Response to Reviewer #2:

[Comment 1] The manuscript focuses on changes in the output of the dynamic global vegetation model (DGVM) LPJ-GUESS due to increased accuracy of the simulated timing of spring and autumn leaf phenology. Spring and autumn leaf phenology mark the start and end of the growing season for deciduous trees and shrubs, thereby affecting biosphere-atmosphere interactions such as the carbon and water cycles simulated by DGVMs. Accurate DGVM simulations under projected future climate are pivotal for adequate climate mitigation policies, which justifies and substantiates the present study. However, I believe (1) that the study and the manuscript need to be completed, (2) that the conclusions and language need to be more precise, (3) that the readability and comprehensibility need to be improved, and (4) that the discussion needs to be deepened.

Response: We thank the reviewer's helpful and constructive comments, and the recognition of the essentiality of our study. We have revised the integrity, readability, and depth of discussion of the manuscript according to the reviewer's comments, as detailed in the follow-up reply and revised manuscript.

[Comment 2] First, the study compares a newly and currently implemented leaf phenology module (hereafter referred to as new and current LPM, respectively). While the new LPM was specifically calibrated the current LPM was not. Thus, before evaluating the effect of the module structure, it must be isolated from the effect of the specific calibration or it will be distorted and probably overestimated. Further, the study stops at the calculation of the difference between LPJ-GUESS simulations based on the current and new LPMs. I feel that these differences should be analyzed further (e.g., by testing the significance of the difference and comparing the differences between regions). The manuscript fails to present all the data used in the study and to present the software used to analyze the data.

Response: We agree with the reviewer that the simulation difference of vegetation phenology between the extended and original LPJ-GUESS could originate from the effect of the module structure and the specific calibration only applied for the extended version. Therefore, in this revision, we further explored the differences between the original phenological module of LPJ-GUESS model before and after parameterization. We used particle swarm optimization to perform parameter calibration for the spring phenological model (Sykes et al., 1996), and because no autumn phenological module in LPJ-GUESS and the autumn phenology, i.e. the dormancy onset date was determined only by global variable APHEN_MAX, which represents leaf longevity=210, we therefore did not for more the model calibration in autumn. We applied two calibration schemes: the first one is based on the original LPJ-GUESS model to determine a common parameter set for all PFTs, and the second one is to determine a unique set of parameters for each PFTs. The results show that the phenology simulation of the original phenological module under the two calibration schemes was inferior to that of the new phenological module based on the co-controls of temperature and photoperiod (Table S3). Therefore, the new analysis confirms our conclusion that coupling the new phenology module improved the LPJ-GUESS model performance in phenology simulation. In the revised manuscript, we added the new analysis as the Table S3, in addition, we avoided using 'significant' and complete the description of the data and software used in this study. Please refer to the revised manuscript and Supplementary information (Line 354-361, 412-413, 428, 543 and Table S3).

Table S3 Model performance of parameterized original phenological module in LPJ-GUESS.

Plant function	Parameters			Calibration			Validation		
type	a	b	k	R^2	NSE	RMSE	R^2	NSE	RMSE
BNS				0.52	0.35	9.07	0.50	0.29	9.64
IBS&TeBS	0.00	1515.08	0.021	0.41	0.36	11.85	0.41	0.33	12.03
Shrub				0.42	0.31	12.61	0.44	0.27	13.06
	0.00	1649.60	0.024	0.50	0.42	8.63	0.50	0.40	8.91
IBS&TeBS	0.00	105.62	0.005	0.56	0.44	11.02	0.54	0.37	11.61
Shrub	0.00	1657.27	0.018	0.40	0.36	12.11	0.41	0.33	12.49

R^2, coefficient of determination, NSE, Nash–Sutcliffe Efficiency, RMSE, Root mean square error. BNS, boreal needle leaved summergreen tree, IBS, Shade-intolerant broadleaved summergreen tree, TeBS, shade-tolerant temperate broadleaved summergreen tree and Shrubs, summergreen shrubs plant function types). As with Sykes et al. (1996), the parameter a is fixed to 0.

