DeepPhenoMem V1.0: Deep learning modelling of canopy greenness dynamics accounting for multi-variate meteorological memory effects on vegetation phenology

Guohua Liu [1, 2], Mirco Migliavacca [3], Christian Reimers [1], Basil Kraft [1], Markus Reichstein [1], Andrew D. Richardson [4], Lisa Wingate [5], Nicolas Delpierre [6], Hui Yang [1], Alexander J. Winkler [1]

[1]Department of Biogeochemical Integration, Max Planck Institute for Biogeochemistry, Jena, 07745, Germany
[2]Jiangsu Key Laboratory of Agricultural Meteorology, Nanjing University of Information Science and Technology, Nanjing, 210044, China
[3]European Commission - Joint Research Centre Via Enrico Fermi, Ispra (VA), 21027, Italy
[4]School of Informatics, Computing, and Cyber Systems, and Center for Ecosystem Science and Society, Northern Arizona University, Flagstaff, AZ 86011, United States
[5]INRAE, Bordeaux Sciences Agro, UMR ISPA, Villenave d'Ornon, 33140, France
[6]Université Paris-Saclay, CNRS, AgroParisTech, Ecologie Systématique et Evolution, 91190, Gif-sur-Yvette, France

Correspondence to: Guohua Liu (gliu@bgc-jena.mpg.de)

Abstract. Vegetation phenology plays a key role in controlling the seasonality of ecosystem processes that modulate carbon, water and energy fluxes between biosphere and atmosphere. Accurate modelling of vegetation phenology in the interplay of Earth's surface and the atmosphere is thus crucial to understand how the coupled system will respond to and shape climatic changes. Phenology is controlled by meteorological conditions at different time scales: on the one hand, changes in key meteorological variables (temperature, water, radiation) can have immediate effects on the vegetation development; on the other hand, phenological changes can be driven by past environmental conditions, known as memory effects. However, the processes governing meteorological memory effects on phenology are not completely understood, resulting in their limited performance of vegetation phenology represented in land surface models. A deep learning model, specifically a long short-term memory network (LSTM), has the potential to capture and model the meteorological memory effects on vegetation phenology. Here, we apply the LSTM to model the vegetation phenology using meteorological drivers and high temporal resolution canopy greenness observations through digital repeat photography by the PhenoCam network. We compare a multiple linear regression model, a no-memory-effect, and a full-memory-effect LSTM model to predict the whole seasonal greenness trajectory and the corresponding phenological transition dates across 50 sites and 317 site-year during 2009-2018, covering deciduous broadleaf forests, evergreen needleleaf forests and grasslands. Results show that the deep learning model outperforms the multiple linear regression model, and the full-memory-effect LSTM model performs better than no-memory-effect model for all three plant function types (median R^2 of 0.878, 0.957, and 0.955 for broadleaf forests, evergreen needleleaf

forests and grasslands). We also find that the full-memory-effect LSTM model is capable of predicting the seasonal dynamic variations of canopy greenness and reproducing trends in shifting phenological transition dates. We also performed a sensitivity analysis of the full-memory-effect LSTM model to assess its plausibility, revealing its coherence with established knowledge of vegetation phenology sensitivity to meteorological conditions, particularly changes in temperature. Our study highlights that 1) multi-variate meteorological memory effects play a crucial role in vegetation phenology, and 2) deep learning opens up new avenues for improving the representation of vegetation phenological processes in land surface models via a hybrid modelling approach.

1 Introduction

Vegetation phenology characterizing key plant development stages such as leaf unfolding and leaf senescence, plays a pivotal role as primary regulator of ecosystem processes and land-atmosphere interactions (Peñuelas et al., 2009; Richardson et al., 2013; Piao et al., 2019). In response to the global change, vegetation phenology has shown divergent shifts in the diverse biomes (Menzel et al., 2006; Cleland et al., 2007; Wolkovich et al., 2012; Fu et al., 2015; Zhang et al., 2022), whilst at the same time exerting a substantial influence on ecosystem productivity and functions through the impact on biogeochemical processes, especially photosynthesis and carbon sequestration (e.g., Richardson et al., 2010) as well as ecosystem respiration (e.g., Migliavacca et al., 2015). Additionally, as green leaves are the primary interface for the exchange of energy, mass, and momentum between the terrestrial surface and planetary layer (Richardson et al., 2012), vegetation phenology plays a fundamental role in controlling seasonal dynamics of water and heat fluxes between the land and the atmosphere (Peñuelas et al., 2009; Richardson et al., 2009; Puma et al., 2013; Jin et al., 2017; Buermann et al., 2018; Koebsch et al., 2020; Wu et al., 2022). Given the significance of vegetation phenology within the Earth system, an accurate representation of vegetation phenology in land surface models (LSMs) is crucial to enhance our understanding of ecosystem processes and their dynamics in response to climate change.

Over decades, much of the modelling efforts have been made to improve the development of accurate phenological models at species-specific and plant-functional-type scale (White et al., 1997; Chuine, 2000; Jolly et al., 2005; Delpierre et al., 2009), including understanding the physiological mechanisms and environmental driving factors controlling phenology (Fu et al., 2020). Currently, vegetation phenological models mainly include statistical models and process-based models. These models are developed to simulate phenological events by integrating meteorological variables which are supposed to drive the processes of vegetation phenology, utilizing ground observations or phenological proxies derived from remote sensing vegetation index data. The most popular approach to represent phenology in land surface models is based on the accumulated growing-degree-days (GDD) (Lawrence et al., 2019; Asse et al., 2020; Pollard et al., 2020). The GDD model assumes that vegetation phenological events occur when the accumulated growing-degree-day sum fulfils a given requirement (i.e., a threshold of accumulated temperature over a certain time period). Considering the physiological processes, plants experience dormancy before entering the growing season, and thus chilling is considered to be essential to break dormancy in phenological

models (Chuine, 2000; Zhang et al., 2022). Besides temperature, also the photoperiod, and soil water availability have been shown to be important drivers for vegetation phenology (Adole et al., 2019; Borchert et al., 2005; Luo et al., 2020). Consequently, models based on GDD models have been improved by incorporating photoperiod and soil water availability effects, which have been applied in many LSMs, such as Biome-BGC (BioGeochemical Cycles) model (Thornton et al., 2002; Thornton and Rosenbloom, 2005), JSBACH (Jena Scheme for Biosphere-Atmosphere Coupling in Hamburg; (Mauritsen et al., 2019)), ORCHIDEE (ORganising Carbon and Hydrology In Dynamic Ecosystems; (Krinner et al., 2005a)) and so on. Large uncertainties and biases in modelling phenology following this ad-hoc concepts have been identified within LSMs and Earth system models (Richardson et al., 2012; Jeong et al., 2012; Murray-Tortarolo et al., 2013; Lawrence et al., 2019; Peano et al., 2021), resulting in inaccurate estimations of primary productivity and the terrestrial ecosystem carbon and water cycle (Migliavacca et al., 2012).

To improve such phenology model performance, one has to consider more complex interactions of meteorological conditions that drive the vegetation phenological development. Phenology is triggered by meteorological conditions at various time scales. Instantaneous meteorological conditions like day-to-day variations in temperature, water, radiation can directly impact vegetation development. Additionally, longer-term and past meteorological conditions from the previous month / year, have legacy effects on phenological changes. For example, a plant might face delays in budburst if the chilling requirements are not fulfilled (Ren et al., 2021). These lasting or delayed impacts are often referred to as memory effects, representing the impact of previous climate conditions on the present or future vegetation development. Studies have revealed that besides the well-known memory effect from temperature like GDD or chilling, other meteorological variables like precipitation or drought also have the memory effects on vegetation growth and subsequent phenological appearance (Walter et al., 2011; Ogle et al., 2015; Ettinger et al., 2018; Liu et al., 2018b; Lian et al., 2021). Due to the complexity and the interplay of various meteorological factors, current modelling efforts face a challenge in incorporating these multi-variate memory effects (that refers to the different memory effects that can be associated to different meteorological drivers) in a mechanistic manner.

Recently, data-driven methods including deep learning techniques have been used to investigate the influence of climatic factors on land surface processes (Forkel et al., 2017; Reichstein et al., 2019; Besnard et al., 2019; Kraft et al., 2019; Callaghan et al., 2021; Zhou et al., 2021), demonstrating their potential in capturing long-term temporal dependencies (Sutskever et al., 2014; Bahdanau et al., 2016). Deep learning models aim to consider the full spectrum of meteorological inputs making predictions and thus hold promise in capturing long-term temporal dependencies from multiple variables (Besnard et al., 2019; Kraft et al., 2019). A recent study already indicated that the deep learning technique is capable of improving the predictability skill of vegetation phenology with respect to conventional methods (Zhou et al., 2021). The long short-term memory network (LSTM), a type of deep learning neural network, designed specifically for sequence prediction problems that deal with the memory effect (Hochreiter and Schmidhuber, 1997). Also, time series of consistent and widely distributed phenological near-surface observations are now long enough so that they reflect short- as well as long-term sensitivities to meteorological conditions. Specifically, the continuous daily dataset from the PhenoCam initiative (Richardson et al., 2018), obtained from images of near-surface digital cameras, offers opportunities to apply deep learning methods for developing vegetation

phenology models that account for the memory effects of multiple meteorological variables (Richardson et al., 2018; Seyednasrollah et al., 2019).

100   We propose to use the PhenoCam data and the LSTM framework to develop a deep learning model that not only is capable of predicting specific transition dates, but also in forecasting the state of the canopy greenness throughout the entire year. Most current phenology modelling studies focus on phenological transition dates only, such as the onset of budburst, flowering, or leaf senescence (Chuine, 2000; Delpierre et al., 2009; Fu et al., 2020). These phenological models effectively capture the processes that lead to individual phenological events, but fail to overlook the dynamical nature characterizing the continuous

105   phenological development throughout the entire annual cycle, where the phenological state itself could influence the subsequent phenological development (Fu et al., 2014). Conversely, models targeting the whole seasonal trajectory (White et al., 1997) possess the capability to provide continuous predictions of the phenological development state that decisively influence other biogeochemical and –physical processes at the land surface.

     More specifically, in this study, we focus on the modelling of the whole seasonal trajectory of canopy greenness, but also

110   the prediction of transition dates in the annual cycle of canopy greenness. Overall, our key objective is to develop a robust deep learning vegetation phenology model on the basis of a LSTM to characterise the memory effects of multiple meteorological variables on canopy greenness using the PhenoCam observations. We also build a statistical model as the baseline to evaluate the performance of our machine learning model. Our study focuses on addressing the following research questions: (1) Can deep learning model perform better than statistical model? And can deep learning model accounting for

115   memory-effects of multiple meteorological variables outperform models without accounting for such memory effects? (2) Does deep learning model successfully capture temporal variations on different time scales of canopy greenness and vegetation phenology? (3) Can deep learning model provide meaningful interpretations of the underlying physical and biological relationships between vegetation greenness, phenology, and a changing climate?

2 Materials and methods

120 **2.1 PhenoCam data**

The phenological data used in this study are acquired from the PhenoCam dataset v2.0 (https://daac.ornl.gov/VEGETATION/guides/PhenoCam_V2.html, (Seyednasrollah et al., 2019)). The dataset is derived from digital images photographed by automated and high-frequency digital cameras at half-hour intervals. These images are analysed to calculate the green chromatic coordinate (GCC), which is the ratio of the green channel digital numbers to the total

125   digital values of the digital Red-Green-Blue images within a predefined region of interest. GCC serves as an indicator of canopy greenness, with variations primarily due to changes in photoprotective pigments, which cause the variation in canopy greenness. For example, the changes in photoprotective pigments cause the canopy to appear more 'red' in winter (indicating low GCC) compared to summer.

     To construct a daily GCC time series, we use the 90th percentile of daily GCC values, reducing the impact of atmospheric

130   conditions on illumination. This daily GCC time series is used in this study to represent the canopy greenness development

and senescence. Specifically, we select the GCC time series from Type I observation sites that follow a standard protocol to ensure data quality and continuity (Richardson et al., 2018). To ensure robust data, we exclude yearly GCC data for years with more than 20 days of missing digital images. We further select sites with continuous observations available for more than 5 years. Missing data in the GCC time series are interpolated using cubic spline interpolation method (Hall and Meyer, 1976).

