, F., Plummer, S., and Schaepman-Strub, G.: An overview of global leaf area index (LAI): methods, products, validation, and applications, *Reviews of Geophysics*, 57, 739–799, <https://doi.org/10.1029/2018RG000608>, 2019.
- Forzieri, G., Alkama, R., Miralles, D. G. & Cescatti, A. Satellites reveal contrasting responses of regional climate to the widespread greening of Earth. *Science* 356, 1180–1184 (2017).
- Gibelin, A.-L., Calvet, J.-C., Roujean, J.-L., Jarlan, L., & Los, S. (2006). Ability of the land surface model ISBA-A-gs to simulate leaf area index at the global scale: Comparison with satellites products. *Journal of Geophysical Research*, 111, D18102. <https://doi.org/10.1029/2005JD006691>
- Hurt, G. C., Chini, L., Sahajpal, R., Froking, S., Bodirsky, B. L., Calvin, K., Doelman, J. C., Fisk, J., Fujimori, S., Klein Goldewijk, K., Hasegawa, T., Havlik, P., Heinemann, A., Humenöder, F., Jungclaus, J., Kaplan, J. O., Kennedy, J., Krisztin, T., Lawrence, D., Lawrence, P., Ma, L., Mertz, O., Pongratz, J., Popp, A., Poulter, B., Riahi, K., Shevliakova, E., Stehfest, E., Thornton, P., Tubiello, F. N., van Vuuren, D. P., and Zhang, X.: Harmonization of global land use change and management for the period 850–2100 (LUH2) for CMIP6, *Geoscientific Model Development*, 13, 5425–5464, <https://doi.org/10.5194/gmd-13-5425-2020>, 2020.
- Krinner, G. and Viovy, N., de Noblet-Ducoudré, N., Ogée, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S., and Prentice, I. C.: A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system, *Global Biogeochem. Cy.*, 19, GB1015, <https://doi.org/10.1029/2003GB002199>, 2005.

Lawrence, D. M., Fisher, R. A., Koven, C. D., Oleson, K. W., Swenson, S. C., Bonan, G., Collier, N., Ghimire, B., van Kampenhout, L., Kennedy, D., Kluzek, E., Lawrence, P. J., Li, F., Li, H., Lombardozzi, D., Riley, W. J., Sacks, W. J., Shi, M., Vertenstein, M., Wieder, W. R., Xu, C., Ali, A. A., Badger, A. M., Bisht, G., van den Broeke, M., Brunke, M. A., Burns, S. P., Buzan, J., Clark, M., Craig, A., Dahlin, K., Drewniak, B., Fisher, J. B., Flanner, M., Fox, A. M., Gentine, P., Hoffman, F., Keppel-Aleks, G., Knox, R., Kumar, S., Lenaerts, J., Leung, L. R., Lipscomb, W. H., Lu, Y., Pandey, A., Pelletier, J. D., Perket, J., Randerson, J. T., Ricciuto, D. M., Sanderson, B. M., Slater, A., Subin, Z. M., Tang, J., Thomas, R. Q., Val Martin, M., and Zeng, X.: The Community Land Model version 5: Description of new features, benchmarking, and impact of forcing uncertainty, *Journal of Advances in Modeling Earth Systems*, 11, 4245–4287, <https://doi.org/10.1029/2018MS001583>, 2019.

Li, W., Migliavacca, M., Forkel, M., Denissen, J. M. C., Reichstein, M., Yang, H., Duveiller, G., Weber, U., & Orth, R. (2022). Widespread increasing vegetation sensitivity to soil moisture. *Nature Communications*, 13(1), 3959. <https://doi.org/10.1038/s41467-022-31667-9>

Li, X., Melaas, E., Carrillo, C. M., Ault, T., Richardson, A. D., Lawrence, P., Friedl, M. A., Seyednasrollah, B., Lawrence, D. M., and Young, A. M.: A comparison of land surface phenology in the Northern Hemisphere derived from satellite remote sensing and the Community Land Model, *Journal of Hydrometeorology*, 23, 859–873, <https://doi.org/10.1175/JHM-D-21-0169.1>, 2022b.