Sykes, M. T., Prentice, I. C., and Cramer, W.: A bioclimatic model for the potential distributions of north European tree species under present and future climates, *J. Biogeogr*., 203-233, 1996.

[Comment 3] Second, the conclusion is compromised by the distorted effect of the module structure (i.e., due to the specific calibration likely increasing the accuracy of the new LPM). Certain results are labeled 'significant', but the study does not apply a measure and corresponding level for statistical significance.

Response: We are total agree with the reviewer's comment and are grateful to the reviewer to point this issue out. As mentioned in Response to [Comment 2], we have now separated the impacts into the changes in phenology module and change from the parameter calibration changes. In the revised manuscript, we removed the word 'significant'. Please refer to the revised manuscript Line 412-413, 428 and 543.

[Comment 4] Third, the readability and understandability are affected by imprecise language and long sentences as well as by mistakes in grammar and syntax.

Response: We have improved the language of the manuscript, corrected errors in grammar and syntax, and avoided using long sentences to increase its readability and comprehensibility. Please refer to the revised manuscript.

[Comment 5] Fourth, while some results are already discussed in the Results section, I would like to see a more focused and in-depth Discussion section. Below some suggestions.

Response: Following the reviewer's suggestion, we have updated the discussion section by adding more in-depth discussion. Please refer to the revised manuscript Line 450-534 and our replies below.

Major comments

[Comment 6] 1. Completeness. Two phenology models (Caffarra et al., 2011; Delpierre et al., 2009) are calibrated and constitute the new LPM implemented in LPJ-GUESS. One of these phenology models simulates the start of the growing season (SOS; Caffarra et al., 2011), while the other simulates the end of the growing season (EOS; Delpierre et al., 2009). Simulated SOS and EOS are outputs of LPJ-GUESS, which further include simulated gross primary productivity (GPP), foliar projection cover (FPC), and actual evapotranspiration (AET).

1.1. According to my understanding of the manuscript, SOS and EOS according to LPJ-GUESS are directly taken from the new versus current LPM, and are evaluated against the same data with which the new LPM was calibrated (i.e., the NDVI of the GIMMS data set; L. 113–121). If this is the case, the results regarding SOS and EOS simulated by LPJ-GUESS (L. 302–310) are technically a comparison of the new versus the current LPM (rather than an evaluation of LPJ-GUESS, which should be clearly stated; see below). Moreover, the new LPM was specifically calibrated with the GIMMS data set (L. 251–254), whereas the current LPM was not (i.e., the module parameters were taken from the current LPJ-GUESS code). Because the accuracy of both LPM in simulating SOS and EOS was also assessed with the GIMMS data set (L. 302, not explained in the Data and Methods section), the increased accuracy of the new LPM (L. 302–310) is expected. It cannot be determined, to what degree this increase in accuracy is the result of the specific calibration or the formulation of the new LPM. To really compare the new and current LPM, I suggest to also specifically calibrate the current LPM (i.e., calibrated constants a, b, and k as well as calibrated longevity for the currently implemented models; L. 181–197), using the same calibration sample that has been used to calibrate the new LPM.
Response: We thank the reviewers for this important point. In this revision, we have now added the calibration for the original phenology module in LPJ-GUESS.

For the original spring phenology module in LPJ-GUESS, all tree PFTs share the same parameters, so we firstly parameterized the a, b and k parameters of tree PFTs. Then, we also calibrated unique parameters for each PFT based on PFT distribution and satellite NDVI data. Since the autumn phenology is based on a fixed global variable APHEN_MAX (leaf longevity, which is equal to 210) in the LPJ-GUESS, the improvement of the phenology simulation performance of the autumn phenology model is due to the introduction of a new autumn phenology module. The results of calibrating the original spring phenology module are shown in Table S3 (can be also found below), which shows that the phenology simulation performance of the original phenological model of LPJ-GUESS after parameter calibration is still worse than that of the DORMPHOT model after parameter calibration. Therefore, the new analysis confirms our conclusion that coupling the new phenology module improved the LPJ-GUESS model performance in phenology simulation.