135 Additionally, we apply a locally weighted scatterplot smoothing method to reduce noise in the GCC time series (Cleveland, 1979). Our study focuses on three main plant functional types (PFT): deciduous broadleaf forest (DB), evergreen needleleaf forest (EN), and grassland (GR), with observed GCC ranges of 0.30-0.46 for DB, 0.32-0.43 for EN, and 0.30-0.43 for GR. Ultimately, a total of 50 sites and 317 site-year observations during 2009-2018 are used in the analysis. Of these, 28 sites with 178 site-year observations are DB sites, 13 sites with 82 site-year observations are EN sites and 9 sites with 57 site-year

140 observations are GR sites. The spatial distribution of the study sites and test sites for each PFT is shown in Fig. 1.
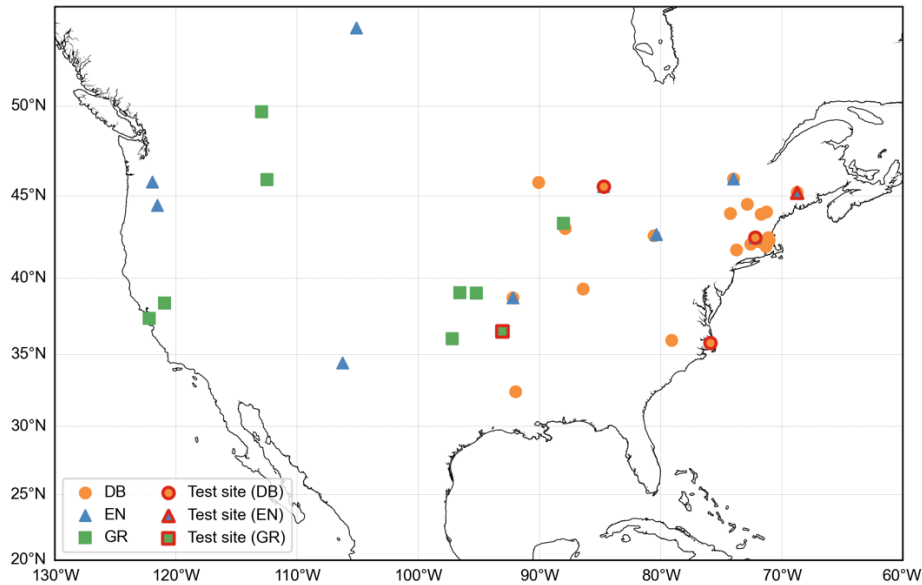
Fig. 1 Geographical distribution of study sites for deciduous broadleaf forest (orange circle), evergreen needleleaf (blue triangle) and grassland (green square). The test sites are specifically highlighted within a red frame.

2.2 Explanatory variables

145 The daily meteorological variables were obtained from the station-level Daymet dataset (https://daymet.ornl.gov/). These variables include daily minimum temperature (T_{min}), daily maximum temperature (T_{max}), daily daylength (DL), daily precipitation (P), daily water vapour pressure of the air (e_a), and daily shortwave radiation (R). We extracte time-series of these meteorological variables for each studied Phenocam site. Given that the soil moisture might have memory effects on vegetation canopy greenness development, it should be included as a driver in the model. In our study, six dynamic variables

150 are used, including T_{min}, T_{max}, DL, R, vapour pressure deficit (VPD) and soil moisture index (SMI). The daily VPD is the

difference between the saturation pressure of water (e_s, calculated from the daily mean temperature) and the actual water vapour pressure of the air (e_a), calculated following Eq. (1) & (2) (Alduchov and Eskridge, 1997). Due to the lack of measured soil moisture data, we have utilized a 30-day backward running mean of precipitation, applying decreasing weights (from 0 to 1) to days further in the past one month, as a proxy for soil moisture. This index has been demonstrated to serve effectively as a proxy for soil moisture where direct soil moisture measurements are unavailable (Migliavacca et al. 2011). The daily SMI is computed using a proxy of the sum of precipitation over the previous month (Eq. (3)). Additionally, two static variables, mean annual temperature (T_{mean}), and mean annual precipitation (P_{mean}), are derived from the records of these two variables in PhenoCam dataset.

$$VPD = e_s - e_a \tag{1}$$

$$e_s = 0.61094 \times e^{\frac{17.625T}{T+243.04}} \tag{2}$$

Where e_s and e_a are saturated and actual vapour pressure, respectively (kPa). T is the mean daily air temperature (°C).

$$SMI_t = \frac{\sum_{i=0}^{30} (P_{t-i} \times \frac{30-i}{30})}{31} \tag{3}$$

where SMI_t is the soil moisture index on day t, P_{t-i} is the precipitation on day $(t-i)$, i is the number of days away from the day of t.

2.3 LSTM modelling approach

Our goal is to predict the whole seasonal trajectory of canopy green chromatic coordinate (GCC) from the time series of the eight predictor variables using one model per PFT for multiple sites. To make this task feasible, we subtract the winter baseline value (the mean of the minimum GCC values in available years) of GCC at each site and PFT, making the measurements more comparable across sites. Further, the predictor variables and targets (GCC) are globally normalized using a min-max transformation for each PFT.

To ensure that our models learn relationships that can be generalized, we evaluate them on unseen data in space and time. For the spatial generalisation, we hold out 10% of all studied sites as unseen test sites. Furthermore, we use all data from the year of 2018 for each PFT as our temporal test dataset. The division of the data is illustrated in Fig. S1. Additionally, we divide the dataset into samples consisting of two years of input predictor variables (accounting for potential memory effects of meteorological variables from previous one year to current-year), along with one year of GCC observations corresponding to the second year of input.

To capture the relationship between the meteorological variables and the GCC, we employ a LSTM network (Hochreiter and Schmidhuber, 1997). We choose this method for several reasons. Firstly, we expect a highly non-linear relationship between the meteorological variables and observed canopy greenness, necessitating a flexible, nonparametric model such as a neural network (Hornik et al., 1989). Secondly, the representation of dynamic meteorological memory effects on canopy

greenness, requires a model that can represent temporal interactions across scales. For this purpose, we utilize a recurrent neural network (RNN), specifically the LSTM which have demonstrated strong performance in prediction related problems involving time series data (Wu et al., 2017; Besnard et al., 2019; Kraft et al., 2019, 2022).

In our study, we employ a single-layer LSTM with 128 nodes, followed by one fully connected (output-) layer and preceded by one fully connected (input-) layer. The mean squared error between the predicted and the observed GCC is optimized using the gradient based AdamW (Loshchilov and Hutter, 2019), an algorithm that adds decoupled weight decay to the Adam (Kingma and Ba, 2017) optimizer. We train an ensemble of models to reach a more stable prediction. During the model development phase, we utilized a leave-one-site-out cross-validation strategy, applying 25-fold for the DB, 12-fold for the EN, and 8-fold for the GR datasets. This approach was integral to identifying the most effective model architectures and hyperparameters, ensuring robustness across various sites. Additionally, we employed these validation datasets for early stopping to prevent overfitting. Specifically, we stop the training when the performance no longer improves on the left-out validation set of 150 epochs (an epoch refers to one complete pass through the entire training dataset). Furthermore, we decay the learning rate (a hyperparameter that determines the size of the steps taken during the optimization of a model) of, initially, 0.01 by a factor of 0.9 after each epoch. For testing we use the mean of the ensembled LSTMs as the prediction for the GCC.

To quantify the importance of the memory effects in the model, we additionally train our model on the same dataset with all data being randomly shuffled in the time dimension (Besnard et al., 2019; Kraft et al., 2019). In this dataset the "instantaneous" relation between the inputs and outputs of the current day is unimpaired but the effects of previous days cannot be learned, as these days are random. This LSTM model does not consider the memory effects, referred to as no-memory-effect LSTM model M_0. In contrast, the original model, which has access to the full history of the input variables, is referred to as the full-memory-effect LSTM model M_{full}. The framework of our LSTM model in predicting canopy greenness GCC using six dynamic and two static predictors is illustrated in Fig. 2.
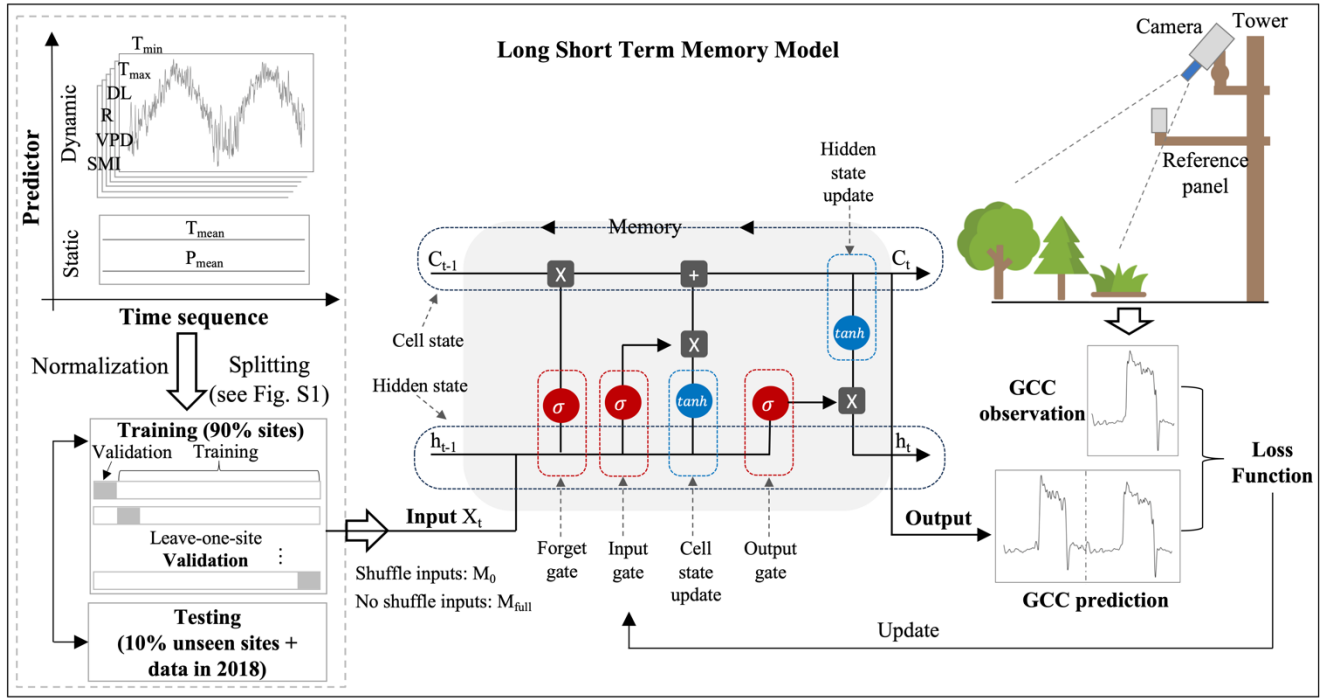
Fig. 2 The framework of canopy greenness modelling using LSTM. The LSTM is composed of a forget gate, input gate, output gate and a candidate and hidden state. The LSTM networks are adapted from Christopher Olah, http://colah.github.io/posts/2015-08-Understanding-LSTMs/.

2.4 Model evaluation

To assess the modelling ability of deep learning model, we develop a baseline model using multiple linear regression (MLR) between the eight predictor variables and GCC (Eq. (4)). The MLR model is trained and tested in the same training and testing dataset as the LSTM models.

$$GCC = a * T_{min} + b * T_{max} + c * DL + d * R + e * VPD + f * SW + g * T_{mean} + k * P_{mean} + res \qquad (4)$$

where GCC is our target, $T_{min}, T_{max}, DL, R, VPD, SW, T_{mean}$, and P_{mean} are predictor variables, res is the residual.