Ling, X. L., C. B. Fu, W. D. Guo, and Z.-L. Yang (2019), Assimilation of Remotely Sensed LAI Into CLM4CN Using DART, *J. Adv. Model. Earth Syst.*, 11, 2768–2786.
[doi:https://doi.org/10.1029/2019MS001634](https://doi.org/10.1029/2019MS001634).

Murray-Tortarolo, G., Anav, A., Friedlingstein, P., Sitch, S., Piao, S., Zhu, Z., et al. (2013). Evaluation of land surface models in reproducing satellite-derived LAI, over the high-latitude Northern Hemisphere. part I: Uncoupled DGVMs. *Remote Sensing*, 5, 4819–4838. <https://doi.org/10.3390/rs5104819>

Peano, D., Materia, S., Collalti, A., Alessandri, A., Anav, A., Bombelli, A., & Gualdi, S. (2019). Global variability of simulated and observed vegetation growing season. *Journal of Geophysical Research: Biogeosciences*, 124, 3569–3587. <https://doi.org/10.1029/2018JG004881>

Peano, D., Hemming, D., Materia, S., Delire, C., Fan, Y., Joetzjer, E., Lee, H., Nabel, J. E. M. S., Park, T., Peylin, P., Wärlind, D., Wiltshire, A., and Zaehle, S.: Plant phenology evaluation of CRESCENDO land surface models – Part 1: Start and end of the growing season, *Biogeosciences*, 18, 2405–2428, <https://doi.org/10.5194/bg-18-2405-2021>, 2021.

Polcher, J. (1994), *Etude de la sensibilité du climat tropical à la déforestation*, Thèse de doctorat, Univ. Pierre et Marie Curie, Paris.

Reich, P. B., Walters, M. B., and Ellsworth, D. S.: Leaf life-span in relation to leaf, plant and stand characteristics among diverse ecosystems, *Ecol. Monogr.*, 62, 365–392, 1992.

Sitch, S., Friedlingstein, P., Gruber, N., Jones, S. D., Murray-Tortarolo, G., Ahlström, A., Doney, S. C., Graven, H., Heinze, C., Huntingford, C., Levis, S., Levy, P. E., Lomas, M., Poulter, B., Viovy, N., Zaehle, S., Zeng, N., Arneth, A., Bonan, G., Bopp, L., Canadell, J. G., Chevallier, F., Ciais, P., Ellis, R., Gloor, M., Peylin, P., Piao, S. L., Le Quéré, C., Smith, B., Zhu, Z., and Myneni, R.: Recent trends and drivers of regional sources and sinks of carbon dioxide, *Biogeosciences*, 12, 653–679, <https://doi.org/10.5194/bg-12-653-2015>, 2015.

Running, S.W.; Coughlan, J.C. A general model of forest ecosystem processes for regional applications I. Hydrologic balance, canopy gas exchange and primary production processes. *Ecol. Model.* 1988, 42, 125–154.

Smith, B., Wårlind, D., Arneth, A., Hickler, T., Leadley, P., Siltberg, J., and Zaehle, S.: Implications of incorporating N cycling and N limitations on primary production in an individual-based dynamic vegetation model, *Biogeosciences*, 11, 2027–2054, <https://doi.org/10.5194/bg-11-2027-2014>, 2014.

Thornton, P. E., and N. E. Zimmermann, 2007: An Improved Canopy Integration Scheme for a Land Surface Model with Prognostic Canopy Structure. *J. Climate*, 20, 3902–3923, <https://doi.org/10.1175/JCLI4222.1>.