In the revised manuscript, we added the new analysis as the Table S3, in addition, we avoided using 'significant' and complete the description of the data and software used in this study. Please refer to the revised manuscript and Supplementary information (Line 354-361, 412-413, 428, 543 and Table S3).

Table S3 Model performance of parameterized original phenological module in LPJ-GUESS.

Plant function type	Parameters			Calibration			Validation		
	a	b	k	R^2	NSE	RMSE	R^2	NSE	RMSE
BNS				0.52	0.35	9.07	0.50	0.29	9.64
IBS&TeBS	0.00	1515.08	0.021	0.41	0.36	11.85	0.41	0.33	12.03
Shrub				0.42	0.31	12.61	0.44	0.27	13.06
BNS	0.00	1649.60	0.024	0.50	0.42	8.63	0.50	0.40	8.91
IBS&TeBS	0.00	105.62	0.005	0.56	0.44	11.02	0.54	0.37	11.61
Shrub	0.00	1657.27	0.018	0.40	0.36	12.11	0.41	0.33	12.49

R^2, coefficient of determination, NSE, Nash–Sutcliffe Efficiency, RMSE, Root mean square error. BNS, boreal needle leaved summergreen tree, IBS, Shade-intolerant broadleaved summergreen tree, TeBS, shade-tolerant temperate broadleaved summergreen tree and Shrubs, summergreen shrubs plant function types). As with Sykes et al. (1996), the parameter a is fixed to 0.

Sykes, M. T., Prentice, I. C., and Cramer, W.: A bioclimatic model for the potential distributions of north European tree species under present and future climates, *J. Biogeogr.*, 203-233, 1996.

[Comment 7] 1.2. While GPP, FPC and AET simulated by LPJ-GUESS were compared between the LPJ-GUESS running with the new and current LPM, GPP simulations were further compared with the (not introduced; see below) VPM GPP product. These comparisons include the results of the LPJ-GUESS running with the new versus current LPM as well as the difference between these results. Here, I would like to see more, such as (1) a comparison of the special distributions of GPP, FPC, and AET when simulated with LPJ-GUESS running with the new versus current LPM and (2) an evaluation against observational data.

Response: Following the reviewer's suggestions, we compared the spatial distribution pattern of original and extended LPJ-GUESS simulated gross primary productivity (GPP) and actual evapotranspiration (AET) during transition period, Spring (March to May) and Autumn (August to November), with VPM GPP and REA ET data. The results show that LPJ-GUESS can accurately simulate the spatial pattern of GPP and AET during transition period. Please refer to the revised manuscript and supplementary information (Line 394-397, 429, Fig. S6 and S7).

Figure S6 Spatial distributions of Spring (March to May) and Autumn (August to November) gross primary productivity (GPP) of LPJ-GUESS simulation and VPM GPP data. (a-c) Spring GPP with original, extended LPJ-GUESS and VPM data. (d-f) Autumn GPP with original, extended LPJ-GUESS and VPM data.