For model evaluation, we primarily use root mean square error (RMSE) and the coefficient of determination (R^2) for model evaluation. These metrics are calculated based on the predicted and observed GCC at each site for each PFT. We compare the model performance of all models in the testing dataset, and select the best model for each PFT by R^2. For the best models, we evaluate their performance in simulating: 1) GCC observations; 2) GCC temporal variation; 3) phenological transition dates in testing datasets. GCC temporal variation includes three time-scales variation: daily variation, monthly mean GCC variation, and interannual variation of the anomalies of median GCC. For daily variation, we also calculate the daily anomalies for observations and LSTM models. For the monthly GCC variation, we further compare the mean monthly canopy development

rate (V_{GCC}) during studied years between observed and predicted GCC time series. The rate of canopy development for each month (from February to December) is calculated at the monthly scale from the GCC time series using Eq. (5).

$$V_{GCC(t)} = GCC_t - GCC_{t-1} \tag{5}$$

where t is time (month), GCC_t and GCC_{t-1} is the mean GCC for a given year at month t and $t-1$, respectively.

The phenological transition dates are estimated intending to define the start of season (SOS) and the end of season (EOS) for one year. We choose the dates corresponding to a 30% of the seasonal amplitude (from the 5th percentile to the 95th percentile) through greening rising and falling to represent the start of season (SOS) and end of season (EOS). SOS and EOS transition dates are estimated from both the predicted and observed GCC time series.

2.5 Model sensitivity analysis

In order to gain insights into the physical implications of deep learning models, we conduct two simple experiments to assess the model sensitivity to meteorological drivers. First, we increase (warming) and decrease (cooling) temperature (T_{min} and T_{max}) by 4 °C throughout the year while keeping all other predictors unchanged. Another experiment involves the same temperature adjustments for T_{min} and T_{max}, but with the VPD varying based on mean temperature while keeping other predictors constant. We then compare the annual canopy greenness cycles under warming and cooling conditions with the actual observations. Furthermore, to better understand how vegetation phenology shifts in response to a warming environment according to the LSTM models, we estimate the SOS and EOS in these two experiments. We evaluate the differences between the treatment of 1°C warming and the observations for SOS and EOS.

3 Results

3.1 Model performance

A comparison of the model performance is conducted between the statistical MLR model and the LSTM models (including the no-memory-effect LSTM model M_0 and the full-memory-effect LSTM model M_{full}) on the test dataset for deciduous broadleaf (DB), evergreen needleleaf (EN) and grassland (GR) (Table 1). The LSTM models achieve a better performance for predicting the GCCs than the MLR model for all PFTs (Table 1). The coefficient of determination R^2 between modelled and observed canopy greenness GCC are much higher in LSTM models than MLR, with the median R^2 increased from MLR to LSTM from 0.779 to more than 0.806 for DB, from 0.777 to more than 0.830 for EN, and from 0.646 to more than 0.914 for GR. Similarly, the Root Mean Square Error (RMSE) values corroborated the R^2 findings, demonstrating that LSTM models generally exhibit lower prediction errors than the MLR model. In conclusion, LSTM models significantly enhance the accuracy of GCC predictions across all three PFTs when compared to the baseline MLR model.

Metric	Model	DB	EN	GR
R^2	MLR	0.779 (± 0.052)	0.777 (± 0.115)	0.646 (± 0.110)

	M_0	0.806 (± 0.073)	0.830 (± 0.155)	0.914 (± 0.033)
	M_{full}	0.878 (± 0.107)	0.957 (± 0.071)	0.955 (± 0.030)
	MLR	0.021 (± 0.005)	0.019 (± 0.008)	0.018 (± 0.003)
RMSE	M_0	0.030 (± 0.015)	0.017 (± 0.006)	0.012 (± 0.003)
	M_{full}	0.028 (± 0.016)	0.012 (± 0.006)	0.009 (± 0.002)

Table 1 Coefficient of determination (R^2) and root mean squared error (RMSE) comparisons (ensemble median ± std estimate of all study
250 sites) for multiple linear regression model (MLR), no-memory-effect LSTM model (M_0) and full-memory-effect LSTM model (M_{full}) on test dataset for deciduous broadleaf (DB), evergreen needleleaf (EN) and grassland (GR).

Furthermore, comparing the two different LSTM models, the full-memory-effect model M_{full} exhibits superior performance in simulating GCC than no-memory-effect model M_0 across all three PFTs (Fig. 3). The median R^2 of all studied sites in the

255 full-memory-effect model exceeds 0.85, specifically 0.878 for DB, 0.957 for EN, and 0.955 for GR. This represents an improvement in model performance of 8.9% for DB, 15.3% for EN, and 4.5% for GR, compared to the no-memory-effect model with R^2 of around 0.806 (DB), 0.830 (EN) and 0.914 (GR). Similarly, there is a reduction in bias of 12.5% (RMSE decreased from 0.036 in M_0 to 0.032 in M_{full}) for DB, 15.4% (RMSE decreased from 0.015 in M_0 to 0.013 in M_{full}) for EN, and 37.5% (RMSE decreased from 0.011 in M_0 to 0.008 in M_{full}) for GR in full-memory-effect models. These findings suggest

260 considering memory effects from multiple meteorological factors can enhance the model performance in simulating GCC compared to models without considering memory effects.
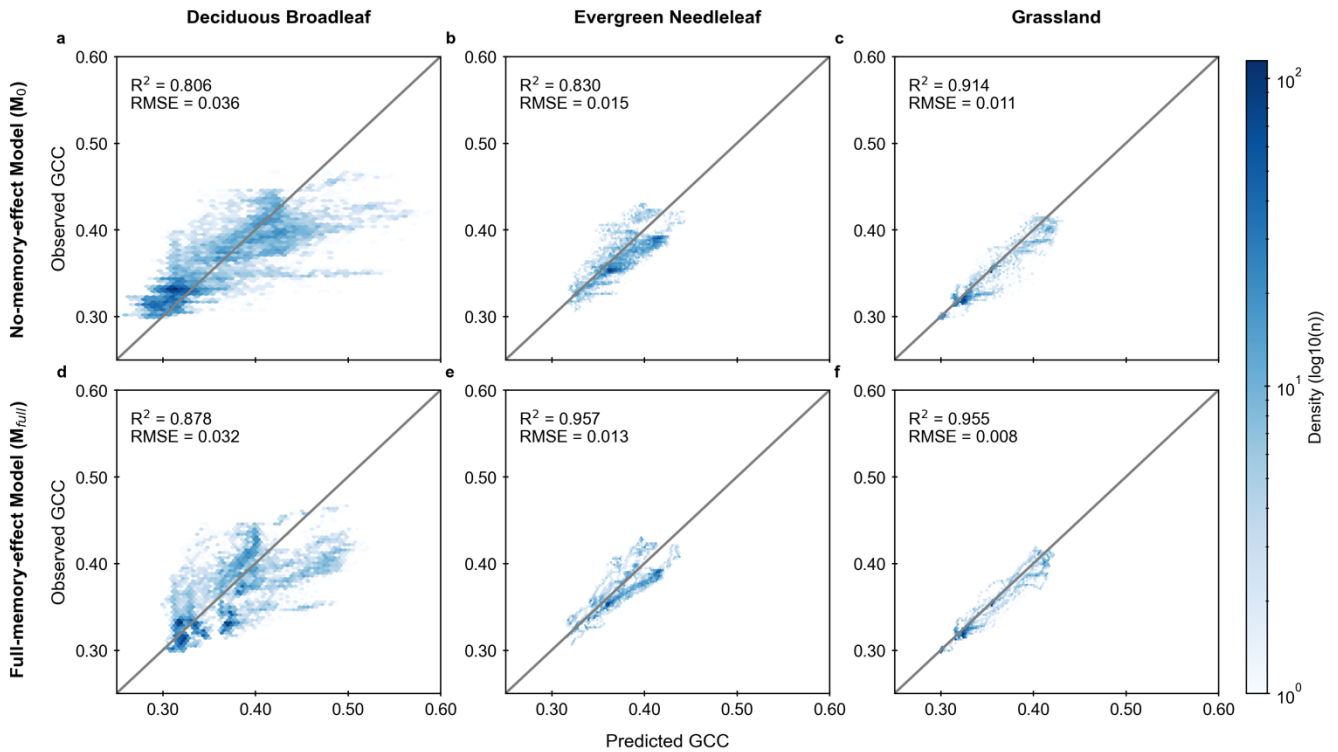
Fig. 3 Model performance of GCC time-series estimation using no-memory-effect model (M_0: **a**, **b**, **c**) and full-memory-effect model (M_{full}: **d**, **e**, **f**) in all test dataset for deciduous broadleaf (DB), evergreen needleleaf (EN) and grassland (GR). The colour indicates the density of points (light blue is lower density, dark blue is higher density). The solid grey lines denote the 1:1 line.

The performance of LSTM models on unseen test sets, both spatially and temporally, shows that full-memory-effect LSTM model M_{full} outperforms no-memory-effect LSTM model M_0 in predicting GCC for unseen site(s) and unseen year across all three studied PFTs (Fig. 4). Specifically, when evaluating performance on unseen sites, M_{full} consistently exhibits higher median R^2 values compared to M_0, with improvements of 9.3%, 13.5%, and 3.5% for DB, EN, and GR respectively. Similarly, for predictions across unseen years, M_{full} demonstrates substantial enhancement in predictive accuracy, with median R^2 values increasing from 0.805 to 0.874 (an 8.6% improvement) for DB, from 0.824 to 0.956 (a 16% improvement) for EN, and from 0.914 to 0.964 (a 5.5% improvement) for GR. The t-test comparing M_0 and M_{full} across all test sets and PFTs indicate a significant difference (P < 0.05), confirming Mfull's superior performance. These findings underscore the robustness of the model in accurately forecasting GCC time series, even when face with previously unobserved spatial and temporal contexts.
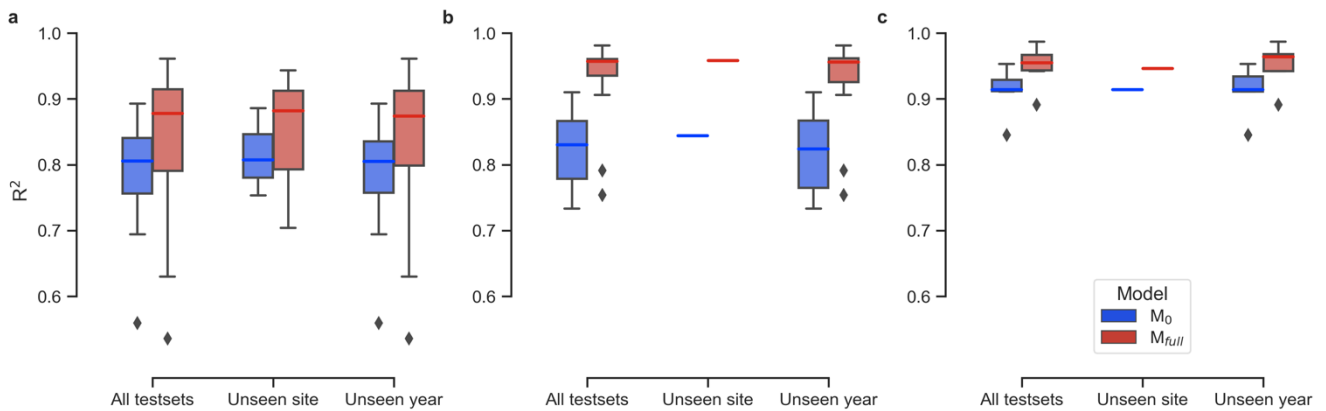
Fig. 4 Coefficient of determination (R^2) comparisons between no-memory-effect model (M_0: blue box) and full-memory-effect model (M_{full}: red box) in all test sets, unseen site and unseen year for deciduous broadleaf (DB, **a**), evergreen needleleaf (EN, **b**) and grassland (GR, **c**). The rhombus in the figure represents the outliers, which are defined as the points beyond 1.5 times the interquartile range (the difference between the 75th and 25th percentiles).