Weedon, G. P., Balsamo, G., Bellouin, N., Gomes, S., Best, M. J., and Viterbo, P.: The WFDEI meteorological forcing dataset: WATCH Forcing Data methodology applied to ERA-Interim reanalysis data, *Water Resources Research*, 50, 7505–7514, <https://doi.org/10.1002/2014WR015638>, 2014.

Zhao, F., Zeng, N., Asrar, G., Friedlingstein, P., Ito, A., Jain, A., Kalnay, E., Kato, E., Koven, C. D., Poulter, B., Rafique, R., Sitch, S., Shu, S., Stocker, B., Viovy, N., Wiltshire, A., and Zaehle, S.: Role of CO₂, climate and land use in regulating the seasonal amplitude increase of carbon fluxes in terrestrial ecosystems: a multimodel analysis, *Biogeosciences*, 13, 5121–5137, <https://doi.org/10.5194/bg-13-5121-2016>, 2016.

Reviewer 2

General comments:

This is an interesting paper comparing global leaf area index (LAI) simulations with satellite-derived LAI observations. Seven land surface models are considered along with 3 satellite LAI data sets. All products and observations are first projected onto the same 0.5 degree x 0.5 degree grid and monthly averages are considered. Maps of simulated and observed LAI peak, trough, and annual amplitude are presented. Mean monthly LAI time series show that the models tend to simulate LAI peaks later than the observations. The paper is reasonably well written, but the discussion section could be improved.

We thank the reviewer for her/his evaluation of the manuscript. We will address the discussion section in the revised version of the manuscript.

Recommendation: major revisions.

Particular comment:

- Phenology simulations may be affected by errors in the atmospheric forcing. What is the quality of the CRUNCEP atmospheric forcing database? Has it been evaluated in previous studies?

We agree with the reviewer that the atmospheric dataset used to force the land surface models has a relevant impact on the final simulated phenology. The selection of CRUNCEP derives from the TRENDY protocol used as a reference for the CRESCENDO experiments. More recent efforts apply other atmospheric forcing, such as GSWP3 or CRUJRA.

We better reference to the choice of CRUNCEP in the methodology section as follow:

“2.3 Experimental setup

All LSMs were forced by near-surface atmospheric variables (2m air temperature, precipitation, wind, surface pressure, short- wave radiation, longwave radiation, and air humidity) from the CRUNCEP version 7 reanalysis dataset (Viovy, 2018), following the TRENDY protocol (Sitch et al., 2015; Zhao et al., 2016), and land cover values from the Land Use Harmonization version 2 (LUH2, Hurtt et al., 2020) [...].”

A direct evaluation of atmospheric forcing is out of the scope of the present study. Moreover, all LSMs are forced by the same atmospheric forcing reducing the impact of atmospheric forcing as a source of difference between model. However, the influence of the atmospheric forcing has been added in the discussion section.

- I am concerned with the definition of "LAI trough". Unlike the LAI peak, low LAI values (e.g. in winter or during a drought) can persist for several months. Commonly used phenology indicators, in addition to peak time, are leaf onset and leaf offset. Why not use these more common indicators? Is LAI trough equivalent to leaf offset?

The present study expands the results presented in its companion paper (Peano et al., 2021), in which the leaf onset and offset timings simulated by the same seven LSMs are compared against the same set of satellite observations. The LAI trough evaluated in the present study represents the peak of dormancy. However, as pointed out by the reviewer, the trough may persist for more than one month. For this reason, we assume the LAI trough/dormancy month is the first month reaching the lowest value. We update the description of peak and trough as follow:

“2.4 Growing Season Analysis

[...]

Peaks are identified as the month with the highest LAI value, representing the apex of the vegetation life cycle. On the contrary, the troughs are identified as the months with the lowest LAI value, depicting the plant's dormancy season. For a situation where the same minimum or maximum LAI values are recorded in multiple months per year, the first month in the year with that value is retained as the peak or trough.

[...]"