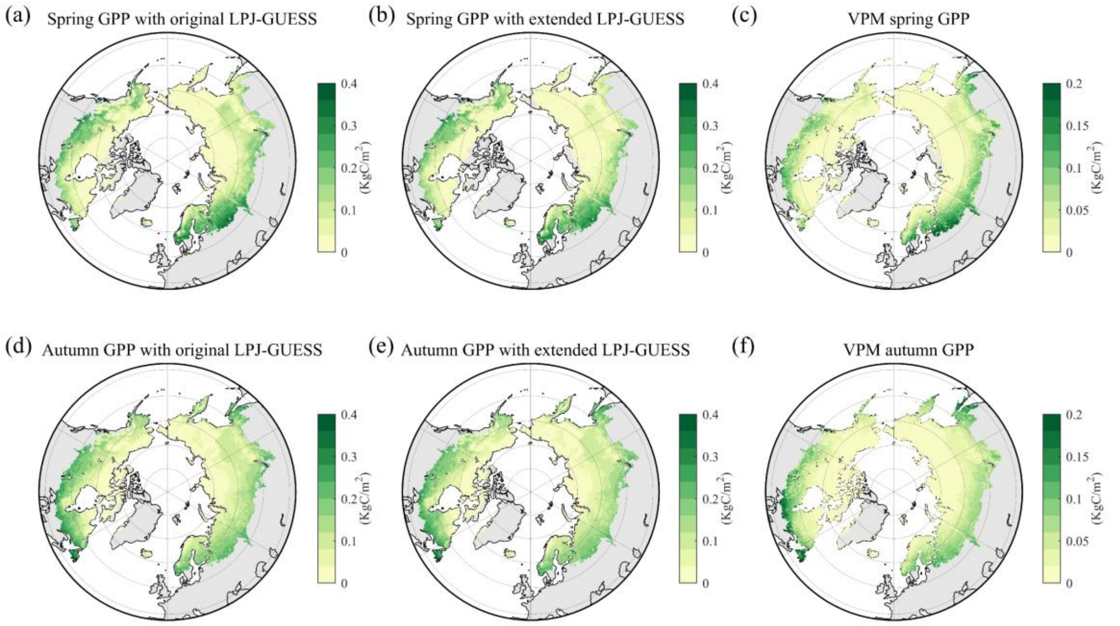
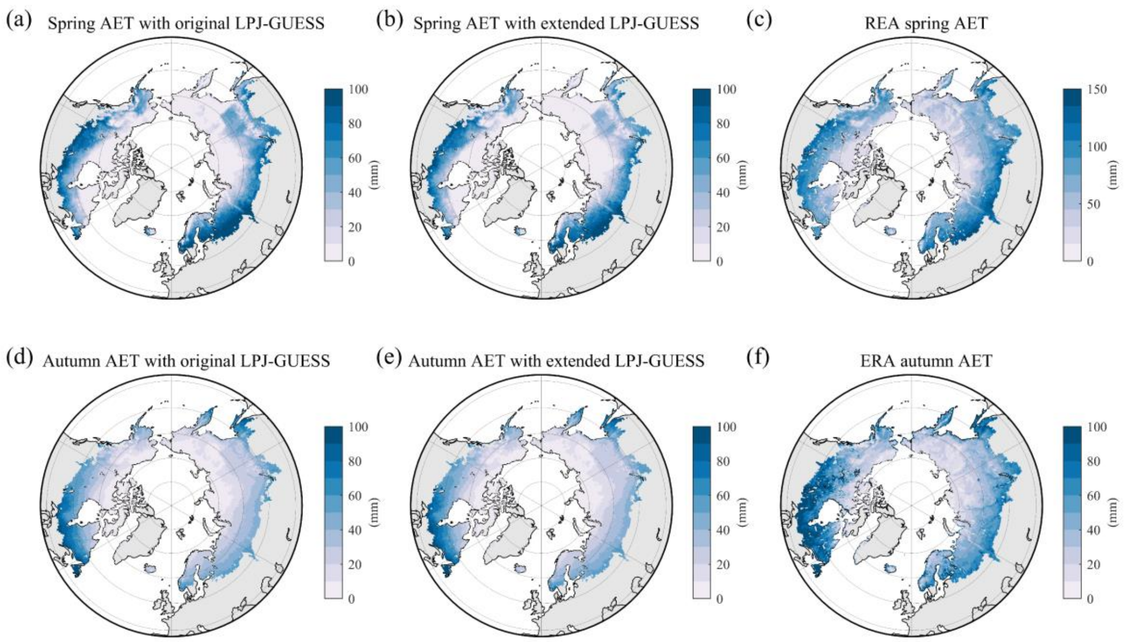


Figure S7 Spatial distributions of Spring (March to May) and Autumn (August to November) actual evapotranspiration (AET) of LPJ-GUESS simulation and REA ET data. (a-c) Spring AET with original, extended LPJ-GUESS and REA ET data. (d-f) Autumn AET with original, extended LPJ-GUESS and REA ET data.



[Comment 8] 1.3. The authors refer to some results as 'significant' (L. 372, 386, and 486) but mention neither a significance level nor a method with which the statistical significance was determined. I am aware that 'significant' has the literal meaning of 'notable' and may be used in that sense. In my opinion however, the term 'significant' usually refers to the value of a significance metric (e.g., the p-value) in scientific studies, why I urge the authors to also use it in this latter sense.