3.2 Modelling temporal variability of GCC at unseen sites: daily to interannual time scales

LSTM models can capture the GCC canopy greenness temporal dynamics, with an initial increase followed by a decrease during the growing season. This is illustrated in Fig.4, which displays the observed and predicted daily variability of GCC, daily GCC anomaly, seasonal variability of GCC and its development rate, and interannual variability of GCC anomalies at unseen sites from 2009 to 2018 using LSTM models. The predicted daily variability of GCC shows a strong correlation with observations across multiple years, with a significant correlation coefficient (r) above 0.9 (Fig. 5 **a** – **c**) in M_{full}. Conversely, M_0 shows increased noise in daily GCC variability, displaying larger biases in predicting both GCC peaks and minimum values, particularly evident for DB (Fig. 5 **a**). From the predictions of daily GCC anomalies which remove the mean seasonal cycle (Fig. 5 **d** – **f**), we can further find M_{full} performs better than M_0 in capturing the daily fluctuation though they are still not good at predicting the daily GCC anomalies as we expected. The R^2 between observed and predicted GCC anomalies are higher in M_{full} than in M_0 for DB (M_0: 0.0007, M_{full}: 0.02), EN (M_0: 0.03, M_{full}: 0.22), and GR (M_0: 0.07, M_{full}: 0.32). However, a discrepancy between observed and predicted absolute GCC at the daily scale is observed for EN in the unseen site, where the predicted GCC is overestimated compared to the observation in both M_{full} and M_0 (Fig. 5 **b**).

The overall seasonal cycle of monthly GCC shows good match between the observation and prediction by LSTM models. The observed GCC typically starts to increase in March (GR) or April (DB and EN), peaks in June (DB and GR) or July (EN), and gradually decreases until November (DB and GR) or December (EN) (Fig. 5 **g** – **i**). Both M_{full} and M_0 can effectively capture this seasonal pattern for EN and GR. However, for DB, the M_{full} model predicts a similar seasonal dynamic pattern of GCC to observations, depicting greening up before June or July followed by greening down until November or December, while M_0 predicts a peak in greenness occurring in July which diverges from observation (Fig. 5 **g**). Similarly, the development

rate of monthly GCC shows similar performance as seasonal cycle of monthly GCC. The largest increase in observed GCC occurs in May during the greening-up period for all three vegetation types, while the speed of greenness decreases significantly accelerates around October (grey box in Fig. 5 **g** – **i**). The LSTMs predict developed rates in greening up and down exhibit a similar pattern to the observations for EN and GR (red box in Fig. 5 **g** – **i**). However, for DB, both M_{full} and M_0 predict the highest development rate in June during the green-up period and in September during the green-down period, in contrast to observations which indicate peak rates in May and October, respectively.

The predicted interannual variability of maximum GCC anomalies shows that both LSTM models with full-memory effects and no-memory effects can generally forecast trends of interannual variability of maximum GCC for EN and GR (Fig. 5**k** – **l**). However, an increasing trend (0.026 per decade) in greenness is observed in maximum GCC anomalies from harvardbarn2 for DB during the period from 2012 to 2018 (see Fig. 5**j**), and this trend is well predicted by the M_{full} model (0.015 per decade) but failed in the M_0 model (-0.004 per decade). Specifically, from 2012 to 2015, the observed maximum GCC exhibits a continual increase, whereas M_0 indicates a continual declining maximum GCC (see Fig. 5j). Furthermore, a larger bias is found in M_0 compared to M_{full} in predicting the annual maximum GCC for DB.
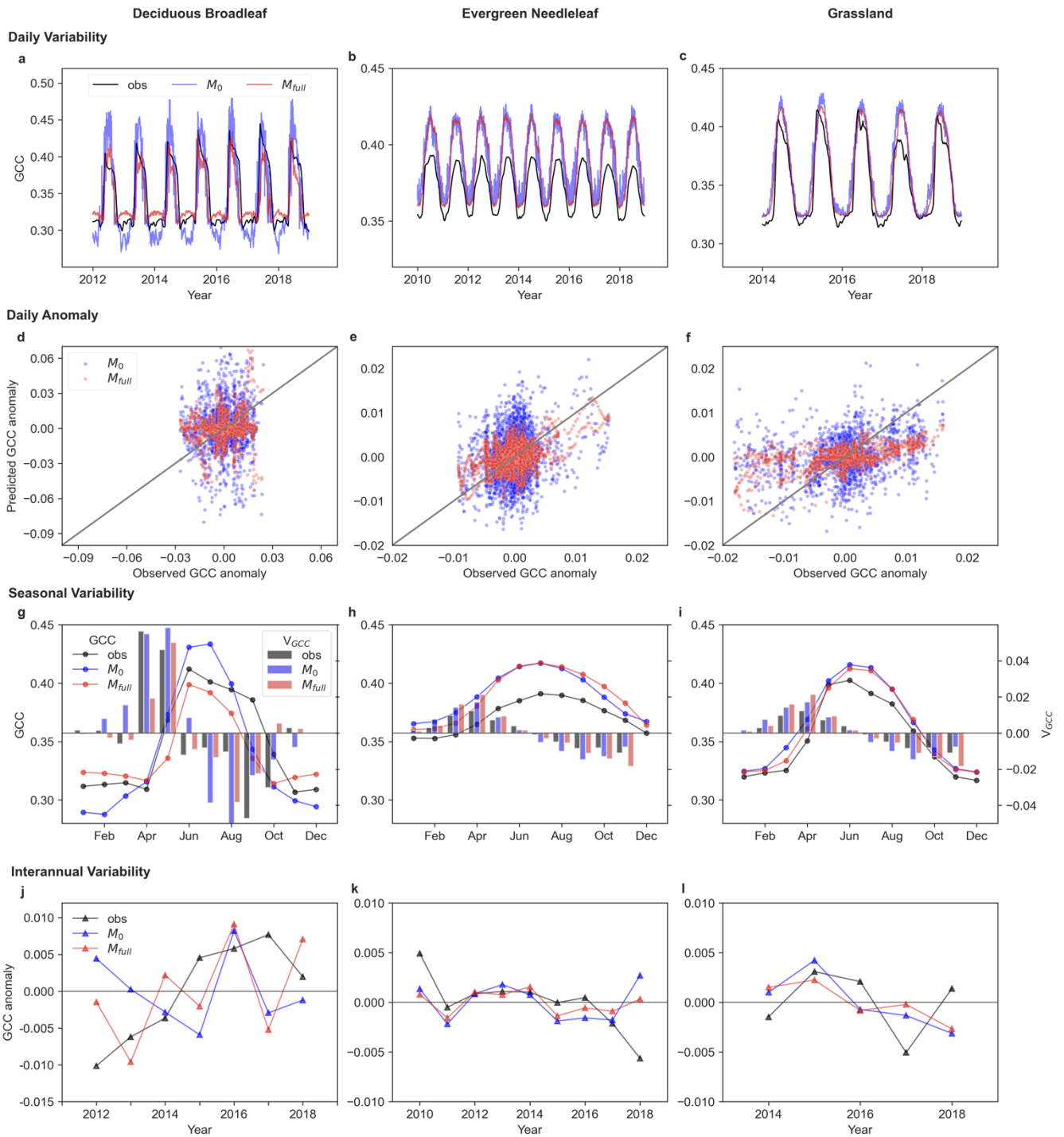
Fig. 5 Observed (obs, grey) and predicted (M₀: blue, M_full: red) daily (**a**, **b**, **c**), seasonal (**g**, **h**, **i**), and interannual (**j**, **k**, **l**) variability of canopy greenness (GCC) and daily GCC anomaly (**d**, **e**, **f**) for deciduous broadleaf (DB), evergreen needleleaf (EN) and grassland (GR) in

315

unseen sites (DB: harvardbarn2, EN: howland1, GR: bullshoals). In panel **d - f**, the solid grey lines denote the 1:1 line. In panel **g - i**, the development rate of monthly GCC (V_{GCC}) is represented by bar plots (right y-axis; red bar: observed V_{GCC}; red bar: predicted V_{GCC}). In panel **j - l**, the interannual variability of annual maximum GCC anomalies are shown.

3.3 Modelling the vegetation canopy phenological transition dates in unseen sites

Figure 6 illustrates that M_{full} can capture the interannual variability of phenological transition dates, outperforming M_0. Regarding the start of season (SOS) (Fig. 6 **a**, **b**, **c**), M_{full} consistently exhibits the same shift direction (sign of the anomaly) as observations in the majority of years. Specifically, we observe concordance in the direction of advance or delay between prediction and observation in 5 out of 7 years (71% of the years) for DB. Similarly, the predicted SOS shifts agreed well with the observations in more than 80% years for EN (89% of the years) and GR (80% of the years). Moreover, a high correlation is evident between the M_{full} predicted interannual variability of SOS anomaly and the observed interannual variability of SOS anomaly. The correlation coefficient between observed and predicted interannual variability of SOS anomaly reaches up to more than 0.9 for EN and GR. In further, the observed delay of SOS with the trend of 2.1 days per year is also reproduced well by the M_{full} (predicted trend is 1.2 days/year) in howland1 for EN during 2010 to 2018 (Fig. 6 **b**). Conversely, compared to M_{full}, M_0 exhibits poor performance in capturing the shift direction and lower correlation between the predicted interannual variability of SOS anomaly and the observed interannual variability of SOS anomaly.

As for EOS, good agreements between the predicted interannual dynamics of EOS anomalies by M_{full} and the observed ones are found in DB and EN, although not in GR (Fig. 6 **d** – **f**). Firstly, M_{full} predicted shift directions of EOS in most years (over 80%) are consistent with the observed shifts for each PFT. For DB, there are 86% of years showing the same shift direction of advance or delay between observed and predicted EOS. For EN and GR, the percentages are 89% and 80% respectively. On the other hand, we find a positive correlation between observed and M_{full} predicted EOS anomalies for DB and EN. The observed delayed trend (2.7 days/year) for DB in harvardbarn2 is also well predicted (1.6 days/year) by the M_{full}. Interestingly, M_{full} also captures the larger advancement of EOS observed in 2018 compared to the mean EOS for EN, indicating its capability to capture extreme interannual anomalies. Compared M_{full}, M_0 shows a relatively poor performance, displaying larger bias in its predictions.
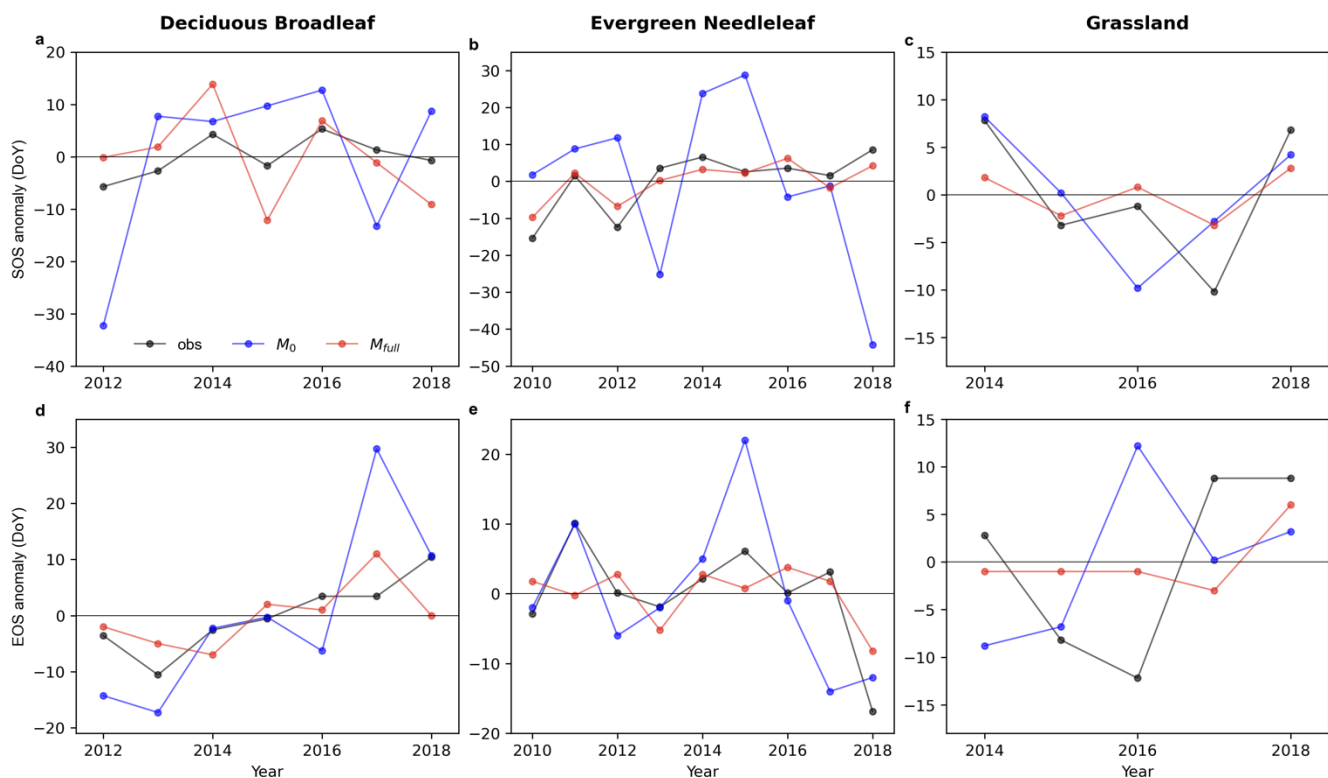
Fig. 6 Observed (obs, black line) and predicted (M$_0$: blue, M$_{full}$: red) the interannual variability in anomaly of start of season (SOS, **a**, **b**, **c**) and end of season (EOS, **d**, **e**, **f**) for deciduous broadleaf, evergreen needleleaf and grassland in unseen sites (DB: harvardbarn2, EN: howland1, GR: bullshoals).