- The observed model LAI peak lags at mid and high latitudes may indicate a problem in the representation of temperature. Temperature is a key driver of leaf emergence in these regions. The temperature relevant to phenology is likely to be close to the land surface temperature resulting from the energy budget calculations. This is particularly true for the ISBA model, where phenology is driven by photosynthesis and leaf temperature. Is daytime leaf temperature underestimated in this model? In a number of ISBA papers (e.g. <https://doi.org/10.5194/hess-21-4861-2017>) the LAI peak time is generally consistent with observations. What has changed in the ISBA settings? Is the surface temperature calculated in the same way as before?

We agree with the reviewer on the high relevance of temperature as a key driver of leaf emergence, growth and decay. For this reason, as noted by the reviewer, the selected atmospheric forcing should be clearly accounted for as a possible source of discrepancy between modelled and observed timings. This point has been mentioned in the revised discussion section:

“4.3 Sources of variability between land surface models

The LSMs involved in this study and its companion (Peano et al., 2021) use state-of-the-art boundary conditions (i.e. atmospheric and land use forcings). The LUH2 land-use evolution (Hurt et al., 2020) employed in the CMIP6 effort is used in this study. It is noteworthy that each LSM implements the same land-use boundary conditions differently due to dissimilarities in the original resolution, number of PFTs, and land-use scheme implemented in each model (Table 1). Nonetheless, the implementation of a common land-use dataset allows all LSM to reproduce the same spatial land coverage evolution, leaving only differences in plant growth and seasonality among them. Besides, the same set of atmospheric variables has been used to force all the LSMs

following the TRENDY protocol (Sitch et al., 2015; Zhao et al., 2016). However, the atmospheric conditions (e.g. temperature and water availability) strongly influence the vegetation growth, and a different source of atmospheric forcing may partially affect the biases obtained in this study. For example, the ISBA-CTRIP model forced by WFDEI forcing (Weedon et al., 2014) exhibits smaller biases compared to the present results (Dewaele et al., 2017). [...]"

Regarding ISBA results, the version used in the present study is based on the CMIP6 configuration described in Decharme et al. (2019) and Delire et al. (2020). As mentioned before, the difference in atmospheric forcing (i.e. CRUNCEP here and WFDEI in Dewaele et al., 2017) could partially explain the discrepancies between the two studies. This point has been added in the revised version of the manuscript.

- For all models, it should be stated how the temperature used in the phenology model is calculated and how reliable it is.

Following the comments from the first reviewer, we introduce a table in which the main drivers and approaches used in representing LAI and phenology in each LSM will be referenced as done in Table 1 in the companion paper (Peano et al., 2021).

“2.2 Land Surface Model

The seven LSM utilised in the CRESCENDO project are evaluated in this study. A summary of their main features is provided below and listed in Table 1. Further details on the LSMs' phenology schemes are provided in Peano et al. (2021).

LSM	Original Resolution	PFT	CFT	Phenology schemes	LAI driver	LAI Reference
CLM4.5	1.25° x 0.9375°	15	1 C3	Evergreen; Seasonal Deciduous; Stress Deciduous	Leaf Carbon Specific Leaf Area	Thornton and Zimmermann (2007)
CLM5.0	0.5° x 0.5°	15	2 C3	Evergreen; Seasonal Deciduous; Stress Deciduous	Leaf Carbon Specific Leaf Area	Thornton and Zimmermann (2007)
JULES-ES	1.875° x 1.25°	13	1 C3, 1 C4	Deciduous Trees	Balanced LAI Temperature	Cox (2001)
JSBACH	1.9° x 1.9°	12	1 C3, 1 C4	Evergreen; summergreen; raingreen; grasses; tropical crops;	Maximum LAI Temperature Soil Moisture	Dalmonech et al. (2015) Böttcher et al. (2016)

				extratropical crops	Net Primary Productivity	
LPJ-GUESS	0.5° x 0.5°	25	3 C3, 2 C4	Evergreen; Seasonal Deciduous; Stress Deciduous	Specific Leaf Area Leaf Biomass	Reich et al. (1992) Smith et al. (2014)
ORCHIDEE	0.5° x 0.5°	15	1 C3, 1 C4	Deciduous; dry and semi-arid; grasses and crops	Temperature	Polcher (1994) Krinner et al. (2005)
ISBA-CTRIP	1° x 1°	16	1 C3, 1 C4	Leaf biomass	Leaf Biomass Specific Leaf Area Leaf Nitrogen	Delire et al. (2020) Gibelin et al. (2006)