Response: We thank the reviewer for making us notice this, and following the reviewer's suggestion, we avoid the misleading use of 'significant' and remove it from the revised manuscript. Please refer to the revised manuscript Line 412-413, 428 and 543.

[Comment 9] 1.4. Following data was used but not introduced (therefore needing introduction in the Data section): (1) CRU NCEP v7 gridded climate data (L. 271), (2) VPM GPP products (L. 354–355).

Response: Following the reviewer's suggestion, in the revised manuscript, we have added corresponding description of all the data used in this study in the Data section. Please refer to the revised manuscript Line 125-130 and 160-177.

[Comment 10] 1.5. The software used for data preparation, model calibration, data analysis, and result visualization is omitted and should be mentioned at the end of section 2.

Response: Following the reviewer's suggestion, we have added the description software used for data processing and analysis in this study as "*All the data processing and analysis in this study were completed in matlab 2020b (www.mathworks.com)*." at the end of section 2. Please refer to the revised manuscript Line 313-314.

[Comment 11] 2. Precision. 2.1. The study implements the SOS model by Caffarra et al. (2011) in LPJ-GUESS. This model is called DORMPHOT and not DROMPHOT, which must be corrected throughout the manuscript (e.g., L. 87, 207, 213, etc.).

Response: We thank the reviewer's for making us notice this spelling mistake, and we have corrected all the spelling as "DORMPHOT", please refer to the revised manuscript.

[Comment 12] 2.2. I feel that the authors sometimes used 'developed' and 'constructed' where 'implemented', 'adopted', 'extended', 'improved', etc. would be more appropriate. For example, did this study really develop/construct SOS and EOS models (L. 20–21 and 479–480) and LPJ-GUESS (L. 24)? Because all these models were taken from the literature, the EOS and SOS were probably rather 'implemented' and LPJ-GUESS was rather 'improved'.

Response: We have seriously considered reviewer's comment and we totally agree that it is more accurate to use 'implemented' when describing phenological models. According to the reviewer's suggestion, we have revised the expression of the spring and autumn phenological model to 'implemented', and unified the description of LPJ-GUESS to 'extended'. Please refer to the revised manuscript.

[Comment 13] 2.3. Lines 393–395 and 492 mention both 'water stress' and 'legacy effects', which must be defined in the Methods section. Moreover, the statement made in lines L. 393–395 seems not justified by any results.

Response: Following the reviewer's suggestion, we have removed the statement which was not justified by any results. Please refer to the revised manuscript Line 437-439 and 551-552.

[Comment 14] 2.4. I have difficulties with the conclusion that "LPJ-GUESS using the modified phenological module substantially improved […] (the) accuracy of spring and autumn phenology compared to […] (when using) the original phenological module" (L. 483–485). In my opinion, the study rather shows that the timing of SOS and EOS was simulated more accurately by the new versus current LPM implemented in LPJ-GUESS, which may partly be because of the module formulation. However, in contrast to the currently implemented phenology module, the new LPM was specifically calibrated (see above). The study cannot untangle the effect of the specific calibration from the effect of the module formulation. In consequence, the results do not allow to conclude on the isolated strength of either one of these effects. I strongly urge the authors to specifically calibrate both the new and current LPM with an identical calibration sample before comparing their accuracy based on an identical validation sample.
Response: We are grateful to this helpful and constructive comment, and as mentioned above in response to Comments 2 and 6, we aim to compare the differences between the LPJ-GUESS model introduced with the DORMPHOT and DM models based on the cooperative regulation mechanism of temperature and photoperiod and the currently widely used LPJ-GUESS model. Retaining the phenological module setup of the original LPJ-GUESS model can best reflect the impact of introducing the new phenological model. Following the reviewer's suggestion, we have also parameterization the parameter sets for original phenological model of LPJ-GUESS applying two schemes (for tree group or for specific PFT), and the results show that the phenology simulation performance of the original phenological module under the two calibration schemes was inferior to that of the new phenological module based on the cooperative control of temperature and photoperiod (Line 354-361 and Table S3). Overall, the added analysis strength our conclusion, and we thank the reviewer for helping us to mention this and improve the manuscript. In the revised manuscript, we thank the reviewer's help in the acknowledgements.