3.4 The model sensitivity analysis

The model sensitivity analysis indicates that full-memory-effect model M$_{full}$ can simulate well the GCC response to temperature. Fig. 7 illustrates the temperature sensitivity of GCC in the LSTM model M$_{full}$ for all three PFTs studied here. Comparison of warming (red line, Fig. 7 **a**, **b**, **c**) and cooling (blue line, Fig. 7 **a**, **b**, **c**) alone treatments (increasing or decreasing 4 °C) to the unchanged temperature (± 0 °C) control (grey line, Fig. 7 **a**, **b**, **c**) reveals that warming led to greener and longer greenness season, while cooling caused to the less greening and shorter vegetation season for studied three PFTs (Fig. 7 **a**, **b**, **c**). When VPD varies with temperature, high temperature along with the high VPD does not have a significant effect on the greenness and length of vegetation period. During the greenness rising and falling period, the canopy greenness is very similar to the actual GCC cycle (the control) for the three PFTs, but the peak greenness is lower than the actual GCC peak values, especially for DB and GR (Fig. 7 **d**, **e**, **f**). In the cooling and lower VPD treatment, it shows similar trends with the cooling condition but unchanged VPD. Cooling and lower VPD result in declining canopy greenness and shortened vegetation periods during the growing season.
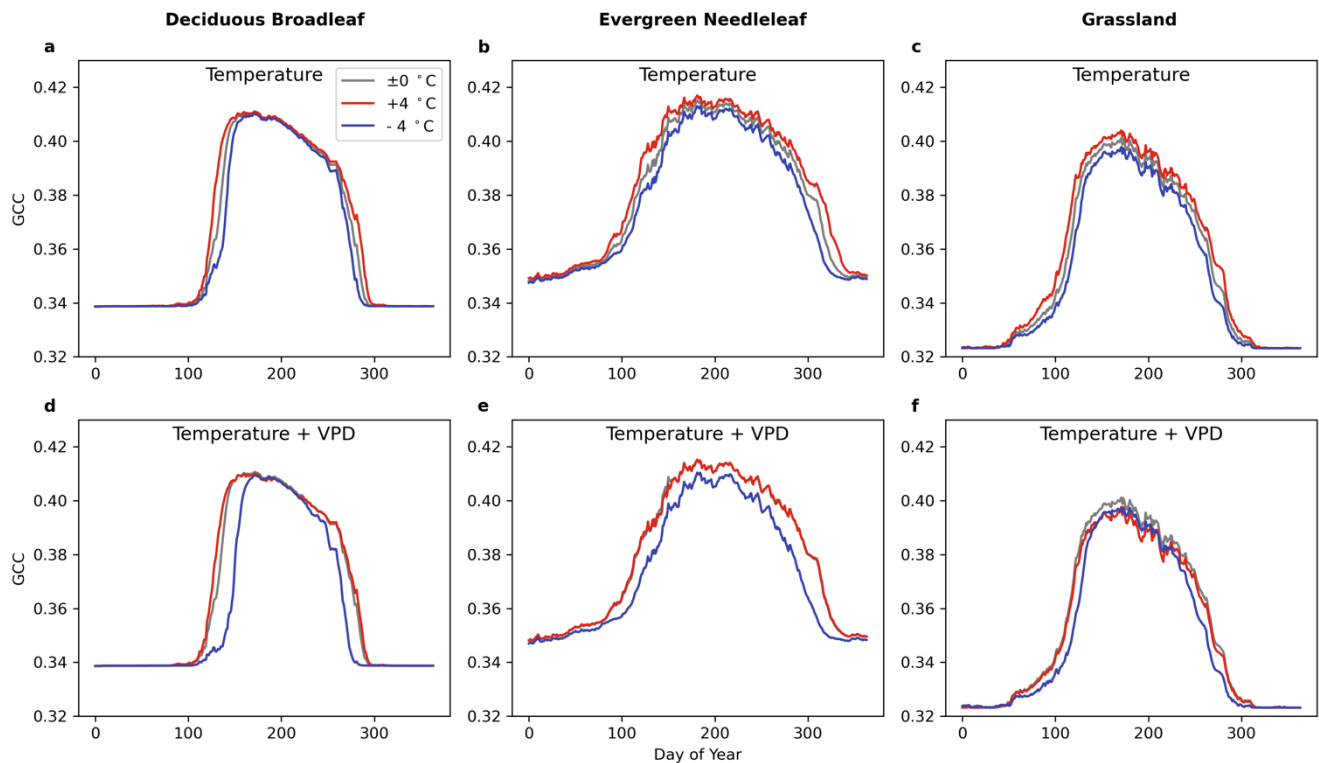
Fig. 7 The sensitivity of canopy greenness (GCC) to temperature change only and both temperature and VPD change under 4 °C warming / cooling (red line / blue line) in the all year using M_{full} for deciduous broadleaf (in howland2), evergreen needleleaf (in laurentides) and grassland (in bullshoals).

Furthermore, we also examine the temperature sensitivity of phenological events (SOS and EOS) (Fig. 8). A one-degree increase in temperature throughout the year resulted in an earlier start of the season (Fig. 8 **a**) and delayed end of the season (Fig. 8 **b**), regardless of low or high VPD. Under warm conditions alone, SOS appears to be one day earlier on average, while under warm and high VPD conditions, it shiftes to two days earlier. The one-degree temperature increase has a similar effect on EOS compared to the one-degree increase accompanied by varied VPD. Through student t-tests on means of the two distributions, no statistically significant (p = 0.12 (SOS), p = 0.69 (EOS)) differences in means are found, indicating that temperature is the most influential meteorological factor affecting the start and end of the season.
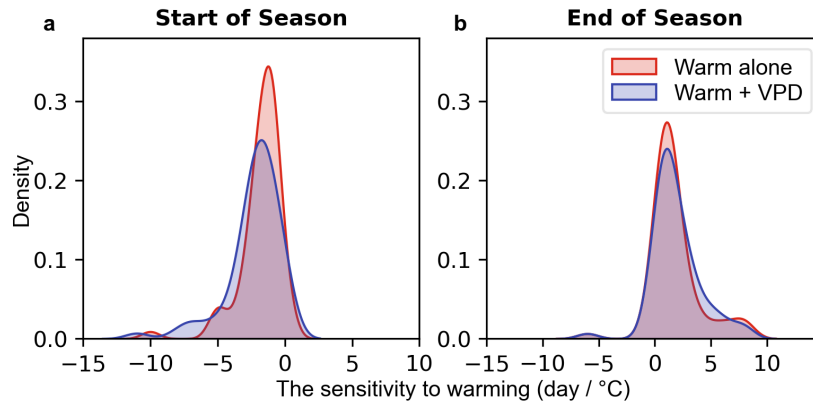
Fig. 8 Temperature sensitivity of start of season (SOS, Fig. 8 **a**) and end of season (EOS, Fig. 8 **b**) under warm alone (red) and warm and high VPD conditions (blue) over all three PFTs.

4 Discussion

4.1 Meteorological memory effects on vegetation canopy greenness

In our study, we have presented a new way to simulate canopy greenness dynamics by applying a data-driven LSTM model accounting for multi-variate meteorological memory. We find that multi-variate meteorological memory is of importance in developing vegetation phenological models. The impact of meteorological factors on vegetation phenological development encompasses both instantaneous and memory effects. Through a comparison of models accounting solely for instantaneous effects (MLR and M_0), with those considering both instantaneous and memory effects of multiple meteorological variables (M_{full}), we have demonstrated that models involving memory effects do outperform models without memory effects (Table 1 & Fig. 3). This suggests that considering both instantaneous and memory effects provides a more comprehensive explanation for vegetation development compared to solely instantaneous effects.

But what specific advantages does the full-memory-effect model offer over the no-memory-effect model? We will explore this question from several perspectives. Firstly, full-memory-effect model exhibits good performance in spatial and temporal extrapolation of canopy greenness. By comparing the model performance of M_{full} and M_0 in unseen site(s) and unseen years (Fig. 4), it becomes clear that the full-memory-effect model outperforms the no-memory-effect model in both unseen site(s) and unseen years for all three PFTs. This indicates that incorporating memory effects into the model enhances performance in unseen sites and years, a challenging task in modelling, especially for unseen sites.

Secondly, the inclusion of memory effects in models improves performance in predicting variabilities across different time scales for unseen sites. At the daily scale, the full-memory-effect model reduces noise on each day and predicts daily anomalies more accurately than the no-memory-effect model (Fig. 5). This underscores that daily changes in canopy greenness are influenced not only by instantaneous climate but also by the memory effects of previous climate on the canopy. Our results align with previous studies indicating that temperature memory (cumulative thermal summation), rather than daily temperature

alone, determines vegetation phenology (Hänninen, 1990; Chuine, 2000). In addition, for the challenges in simulating daily anomalies, our findings reveals that full-memory-effect model performs better in predicting daily anamalies compared to no-memory-effect model (Fig. 3). This finding indicates the significance of memory effects in enhancing the model's capability to simulate daily anomalies. Regarding seasonal dynamics and interannual variability, our study finds that memory effects vary among PFTs. For deciduous broadleaf trees, the full-memory-effect model demonstrates a significant advantage in predicting seasonal and interannual dynamics (Fig. 5 **g** – **l**). It can capture well the seasonal dynamic pattern and the greening trend, which the no-memory-effect model fails to predict. This suggests that changes in canopy greenness over long time scales for deciduous broadleaf trees are sensitive to relatively long-term meteorological changes. This may be attributed to lagged effects of precipitation (Joshi et al., 2022), drought (Peng et al., 2019) and other factors (Gömöry et al., 2015; Ding et al., 2020; Joshi et al., 2022; Zhou et al., 2022; Liu et al., 2018, GCB) from previous months on canopy greenness. However, for evergreen and grasslands, both the full-memory-effect model and the no-memory-effect model show similar performance in predicting seasonal dynamics and interannual variability. It is noteworthy that the memory effect of precipitation in our study is already included in the no-memory-effect model, as the meteorological variable of soil moisture is calculated based on the weighted mean of precipitation in the previous month (due to not available soil moisture data). This implies that such memory effects may offset the performance difference between the full-memory-effect model and the no-memory-effect model.

Lastly, incorporating multi-variate meteorological memory effects into the LSTM model improves performance in predicting vegetation phenology (Fig. 6). Our results suggest that phenological shifts are influenced by meteorological memory effects, consistent with the notion that vegetation phenology is highly variable and responsive to long-term variation in climate (Sparks and Carey, 1995). Specifically, winter chilling (Chuine et al., 2016; Ettinger et al., 2020; Zhang et al., 2022), and the growing season temperature (Liu et al., 2018, GCB) can impact on the spring phenology and autumn development. However, unlike models primarily accounting for temperature memory effects alone, such as GDD (Hänninen, 1990; Chuine, 2000), our LSTM memory effect model shows promise in the incorporation of multiple memory effects from different meteorological variables.

