Interactive comment on “Probing the past 30 year phenology trend of US deciduous forests” by X. Yue et al.

Anonymous Referee #2

We thank the referee for the positive, helpful and comprehensive review that has improved the manuscript. Detailed point-by-point responses to the reviewer comments are provided below. The reviewers’ comments are shown in black with our responses are marked as blue. The line numbers below refer to the revised manuscript to be submitted separately.

I have read the discussion paper "Probing the past 30-year phenology trend of U.S. deciduous forests" by Yue et al. with great interest.

In this paper, the authors test a suite of phenology models against site-level data in the US. Models represent spring budburst and autumn dormancy with various degrees of complexity as shown by differences in the required input data and the number of parameters. Models are first calibrated against phenology observations at four deciduous forests. Then, the selected best spring and autumn models are used to produce time series of phenology for the last 30 years over the whole conterminous US. The authors conclude to a temperature-driven advance in the spring budburst in the East and a delay in dormancy in the Northeast and West with large regional variations.

Phenology in ecosystem models is a topical question and this study falls well within the scope of the journal. The paper is well-written and easy to follow but I have several comments about the modelling procedure that needs to be clarified before publication (see below).

My main comment is that most of the spring phenology models have been derived from the same site, and calibrated against only 4 sites. How reliable are these models once applied to represent forests with varying species composition under very different climatic conditions? An assessment of the uncertainty in at least the models used to produce the regional maps is needed to check whether trends are a robust, or a result of over-fitted models.

→ We fully agree that the assessment of modeling uncertainty is important for the understanding of the robustness of our model predictions. In this study, we have quantified impacts of climate variability, model structure, and species aggregation on the predicted phenological trends. We also discussed the uncertainties from other possible sources, such as the incompleteness of observations, unrealistic representation of chilling/photoperiod limits, incompatibility of model parameters at site-level and continental scales, and so on.

We performed two additional sensitivity tests and added a new figure S18 following the reviewer’s comment. “We analyze species-specific temperature sensitivity of tree phenology at Harvard Forest (section 3.4.3). Based on these results, we perform two additional sensitivity tests to evaluate modeling uncertainties from the intraspecific variations. In the first run (simulation 4), phenological parameters are derived based on records of species with the lowest temperature sensitivity for both spring (Sweet Birch, *Betula Lenta*) and autumn (Paper Birch, *Betula Papyrifera*). In the other run (simulation 5), parameters are derived using records of species with the highest temperature sensitivity for spring (Striped Maple, *Acer Pensylvanicum*) and autumn (Black Oak, *Quercus Velutina*). We applied the derived parameters for the whole domain of U.S. by ignoring the realistic fractional coverage of specific species, so as to estimate the maximum uncertainty of prediction due to the intraspecific variations." (Lines 374-384)

"We perform two sensitivity runs to evaluate the modeling uncertainties due to intraspecific variations (Fig. S18). Simulations with either the lowest (simulation 4) or the highest (simulation 5) temperature sensitivity yield very similar phenological trends as that in the control simulation (simulation 2). In the East, simulation 4 predicts a spring advance by 0.33 day yr^{-1} while simulation 5 predicts an advance by 0.35 day yr^{-1}, both of which are close to the 0.34 day yr^{-1} from the control run. In the West and Northeast, both sensitivity runs predict autumn delay by 0.13-0.15 day yr^{-1}, lower than value of 0.14-0.16 day yr^{-1} from control run, suggesting that site-level responses may not be necessarily consistent with responses at the continental scale. Both the similar temperature sensitivity at site level (Fig. 8) and the predicted phenological trends at continental scale (Fig. S18) support the concept of phenological modeling at the forest and PFT level, and corroborates the further investigation of phenology-climate interactions at the continental and global scale." (Lines 656-668)

Specific comments

P6042 L2: Please define "long-term"? Perhaps the term "decadal" would be more suited to describe the station data.

→ Yes. We have revised "long-term" to "decadal" as suggested.

P6043 L18: Would it be more correct to estimate budburst and dormancy based on significant changes in the LAI time series rather than a threshold?

→ It is a good suggestion to estimate phenological dates based on rapid changes in LAI. However, LAI records are not continuous, making it difficult to identify those phase changes. For example, measurements at US-UMB in 1999 started at DOY 121 with LAI of 1.5 m^2 m^{-2}, which could be considered as the background value of LAI. However, in 2007, measurements began on DOY 132 with LAI of 1.93 m^2 m^{-2}. For the year 2007, budburst date could be calculated only by extrapolating LAI to the day with value of 1.5 m^2 m^{-2}.

P6045 L1: See my main comments. Also, why is Jolly et al.'s (2005) phenology model not used for spring?

→ We have responded to your main comments. All spring models used for the inter-comparison considers chilling requirement while Jolly et al. (2005) does not implement such an effect.

P6046 L11: More detailed are needed to describe the calibration method. I am concerned that calibrating "by hand", as it appears, rather than using an automated tool does not allow to find the "true" optimum. Generally, using different model structures introduces more uncertainty than the equifinality of parameters of a single model. However, a few words about how well-defined the parameters are is required here.

→ In the text, we explained that: "For each model in Table 5, we apply the exhaustive enumeration method to evaluate all combinations of the discrete parameters. We select the optimized parameters that jointly predict the lowest RMSE for the long-term budburst dates at the four calibration sites." (Lines 279-282) We have shown the values of optimized parameters in Tables S3 and S4.

P6051 Section 3.1.2: Can you compare these results with satellite based estimates from Buitenwerf et al. (2015)?

→ Section 3.1.2 evaluates spatial distribution of simulated phenology while Buitenwerf et al. (2015) estimates changes in the standard deviation of phenological metrics. The two studies are investigating different aspects of phenology, making it difficult to perform the comparison directly.

P6052 L12: $p < 0.2$ is a very generous threshold

→ We selected $p<0.2$ because time series is relatively short and correlation coefficient is largely affected by interannual variability (as we discussed in section 3.4.1).

Figures Please add coordinates on Figures 3 to 7.

→ Coordinates have been added as suggested.

In Figure 3, why do the coloured surface appear to have different shapes in panels a and c, especially in the Southeast? Is it an artefact due to the colour scale that draws missing areas and areas with lower values in white? More generally, why are the coloured areas in all maps different from the map in Figure 1?

→ Yes, the inconsistency between a and c is due to lower values in white. A full coverage of DBF has been shown in Figure 1. We have clarified in the caption of Figure 3 as follows: "The coverage of colored patches in (a) and (c) differs from that in Figure 1 because values at and beyond the low end of color scales have been shown in white."

Reference

Buitenwerf, R., Rose, L., and Higgins, S. I.: Three decades of multi-dimensional change in global leaf phenology, Nat Clim Change, 5, 364-368, 2015.

Jolly, W. M., Nemani, R., and Running, S. W.: A generalized, bioclimatic index to predict foliar phenology in response to climate, Global Change Biol, 11, 619-632, doi:10.1111/J.1365-2486.2005.00930.X, 2005.