[Comment 15] 3. Readability and understanding. 3.1. The grammar and syntax needs serous improvement to increase readability and understanding of the manuscript. Examples are: «Vegetation phenological shifts impact [...], and affects» (L. 14), «we developed and coupled the spring and autumn phenology models into [...] LPJ-GUESS»(L 20–21), «These process-based phenology models are driven by temperature and photoperiod, and are parameterized for deciduous trees and shrubs by using remote sensing based phenological observations and reanalysised climate dataset ERA5 land» (L. 21– 24), and «the simulated RMSE for deciduous trees and shrubs» (L. 26–26).
Response: We thank the reviewer for this helpful comment, following the reviewer's suggestion, we have modified the grammar and syntax of the manuscript, please refer to the revised manuscript.

[Comment 16] 3.2. The manuscript contains many long sentences (e.g., L. 61–71, 101–106, 213–216, 271–278, and 443–446), which arguably can only be understood with additional effort. To increase the readability of the manuscript, I suggest to shorten most of the long sentences, for example by splitting the sentences.

Response: Following to the reviewer's suggestion, we changed the excessively long sentences into short sentences. Please refer to the revised manuscript Line 62-72, 103-108, 239-243, 298-305 and 497-503.

[Comment 17] 4. Discussion. 4.1. I do not understand the relevance of section 4.1 (L. 406–436) for this manuscript. Models to simulate SOS and EOS have been calibrated with remote sensed data before (e.g., Keenan & Richardson, 2015; White et al., 1997). Moreover, in my opinion, because the study does not assess the accuracy of vegetation indices derived from remote sensed observations, this procedure does not need to be discussed here.

Response: We have seriously considered reviewer's comment and we agree that models to simulate SOS and EOS have been calibrated with remote sensed data before, and there is still a key issue which toned to be discussed that the information obtained by remote sensing is generally from mixed pixels, and the usual phenological model cannot simulate the changes of components in pixels. However, LPJ-GUESS is a vegetation dynamic process model, and its dynamic vegetation process can just solve this problem, simulate the succession process of ecosystems, and simulate the dynamic changes of components in mixed pixels. The improvement of the simulation accuracy of different vegetation types therefore can provide an opportunity for the precise simulation of mixed pixel phenology. We are sorry for not clearly stating this point, and in the revised manuscript, we have revised the discussion in section 4.1. Please refer to the revised manuscript Line 452-490.

[Comment 18] 4.2. The second paragraph of section 4.1 (L. 420 – 436) does not contain any references to the literature. In addition, I felt that this paragraph is rather an opinion than a discussion of results. Please refer to your results and corresponding literature or consider omitting the paragraph.

Response: Following the reviewer's suggestion, we have added corresponding references and revised the second paragraph of section 4.1. Please refer to the revised manuscript Line 470-490.

[Comment 19] 4.3. The way advancing spring phenology is discussed now, it appears that an advancement always results in an advantage for the concerned species at high elevations (L. 443–446). I doubt that this is true. Many studies have shown that earlier spring phenology also relates with an increased risk in damaged tissue and shoots due to late frost and the weight of late snow fall (e.g., Augspurger, 2009; Bigler & Bugmann, 2018; Drepper et al., 2022). This aspect of advancing spring phenology should be mentioned in the discussion.

Response: We thank the reviewer for this helpful comment, we have added corresponding discussion about the increasing risk of frost with earlier spring onset as

"*In high latitude regions, plants gain a competitive niche through the advancement of spring phenology if there is no damaged tissue and shoots induced by late frost and the weight of late snow fall (Augspurger, 2009; Bigler and Bugmann, 2018; Drepper et al., 2022; Liu et al., 2018).*". Please refer to the revised manuscript Line 502-505.

Augspurger, C. K.: Spring 2007 warmth and frost: phenology, damage and refoliation in a temperate deciduous forest, *Funct. Ecol.*, 23, 1031-1039, 2009.