It should be noted that although our study emphasizes the importance of memory effects of multiple meteorological variables, the specific contributions of different meteorological factors to memory effects on vegetation development remain unclear. Further in-depth studies of memory effects are still needed to discern the relative importance of memory for each meteorological factor, and their memory length across various developmental stages.

4.2 Machine learning modelling of vegetation phenology

In our study, we explore the potential of a deep learning approach using LSTM to predict vegetation phenology based on canopy greenness, specifically GCC annual cycles, using only meteorological variables as inputs. The results indicate that the superior performance of our deep learning model compared to a multiple linear regression model (Table 1), highlighting that deep learning models are capable of capturing nonlinear relationships between inputs and targets. This holds promises for improving the performance of current vegetation phenology models and a significant step toward a better representation of

phenology on earth system models using deep learning approaches. However, comparing our deep learning model with process-based model is still challenging, as their modelling targets are different in many cases. Our deep learning models focus on the whole annual cycle of canopy greenness, whereas most process-based models concentrate on specific phenological events.

430   The deep learning model performance across PFT (Table 1, Fig3) shows that the model performs better for EN compared DB. The superior performance in estimating GCC for the EN compared to the DB might be attributed to two main factors: 1) Spring and autumn variability in GCC: The EN exhibits more gradual changes in GCC during spring and autumn, facilitating more accurate model simulations of GCC. This contrasts with the DB, where abrupt environmental changes lead to more volatile GCC values. 2) Memory effect: The EN may have a longer memory effect, meaning its GCC values are influenced by past conditions over a longer period. In contrast, the GCC for DB respond more immediately to current environmental factors.

435

Our findings demonstrate that full-memory-effect LSTM model can generally explain daily-scale variations and seasonal dynamic changes, as evidenced by the high correlations between predicted and observed GCC (Fig. 5). This is likely because the full-memory-effect LSTM model can better learn the complex relationship between climatic dynamics and canopy greenness dynamics. However, our deep learning model has had less success in accurately predicting absolute GCC values and peak values (Fig. 5, Fig. S2). For instance, in the case of "howland1" site for evergreen needleleaf, the full-memory-effect LSTM model can predict the dynamics of canopy greenness well, but the absolute GCC values are overestimated (Fig. 5 **b**). Indeed, although full-memory-effect model show a good performance, it also overestimates canopy greenness in some sites (Fig. 5 **b,** Fig. 8 **a & c**, Fig. S2), and underestimates in other sites (Fig. 9 **b & d**), Fig. S2). Possible reasons for this discrepancy could be: 1) different climatic drivers for different species among PFTs, 2) incomparable GCC data among sites, and 3)

440

445   inadequate learning of site-specific characteristics. We aim to build a more general model in this study, but it should be noted that even within species, GCC can respond to climate differently to meteorological conditions (Denéchère et al., 2021). Additionally, the combination of GCC data from all sites for a specific PFT in building the model may introduce errors and bias, as GCC data is not consistently calibrated and the colour signals can be sensitive to various parameters, such as camera type, species (foliage colours are different colours of green) or spectral properties of incoming light (Wingate et al., 2015; Richardson et al., 2018). Moreover, the use of static variables (climatological mean temperature and precipitation) to indicate spatial differences may not sufficiently capture site-specific information, leading to overestimation or underestimation of specific sites. Although the deep learning models have limitation in simulating absolute GCC values accurately, it's important to note that the bias in absolute values is less significant compared to the seasonal dynamics of GCC which is used for detecting phenology. Regarding interannual variability, we find that the predicted changes (increase or decrease) in peak GCC are

450

455   consistent with observations in most years (Fig. 5 **g** – **i**), indicating that the model's capability to reproduce basic response of canopy greenness to climate changes. Furthermore, our models can also capture well the interannual variability in GCC (Fig. 5 **k**), and also the trends of interannual variability in anomalies of peak GCC were generated well by our data-driven models (Fig. 5**j**). Overall, these results demonstrate the ability of the LSTM model to reliably predict temporal variability. In terms of spatial performance, we find that the LSTM model has a good agreement with observations in most studied sites for all three

460   PFTs (Fig. S3). This means our model is able to capture spatial variation within each PFT providing support that the model might represent a general model for each PFT.
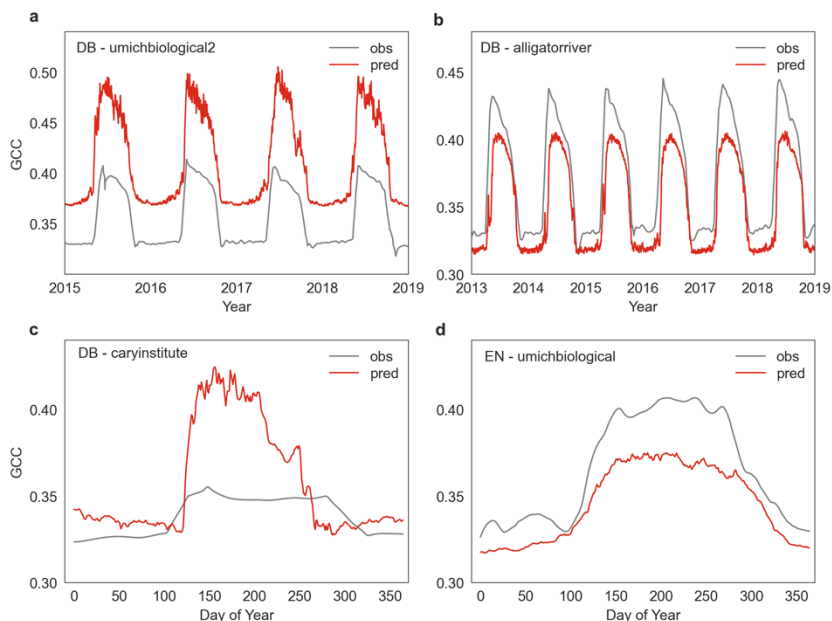
Fig. 9 The overestimated and underestimated canopy greenness (GCC) by the LSTM model in some sites (umichbiological2, alligatorriver, caryinstitute, umichbiological) for deciduous broadleaf (DB) and evergreen needleleaf (EN).

465

The modelling results of vegetation phenology reveal that the full-memory-effect LSTM model is also capable of predicting the shift in phenological transition dates (advance or delay of start of season (SOS) and end of season (EOS)) in most years in the unseen dataset (Fig. 7). The ability to predict the advancement or delay in phenology is crucial for estimating other key processes in the ecosystem functioning, such as ecosystem productivity, as the advancement of spring phenology and the delay

470   of autumn phenology are typically associated with higher productivity (Richardson et al., 2010). Overall, our model's skill to accurately predict the average advance or delay in phenology is encouraging, although, it remains challenging to predict the exact phenological dates given the potential and systematic overestimation or underestimation in the GCC cycle (Fig. S4).

4.3 Can the deep learning model of vegetation phenology learn physically plausible relationships?

The sensitivity analysis of our deep learning model sheds light on its ability to learn meaningful physical insights. The

475   model responds to warmer temperatures by predicting an earlier spring onset and later autumn senescence, which is in line with findings from other studies (Menzel et al., 2006; Jeong, 2020). These results underscore the capability of our deep learning framework to reproduce the sensitivity of canopy greenness and phenology to temperature. Our current study primarily focuses on developing the deep learning model. We also conduct a basic sensitivity analysis that has the potential to dismantle the LSTM model and learn from the identified relationship in the data how the response (i.e. GCC. A more extensive and

comprehensive sensitivity analyses of the LSTM model, as well as interventional experiments, could offer insights into understanding phenology by identifying which predictors are influential and when. Especially, such approach might help to uncover the control of autumn phenology and its modelling – a long-standing challenge faced by process-based models that may struggle due to inadequate predictors inclusion or response functions (Delpierre et al., 2009; Liu et al., 2019). Moreover, the hybrid models by the integration of physic knowledge into the deep learning models might enhance our understanding of how climate change impact on phenology and associated consequences for the ecosystem, another key challenge in phenology modelling. In numerous Earth system models, such as CLM 4.5 and LPJ (Peano et al. 2021), phenology is modelled based solely on climatic drivers, employing PFT-specific thresholds for factors like chilling and growing degree days. Some models also integrate a more realistic connection to the carbon cycle, where leaf growth incurs a carbon cost. For the phenology representation strictly from climate variables in ESMs, our data-driven approach can serve as a direct replacement for traditional empirical formulations. However, for ESMs that consider phenology to be dependent on available carbon resources, it becomes necessary to evolve our data-driven method into a hybrid model (e.g., ElGhawi et al. 2023 for land-atmosphere fluxes) that also accounts for carbon resources in its inputs and loss function. This entails modelling leaf growth as depleting carbon from the reserve pool and adding dropped leaves to the humus pool, thereby ensuring carbon mass balance. Consequently, carbon mass balance serves as a critical constraint for the data-driven phenology model.

5 Conclusions

In this study, we develop a novel deep learning modelling framework incorporating multiple meteorological memory effects to predict the whole seasonal trajectory of canopy greenness and transition dates for each plant functional type using LSTM. Our key findings can be summarized as follows:

1) The general deep learning model, trained for each PFT using LSTM, demonstrates the ability to generalize to unseen sites, indicating that the deep learning approach effectively captures the underlying mechanics of canopy greenness development.

2) The incorporation of multi-variate meteorological memory effects proves crucial in canopy greenness modelling. The LSTM model, accounting for these memory effects, can reproduce general temporal dynamics of canopy greenness across various time scales, from daily to inter-annual variability. Furthermore, it captures phenological shift directions, enhancing the model's comprehensive representation.

3) Our sensitivity analysis demonstrates the LSTM model's capability to learn plausible relationships, revealing its proficiency in acquiring fundamental physical knowledge about vegetation greenness and phenological development.

Our deep learning model accounting for multi-variate memory effects holds promise for improving our understanding of vegetation responses to climatic variability. In future, the integration of deep learning phenology models into coupled land-

surface and earth system models, may further enhance our ability to comprehend and simulate complex interactions and feedback within these systems.

Code availability

All the code are available on Zenodo (https://doi.org/10.5281/zenodo.10668592).

515 **Author contributions**

GL, AW, CR and MM led the LSTM model development and evaluation process and prepared the manuscript. GL, CR and BK contributed to the development of the codes. All co-authors contributed to the writing and editing of the manuscript.

Competing interests

The contact author has declared that none of the authors has any competing interests.

520 **Disclaimer**

Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Acknowledgements

Guohua Liu acknowledges financial support from Sino-German (CSC-DAAD) Postdoc Scholarship to finish this study. Mirco
525 Migliavacca and Alexander J. Winkler acknowledges support by Deutsche Forschungsgemeinschaft (DFG) [PhenoFeedBacks project, grant number MI 2418/4-1]. Christian Reimers and Alexander J. Winkler were supported by the European Research Council (ERC) Synergy Grant "Understanding and Modelling the Earth System with Machine Learning (USMILE)" under the Horizon 2020 research and innovation 415 programme (Grant agreement No. 855187). We thank Ana Bastos for insightful suggestions for revising the manuscript. We thank our many collaborators, including site PIs and technicians, for their efforts
530 in support of PhenoCam. The development of PhenoCam has been funded by the Northeastern States Research Cooperative, NSF's Macrosystems Biology program (awards EF-1065029 and EF-1702697), and DOE's Regional and Global Climate Modeling program (award DE-SC0016011). We acknowledge additional support from the US National Park Service Inventory and Monitoring Program and the USA National Phenology Network (grant number G10AP00129 from the United States Geological Survey), and from the USA National Phenology Network and North Central Climate Science Center (cooperative
535 agreement number G16AC00224 from the United States Geological Survey), and from the Long-Term Agroecosystem

Research (LTAR) network which is supported by the United States Department of Agriculture (USDA) (cooperative agreement 59-3050-2-002). Additional funding, through the National Science Foundation's LTER program, has supported research at Harvard Forest (DEB-1237491) and Bartlett Experimental Forest (DEB-1114804). We also thank the USDA Forest Service Air Resource Management program and the National Park Service Air Resources program for contributing their camera imagery to the PhenoCam archive.