Table 1: Grid spatial resolution used for each land surface model (LSM) and references for their principal features about Phenology and Leaf Area Index (LAI) computations. PFT stands for plant functional type, and CFT stands for crop functional type.”

References

Böttcher, K., Markkanen, T., Thum, T., Aalto, T., Aurela, M., Reick, C., Kolari, P., Arslan, A., and Pulliainen, J.: Evaluating biosphere model estimates of the start of the vegetation active season in boreal forests by satellite observations, *Remote Sensing*, 8, 580, <https://doi.org/10.3390/rs8070580>, 2016.

Cox, P. M.: *Description of the TRIFFID Dynamic Global Vegetation Model*, Hadley Centre Technical Note 24, Hadley Centre, Met Office, Bracknell, UK, 2001.

Dalmonech, D., Zaehle, S., Schürmann, G., Brovkin, V., Reick, C. H., and Schnur, R.: Separation of the effects of land and climate model errors on simulated contemporary land carbon cycle trends in the MPI Earth System Model version 1, *J. Climate*, 28, 272–291, <https://doi.org/10.1175/JCLI-D-13-00593.1>, 2015.

Decharme, B., Delire, C., Minvielle, M., Colin, J., Vergnes, J.-P., Alias, A., Saint-Martin, D., Séférian, R., Sénési, S., and Voldoire, A.: Recent changes in the ISBA-CTRIP land surface system for use in the CNRM-CM6 climate model and in global off-line hydrological applications, *Journal of Advances in Modeling Earth Systems*, 11, 1207–1252, <https://doi.org/10.1029/2018MS001545>, 2019.

Delire, C., Séférian, R., Decharme, B., Alkama, R., Calvet, J.-C., Carrer, D., Gibelin, A.-L., Joetzjer, E., Morel, X., Rocher, M., and Tzanos, D.: The global land carbon cycle simulated with ISBA: improvements over the last decade, *Journal of Advances in Modeling Earth Systems*, 12, e2019MS001886, <https://doi.org/10.1029/2019MS001886>, 2020.

Dewaele, H., Munier, S., Albergel, C., Planque, C., Laanaia, N., Carrer, D., and Calvet, J.-C.: Parameter optimisation for a better representation of drought by LSMs: inverse modelling vs. sequential data assimilation, *Hydrol. Earth Syst. Sci.*, 21, 4861–4878, <https://doi.org/10.5194/hess-21-4861-2017>, 2017.

Gibelin, A.-L., Calvet, J.-C., Roujean, J.-L., Jarlan, L., & Los, S. (2006). Ability of the land surface model ISBA-A-gs to simulate leaf area index at the global scale: Comparison with satellites products. *Journal of Geophysical Research*, 111, D18102. <https://doi.org/10.1029/2005JD006691>

Hurt, G. C., Chini, L., Sahajpal, R., Froking, S., Bodirsky, B. L., Calvin, K., Doelman, J. C., Fisk, J., Fujimori, S., Klein Goldewijk, K., Hasegawa, T., Havlik, P., Heinemann, A., Humenöder, F., Jungclaus, J., Kaplan, J. O., Kennedy, J., Krisztin, T., Lawrence, D., Lawrence, P., Ma, L., Mertz, O., Pongratz, J., Popp, A., Poulter, B., Riahi, K., Shevliakova, E., Stehfest, E., Thornton, P., Tubiello, F. N., van Vuuren, D. P., and Zhang, X.: Harmonization of global land use change and management for the period 850–2100 (LUH2) for CMIP6, *Geoscientific Model Development*, 13, 5425–5464, <https://doi.org/10.5194/gmd-13-5425-2020>, 2020.