Bigler, C. and Bugmann, H.: Climate-induced shifts in leaf unfolding and frost risk of European trees and shrubs, *Sci. Rep.*, 8, 9865, 2018.

Drepper, B., Gobin, A., and Van Orshoven, J.: Spatio-temporal assessment of frost risks during the flowering of pear trees in Belgium for 1971–2068, *Agric. For. Meteorol.*, 315, 108822, 2022.

Liu, Q., Piao, S., Janssens, I. A., Fu, Y., Peng, S., Lian, X., Ciais, P., Myneni, R. B., Peñuelas, J., and Wang, T.: Extension of the growing season increases vegetation exposure to frost, *Nature communications*, 9, 426, 2018.

Minor comments

[Comment 20] 1. To my understanding, particle swarm optimization was used to calibrate the newly implemented phenology models DORMPHOT and DM (L. 257–258). Therefore, the yellow box 'Particle swarm optimization' in Figure 2 (L. 209) should probably come after each of the yellow boxes 'DORMPHOT' and 'DM' rather than between the grey boxes for 'SOS'/'EOS' and 'GLC 2000'.

Response: Following the reviewer's suggestion, we modified Figure 2. Please refer to the revised manuscript Fig.2.

[Comment 21] 2. In my opinion, the result regarding the leaf area index (L. 318–322) is unrelated to the results regarding GPP, FPC, and AET, and therefore irrelevant for this study. I suggest omitting it.
Response: We thank the reviewer for this helpful comment, and we agree that the description of simulation of leaf area index is irrelevant with the results section regarding GPP, FPC and AET. In this study, we used leaf area index simulation to dynamically reveal simulated differences in vegetation phenology, so the description of LAI was replaced in section 3.1. Please refer to the revised manuscript Line 350-354.

[Comment 22] 3. While the references for Figures 4 and 5 in the text match the figure captions, the actual figures are mixed up.
Response: We thank the reviewer for making us notice this, we have adjusted Figure 4 and 5 to the correct position. Please refer to the revised manuscript.

[Comment 23] 4. Some results are already being discussed in the Result section (e.g., L. 392 – 395). Please move all discussion the Discussion section.
Response: We thank the reviewer for this helpful comment, we have removed the

relevant sentences. Please refer to the revised manuscript Line 437-439.

Refernces

Augspurger, C. K. (2009). Spring 2007 warmth and frost: Phenology, damage and refoliation in a temperate deciduous forest. *Functional Ecology*, 23(6), 1031–1039.

https://doi.org/10.1111/j.1365-2435.2009.01587.x

Bigler, C., & Bugmann, H. (2018). Climate-induced shifts in leaf unfolding and frost risk of European trees and shrubs. *Sci Rep*, 8(1), 9865. https://doi.org/10.1038/s41598-018-27893-1

Caffarra, A., Donnelly, A., & Chuine, I. (2011). Modelling the timing of Betula pubescens budburst. II. Integrating complex effects of photoperiod into process-based models. *Climate Research*, 46(2), 159–170. https://doi.org/10.3354/cr00983

Delpierre, N., Dufrene, E., Soudani, K., Ulrich, E., Cecchini, S., Boe, J., & Francois, C. (2009). Modelling interannual and spatial variability of leaf senescence for three deciduous tree species in France. *Agricultural and Forest Meteorology*, 149(6–7), 938–948. https://doi.org/10.1016/j.agrformet.2008.11.014

Drepper, B., Gobin, A., & Van Orshoven, J. (2022). Spatio-temporal assessment of frost risks during the flowering of pear trees in Belgium for 1971–2068. *Agricultural and Forest Meteorology*, 315, 108822. https://doi.org/10.1016/j.agrformet.2022.108822

Keenan, T. F., & Richardson, A. D. (2015). The timing of autumn senescence is affected by the timing of spring phenology: Implications for predictive models. *Glob Chang Biol*, 21(7), 2634–2641. https://doi.org/10.1111/gcb.12890

White, M. A., Thornton, P. E., & Running, S. W. (1997). A continental phenology model for monitoring vegetation responses to interannual climatic variability. *Global Biogeochemical Cycles*, 11(2), 217–234. https://doi.org/10.1029/97gb00330