References

Adole, T., Dash, J., Rodriguez-Galiano, V., and Atkinson, P. M.: Photoperiod controls vegetation phenology across Africa, Commun Biol, 2, 1–13, https://doi.org/10.1038/s42003-019-0636-7, 2019.

Alduchov, O. A. and Eskridge, R. E.: Improved magnus form approximation of saturation vapor pressure, Journal of applied meteorology, 35, 601–609, https://doi.org/10.2172/548871, 1997.

Asse, D., Randin, C. F., Bonhomme, M., Delestrade, A., and Chuine, I.: Process-based models outcompete correlative models in projecting spring phenology of trees in a future warmer climate, Agricultural and Forest Meteorology, 285–286, 107931, https://doi.org/10.1016/j.agrformet.2020.107931, 2020.

Bahdanau, D., Cho, K., and Bengio, Y.: Neural Machine Translation by Jointly Learning to Align and Translate, arXiv, https://doi.org/10.48550/arXiv.1409.0473, 2016.

Besnard, S., Carvalhais, N., Arain, M. A., Black, A., Brede, B., Buchmann, N., Chen, J., Clevers, J. G. P. W., Dutrieux, L. P., Gans, F., Herold, M., Jung, M., Kosugi, Y., Knohl, A., Law, B. E., Paul-Limoges, E., Lohila, A., Merbold, L., Roupsard, O., Valentini, R., Wolf, S., Zhang, X., and Reichstein, M.: Memory effects of climate and vegetation affecting net ecosystem CO2 fluxes in global forests, PLOS ONE, 14, e0211510, https://doi.org/10.1371/journal.pone.0211510, 2019.

Bonan, G.: Climate Change and Terrestrial Ecosystem Modeling, 1st ed., Cambridge University Press, https://doi.org/10.1017/9781107339217, 2019.

Borchert, R., Robertson, K., Schwartz, M. D., and Williams-Linera, G.: Phenology of temperate trees in tropical climates, Int J Biometeorol, 50, 57–65, https://doi.org/10.1007/s00484-005-0261-7, 2005.

Buermann, W., Forkel, M., O'Sullivan, M., Sitch, S., Friedlingstein, P., Haverd, V., Jain, A. K., Kato, E., Kautz, M., Lienert, S., Lombardozzi, D., Nabel, J. E. M. S., Tian, H., Wiltshire, A. J., Zhu, D., Smith, W. K., and Richardson, A. D.: Widespread seasonal compensation effects of spring warming on northern plant productivity, Nature, 562, 110–114, https://doi.org/10.1038/s41586-018-0555-7, 2018.

Callaghan, M., Schleussner, C.-F., Nath, S., Lejeune, Q., Knutson, T. R., Reichstein, M., Hansen, G., Theokritoff, E., Andrijevic, M., Brecha, R. J., Hegarty, M., Jones, C., Lee, K., Lucas, A., van Maanen, N., Menke, I., Pfleiderer, P., Yesil, B., and Minx, J. C.: Machine-learning-based evidence and attribution mapping of 100,000 climate impact studies, Nat. Clim. Chang., 11, 966–972, https://doi.org/10.1038/s41558-021-01168-6, 2021.

Chen, X. and Xu, L.: Temperature controls on the spatial pattern of tree phenology in China's temperate zone, Agricultural and Forest Meteorology, 154–155, 195–202, https://doi.org/10.1016/j.agrformet.2011.11.006, 2012.

Chen, Z., Liu, H., Xu, C., Wu, X., Liang, B., Cao, J., and Chen, D.: Modeling vegetation greenness and its climate sensitivity with deep-learning technology, Ecology and Evolution, 11, 7335–7345, https://doi.org/10.1002/ece3.7564, 2021.

Chuine, I.: A Unified Model for Budburst of Trees, Journal of Theoretical Biology, 207, 337–347, https://doi.org/10.1006/jtbi.2000.2178, 2000.

Chuine, I., Morin, X., and Bugmann, H.: Warming, Photoperiods, and Tree Phenology, Science, 329, 277–278, https://doi.org/10.1126/science.329.5989.277-e, 2010.

Chuine, I., Bonhomme, M., Legave, J.-M., García de Cortázar-Atauri, I., Charrier, G., Lacointe, A., and Améglio, T.: Can phenological models predict tree phenology accurately in the future? The unrevealed hurdle of endodormancy break, Global Change Biology, 22, 3444–3460, https://doi.org/10.1111/gcb.13383, 2016.

Cleland, E. E., Chuine, I., Menzel, A., Mooney, H. A., and Schwartz, M. D.: Shifting plant phenology in response to global change, Trends in Ecology & Evolution, 22, 357–365, https://doi.org/10.1016/j.tree.2007.04.003, 2007.

Cleveland, W. S.: Robust Locally Weighted Regression and Smoothing Scatterplots, Journal of the American Statistical Association, 74, 829–836, https://doi.org/10.1080/01621459.1979.10481038, 1979.

Delpierre, N., Dufrêne, E., Soudani, K., Ulrich, E., Cecchini, S., Boé, J., and François, C.: Modelling interannual and spatial variability of leaf senescence for three deciduous tree species in France, Agricultural and Forest Meteorology, 149, 938–948, https://doi.org/10.1016/j.agrformet.2008.11.014, 2009.

Denéchère, R., Delpierre, N., Apostol, E. N., Berveiller, D., Bonne, F., Cole, E., Delzon, S., Dufrêne, E., Gressler, E., Jean, F., Lebourgeois, F., Liu, G., Louvet, J.-M., Parmentier, J., Soudani, K., and Vincent, G.: The within-population variability of leaf spring and autumn phenology is influenced by temperature in temperate deciduous trees, Int J Biometeorol, 65, 369–379, https://doi.org/10.1007/s00484-019-01762-6, 2021.

Diffenbaugh, N. S. and Barnes, E. A.: Data-driven predictions of the time remaining until critical global warming thresholds are reached, Proceedings of the National Academy of Sciences, 120, e2207183120, https://doi.org/10.1073/pnas.2207183120, 2023.

Ding, Y., Li, Z., and Peng, S.: Global analysis of time-lag and -accumulation effects of climate on vegetation growth, International Journal of Applied Earth Observation and Geoinformation, 92, 102179, https://doi.org/10.1016/j.jag.2020.102179, 2020.

ElGhawi, R., Kraft, B., Reimers, C., Reichstein, M., Körner, M., Gentine, P., and Winkler, A. J.: Hybrid modeling of evapotranspiration: inferring stomatal and aerodynamic resistances using combined physics-based and machine learning, Environ. Res. Lett., 18, 034039, https://doi.org/10.1088/1748-9326/acbbe0, 2023.

Ettinger, A. K., Gee, S., and Wolkovich, E. M.: Phenological sequences: how early-season events define those that follow, American Journal of Botany, 105, 1771–1780, https://doi.org/10.1002/ajb2.1174, 2018.

Ettinger, A. K., Chamberlain, C. J., Morales-Castilla, I., Buonaiuto, D. M., Flynn, D. F. B., Savas, T., Samaha, J. A., and Wolkovich, E. M.: Winter temperatures predominate in spring phenological responses to warming, Nat. Clim. Chang., 10, 1137–1142, https://doi.org/10.1038/s41558-020-00917-3, 2020.

Flynn, D. F. B. and Wolkovich, E. M.: Temperature and photoperiod drive spring phenology across all species in a temperate forest community, New Phytologist, 219, 1353–1362, https://doi.org/10.1111/nph.15232, 2018.

605   Forkel, M., Dorigo, W., Lasslop, G., Teubner, I., Chuvieco, E., and Thonicke, K.: A data-driven approach to identify controls on global fire activity from satellite and climate observations (SOFIA V1), Geoscientific Model Development, 10, 4443–4476, https://doi.org/10.5194/gmd-10-4443-2017, 2017.

Fu, Y., Li, X., Zhou, X., Geng, X., Guo, Y., and Zhang, Y.: Progress in plant phenology modeling under global climate change, Sci. China Earth Sci., 63, 1237–1247, https://doi.org/10.1007/s11430-019-9622-2, 2020.

610   Fu, Y. H., Zhao, H., Piao, S., Peaucelle, M., Peng, S., Zhou, G., Ciais, P., Huang, M., Menzel, A., Peñuelas, J., Song, Y., Vitasse, Y., Zeng, Z., and Janssens, I. A.: Declining global warming effects on the phenology of spring leaf unfolding, Nature, 526, 104–107, https://doi.org/10.1038/nature15402, 2015.

Fu, Y. S. H., Campioli, M., Vitasse, Y., De Boeck, H. J., Van den Berge, J., AbdElgawad, H., Asard, H., Piao, S., Deckmyn, G., and Janssens, I. A.: Variation in leaf flushing date influences autumnal senescence and next year's flushing date in two
615   temperate   tree   species,   Proceedings   of   the   National   Academy   of   Sciences,   111,   7355–7360, https://doi.org/10.1073/pnas.1321727111, 2014.

Gömöry, D., Foffová, E., Longauer, R., and Krajmerová, D.: Memory effects associated with early-growth environment in Norway spruce and European larch, Eur J Forest Res, 134, 89–97, https://doi.org/10.1007/s10342-014-0835-1, 2015.

Hall, C. A. and Meyer, W. W.: Optimal error bounds for cubic spline interpolation, Journal of Approximation Theory, 16,
620   105–122, https://doi.org/10.1016/0021-9045(76)90040-X, 1976.

Hänninen, H.: Modelling bud dormancy release in trees from cool and temperate regions. (Acta forestalia Fennica). Society of Forestry in Finland, 1990.

Hochreiter,   S.   and   Schmidhuber,   J.:   Long   Short-Term   Memory,   Neural   Computation,   9,   1735–1780, https://doi.org/10.1162/neco.1997.9.8.1735, 1997.

625   Hornik, K., Stinchcombe, M., and White, H.: Multilayer feedforward networks are universal approximators, Neural Networks, 2, 359–366, https://doi.org/10.1016/0893-6080(89)90020-8, 1989.

Hua, Y., Zhao, Z., Li, R., Chen, X., Liu, Z., and Zhang, H.: Deep Learning with Long Short-Term Memory for Time Series Prediction, IEEE Communicatio0ns Magazine, 57, 114–119, https://doi.org/10.1109/MCOM.2019.1800155, 2019.

Jeong, S.-J., Medvigy, D., Shevliakova, E., and Malyshev, S.: Uncertainties in terrestrial carbon budgets related to spring
630   phenology, Journal of Geophysical Research: Biogeosciences, 117, https://doi.org/10.1029/2011JG001868, 2012.

Jeong, S.: Autumn greening in a warming climate, Nat. Clim. Chang., 10, 712–713, https://doi.org/10.1038/s41558-020-0852-7, 2020.

Jin, J., Wang, Y., Zhang, Z., Magliulo, V., Jiang, H., and Cheng, M.: Phenology Plays an Important Role in the Regulation of Terrestrial   Ecosystem   Water-Use   Efficiency   in   the   Northern   Hemisphere,   Remote   Sensing,   9,   664,
635   https://doi.org/10.3390/rs9070664, 2017.

Jolly, W. M., Nemani, R., and Running, S. W.: A generalized, bioclimatic index to predict foliar phenology in response to climate, Global Change Biology, 11, 619–632, https://doi.org/10.1111/j.1365-2486.2005.00930.x, 2005.

Joshi, R. C., Sheridan, G. J., Ryu, D., and Lane, P. N. J.: How long is the memory of forest growth to rainfall in asynchronous climates?, Ecological Indicators, 140, 109057, https://doi.org/10.1016/j.ecolind.2022.109057, 2022.

640   Kingma, D. P. and Ba, J.: Adam: A Method for Stochastic Optimization, arXiv, https://doi.org/10.48550/arXiv.1412.6980, 2017.

Koebsch, F., Sonnentag, O., Järveoja, J., Peltoniemi, M., Alekseychik, P., Aurela, M., Arslan, A. N., Dinsmore, K., Gianelle, D., Helfter, C., Jackowicz-Korczynski, M., Korrensalo, A., Leith, F., Linkosalmi, M., Lohila, A., Lund, M., Maddison, M., Mammarella, I., Mander, Ü., Minkkinen, K., Pickard, A., Pullens, J. W. M., Tuittila, E.-S., Nilsson, M. B., and Peichl, M.:
645   Refining the role of phenology in regulating gross ecosystem productivity across European peatlands, Global Change Biology, 26, 876–887, https://doi.org/10.1111/gcb.14905, 2020.