Krinner, G. and Viovy, N., de Noblet-Ducoudré, N., Ogée, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S., and Prentice, I. C.: A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system, *Global Biogeochem. Cy.*, 19, GB1015, <https://doi.org/10.1029/2003GB002199>, 2005.

Peano, D., Hemming, D., Matera, S., Delire, C., Fan, Y., Joetzjer, E., Lee, H., Nabel, J. E. M. S., Park, T., Peylin, P., Wårlind, D., Wiltshire, A., and Zaehle, S.: Plant phenology evaluation of CRESCENDO land surface models – Part 1: Start and end of the growing season, *Biogeosciences*, 18, 2405–2428, <https://doi.org/10.5194/bg-18-2405-2021>, 2021.

Polcher, J. (1994), *Etude de la sensibilité du climat tropical à la déforestation*, Thèse de doctorat, Univ. Pierre et Marie Curie, Paris.

Reich, P. B., Walters, M. B., and Ellsworth, D. S.: Leaf life-span in relation to leaf, plant and stand characteristics among diverse ecosystems, *Ecol. Monogr.*, 62, 365–392, 1992.

Sitch, S., Friedlingstein, P., Gruber, N., Jones, S. D., Murray-Tortarolo, G., Ahlström, A., Doney, S. C., Graven, H., Heinze, C., Huntingford, C., Levis, S., Levy, P. E., Lomas, M., Poulter, B., Viovy, N., Zaehle, S., Zeng, N., Arneth, A., Bonan, G., Bopp, L., Canadell, J. G., Chevallier, F., Ciais, P., Ellis, R., Gloor, M., Peylin, P., Piao, S. L., Le Quéré, C., Smith, B., Zhu, Z., and Myneni, R.: Recent trends and drivers of regional sources and sinks of carbon dioxide, *Biogeosciences*, 12, 653–679, <https://doi.org/10.5194/bg-12-653-2015>, 2015.

Smith, B., Wårlind, D., Arneth, A., Hickler, T., Leadley, P., Siltberg, J., and Zaehle, S.: Implications of incorporating N cycling and N limitations on primary production in an individual-based dynamic vegetation model, *Biogeosciences*, 11, 2027–2054, <https://doi.org/10.5194/bg-11-2027-2014>, 2014.

Thornton, P. E., and N. E. Zimmermann, 2007: An Improved Canopy Integration Scheme for a Land Surface Model with Prognostic Canopy Structure. *J. Climate*, 20, 3902–3923, <https://doi.org/10.1175/JCLI4222.1>.

Viovy, N.: CRUNCEP Version 7 – Atmospheric Forcing Data for the Community Land Model, Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory, Boulder, Colo., available at: <http://rda.ucar.edu/datasets/ds314.3/>, last access: 15 July 2018, 2018.

Weedon, G. P., Balsamo, G., Bellouin, N., Gomes, S., Best, M. J., and Viterbo, P.: The WFDEI meteorological forcing dataset: WATCH Forcing Data methodology applied to ERA-Interim reanalysis data, *Water Resources Research*, 50, 7505–7514, <https://doi.org/10.1002/2014WR015638>, 2014.

Zhao, F., Zeng, N., Asrar, G., Friedlingstein, P., Ito, A., Jain, A., Kalnay, E., Kato, E., Koven, C. D., Poulter, B., Rafique, R., Sitch, S., Shu, S., Stocker, B., Viovy, N., Wiltshire, A., and Zaehle, S.: Role of CO₂, climate and land use in regulating the seasonal amplitude increase of carbon fluxes in terrestrial ecosystems: a multimodel analysis, *Biogeosciences*, 13, 5121–5137, <https://doi.org/10.5194/bg-13-5121-2016>, 2016.