Kraft, B., Jung, M., Körner, M., Requena Mesa, C., Cortés, J., and Reichstein, M.: Identifying Dynamic Memory Effects on Vegetation State Using Recurrent Neural Networks, Frontiers in Big Data, 2, 2019.

Kraft, B., Jung, M., Körner, M., Koirala, S., and Reichstein, M.: Towards hybrid modeling of the global hydrological cycle,
650   Hydrology and Earth System Sciences, 26, 1579–1614, https://doi.org/10.5194/hess-26-1579-2022, 2022.

Krinner, G., 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 Biogeochemical Cycles, 19, 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
655   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,
660   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.

Lian, X., Piao, S., Chen, A., Wang, K., Li, X., Buermann, W., Huntingford, C., Peñuelas, J., Xu, H., and Myneni, R. B.: Seasonal biological carryover dominates northern vegetation growth, Nat Commun, 12, 983, https://doi.org/10.1038/s41467-021-21223-2, 2021.

665   Liu, G., Chen, X., Zhang, Q., Lang, W., and Delpierre, N.: Antagonistic effects of growing season and autumn temperatures on the timing of leaf coloration in winter deciduous trees, Global Change Biology, 24, 3537–3545, https://doi.org/10.1111/gcb.14095, 2018.

Liu, L., Zhang, Y., Wu, S., Li, S., and Qin, D.: Water memory effects and their impacts on global vegetation productivity and resilience, Sci Rep, 8, 2962, https://doi.org/10.1038/s41598-018-21339-4, 2018.

670   Liu, G., Chen, X., Fu, Y., and Delpierre, N.: Modelling leaf coloration dates over temperate China by considering effects of leafy season climate, Ecological Modelling, 394, 34–43, https://doi.org/10.1016/j.ecolmodel.2018.12.020, 2019.

Loshchilov, I. and Hutter, F.: Decoupled Weight Decay Regularization, arXiv, https://doi.org/10.48550/arXiv.1711.05101, 2019.

Luo, Y., El-Madany, T., Ma, X., Nair, R., Jung, M., Weber, U., Filippa, G., Bucher, S. F., Moreno, G., Cremonese, E., Carrara,
675   A., Gonzalez-Cascon, R., Cáceres Escudero, Y., Galvagno, M., Pacheco-Labrador, J., Martín, M. P., Perez-Priego, O., Reichstein, M., Richardson, A. D., Menzel, A., Römermann, C., and Migliavacca, M.: Nutrients and water availability constrain the seasonality of vegetation activity in a Mediterranean ecosystem, Global Change Biology, 26, 4379–4400, https://doi.org/10.1111/gcb.15138, 2020.

Mauritsen, T., Bader, J., Becker, T., Behrens, J., Bittner, M., Brokopf, R., Brovkin, V., Claussen, M., Crueger, T., Esch, M., Fast, I., Fiedler, S., Fläschner, D., Gayler, V., Giorgetta, M., Goll, D. S., Haak, H., Hagemann, S., Hedemann, C., Hohenegger, C., Ilyina, T., Jahns, T., Jimenéz-de-la-Cuesta, D., Jungclaus, J., Kleinen, T., Kloster, S., Kracher, D., Kinne, S., Kleberg, D., Lasslop, G., Kornblueh, L., Marotzke, J., Matei, D., Meraner, K., Mikolajewicz, U., Modali, K., Möbis, B., Müller, W. A., Nabel, J. E. M. S., Nam, C. C. W., Notz, D., Nyawira, S.-S., Paulsen, H., Peters, K., Pincus, R., Pohlmann, H., Pongratz, J., Popp, M., Raddatz, T. J., Rast, S., Redler, R., Reick, C. H., Rohrschneider, T., Schemann, V., Schmidt, H., Schnur, R., Schulzweida, U., Six, K. D., Stein, L., Stemmler, I., Stevens, B., von Storch, J.-S., Tian, F., Voigt, A., Vrese, P., Wieners, K.-H., Wilkenskjeld, S., Winkler, A., and Roeckner, E.: Developments in the MPI-M Earth System Model version 1.2 (MPI-ESM1.2) and Its Response to Increasing CO2, Journal of Advances in Modeling Earth Systems, 11, 998–1038, https://doi.org/10.1029/2018MS001400, 2019.

Menzel, A., Sparks, T. H., Estrella, N., Koch, E., Aasa, A., Ahas, R., Alm-Kübler, K., Bissolli, P., Braslavská, O., Briede, A., Chmielewski, F. M., Crepinsek, Z., Curnel, Y., Dahl, Å., Defila, C., Donnelly, A., Filella, Y., Jatczak, K., Måge, F., Mestre, A., Nordli, Ø., Peñuelas, J., Pirinen, P., Remišová, V., Scheifinger, H., Striz, M., Susnik, A., Van Vliet, A. J. H., Wielgolaski, F.-E., Zach, S., and Zust, A.: European phenological response to climate change matches the warming pattern, Global Change Biology, 12, 1969–1976, https://doi.org/10.1111/j.1365-2486.2006.01193.x, 2006.

Menzel, A., Seifert, H., and Estrella, N.: Effects of recent warm and cold spells on European plant phenology, Int J Biometeorol, 55, 921–932, https://doi.org/10.1007/s00484-011-0466-x, 2011.

Migliavacca, M., Reichstein, M., Richardson, A. D., Colombo, R., Sutton, M. A., Lasslop, G., Tomelleri, E., Wohlfahrt, G., Carvalhais, N., Cescatti, A., Mahecha, M. D., Montagnani, L., Papale, D., Zaehle, S., Arain, A., Arneth, A., Black, T. A., Carrara, A., Dore, S., Gianelle, D., Helfter, C., Hollinger, D., Kutsch, W. L., Lafleur, P. M., Nouvellon, Y., Rebmann, C., Da ROCHA, H. R., Rodeghiero, M., Roupsard, O., Sebastià, M.-T., Seufert, G., Soussana, J.-F., and Van Der MOLEN, M. K.: Semiempirical modeling of abiotic and biotic factors controlling ecosystem respiration across eddy covariance sites, Global Change Biology, 17, 390–409, https://doi.org/10.1111/j.1365-2486.2010.02243.x, 2011.

Migliavacca, M., Sonnentag, O., Keenan, T. F., Cescatti, A., O'Keefe, J., and Richardson, A. D.: On the uncertainty of phenological responses to climate change, and implications for a terrestrial biosphere model, Biogeosciences, 9, 2063–2083, https://doi.org/10.5194/bg-9-2063-2012, 2012.

Migliavacca, M., Reichstein, M., Richardson, A. D., Mahecha, M. D., Cremonese, E., Delpierre, N., Galvagno, M., Law, B. E., Wohlfahrt, G., Andrew Black, T., Carvalhais, N., Ceccherini, G., Chen, J., Gobron, N., Koffi, E., William Munger, J., Perez-Priego, O., Robustelli, M., Tomelleri, E., and Cescatti, A.: Influence of physiological phenology on the seasonal pattern of ecosystem respiration in deciduous forests, Global Change Biology, 21, 363–376, https://doi.org/10.1111/gcb.12671, 2015.

Murray-Tortarolo, G., Anav, A., Friedlingstein, P., Sitch, S., Piao, S., Zhu, Z., Poulter, B., Zaehle, S., Ahlström, A., Lomas, M., Levis, S., Viovy, N., and Zeng, N.: 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, 2013.

O, S., Dutra, E., and Orth, R.: Robustness of Process-Based versus Data-Driven Modeling in Changing Climatic Conditions, Journal of Hydrometeorology, 21, 1929–1944, https://doi.org/10.1175/JHM-D-20-0072.1, 2020.

Ogle, K., Barber, J. J., Barron-Gafford, G. A., Bentley, L. P., Young, J. M., Huxman, T. E., Loik, M. E., and Tissue, D. T.: Quantifying ecological memory in plant and ecosystem processes, Ecology Letters, 18, 221–235, https://doi.org/10.1111/ele.12399, 2015.

Peano, D., Hemming, D., Materia, 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.

Peng, J., Wu, C., Zhang, X., Wang, X., and Gonsamo, A.: Satellite detection of cumulative and lagged effects of drought on autumn leaf senescence over the Northern Hemisphere, Global Change Biology, 25, 2174–2188, https://doi.org/10.1111/gcb.14627, 2019.

Peñuelas, J., Rutishauser, T., and Filella, I.: Phenology Feedbacks on Climate Change, Science, 324, 887–888, https://doi.org/10.1126/science.1173004, 2009.

Piao, S., Liu, Q., Chen, A., Janssens, I. A., Fu, Y., Dai, J., Liu, L., Lian, X., Shen, M., and Zhu, X.: Plant phenology and global climate change: Current progresses and challenges, Global Change Biology, 25, 1922–1940, https://doi.org/10.1111/gcb.14619, 2019.

Pichler, M. and Hartig, F.: Machine learning and deep learning—A review for ecologists, Methods in Ecology and Evolution, 14, 994–1016, https://doi.org/10.1111/2041-210X.14061, 2023.

Pollard, C. P., Griffin, C. T., Andrade Moral, R. de, Duffy, C., Chuche, J., Gaffney, M. T., Fealy, R. M., and Fealy, R.: phenModel: A temperature-dependent phenology/voltinism model for a herbivorous insect incorporating facultative diapause and budburst, Ecological Modelling, 416, 108910, https://doi.org/10.1016/j.ecolmodel.2019.108910, 2020.

Puma, M. J., Koster, R. D., and Cook, B. I.: Phenological versus meteorological controls on land-atmosphere water and carbon fluxes, Journal of Geophysical Research: Biogeosciences, 118, 14–29, https://doi.org/10.1029/2012JG002088, 2013.

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., and Prabhat: Deep learning and process understanding for data-driven Earth system science, Nature, 566, 195–204, https://doi.org/10.1038/s41586-019-0912-1, 2019.

Ren, P., Liang, E., Raymond, P., and Rossi, S.: Bud break in sugar maple submitted to changing conditions simulating a northward migration, Can. J. For. Res., 51, 842–847, https://doi.org/10.1139/cjfr-2020-0365, 2021.

Richardson, A. D., Hollinger, D. Y., Dail, D. B., Lee, J. T., Munger, J. W., and O'keefe, J.: Influence of spring phenology on seasonal and annual carbon balance in two contrasting New England forests, Tree Physiol, 29, 321–331, https://doi.org/10.1093/treephys/tpn040, 2009.

Richardson, A. D., Andy Black, T., Ciais, P., Delbart, N., Friedl, M. A., Gobron, N., Hollinger, D. Y., Kutsch, W. L., Longdoz, B., Luyssaert, S., Migliavacca, M., Montagnani, L., William Munger, J., Moors, E., Piao, S., Rebmann, C., Reichstein, M., Saigusa, N., Tomelleri, E., Vargas, R., and Varlagin, A.: Influence of spring and autumn phenological transitions on forest ecosystem productivity, Philosophical Transactions of the Royal Society B: Biological Sciences, 365, 3227–3246, https://doi.org/10.1098/rstb.2010.0102, 2010.

Richardson, A. D., Anderson, R. S., Arain, M. A., Barr, A. G., Bohrer, G., Chen, G., Chen, J. M., Ciais, P., Davis, K. J., Desai, A. R., Dietze, M. C., Dragoni, D., Garrity, S. R., Gough, C. M., Grant, R., Hollinger, D. Y., Margolis, H. A., McCaughey, H., Migliavacca, M., Monson, R. K., Munger, J. W., Poulter, B., Raczka, B. M., Ricciuto, D. M., Sahoo, A. K., Schaefer, K., Tian, H., Vargas, R., Verbeeck, H., Xiao, J., and Xue, Y.: Terrestrial biosphere models need better representation of vegetation phenology: results from the North American Carbon Program Site Synthesis, Global Change Biology, 18, 566–584, https://doi.org/10.1111/j.1365-2486.2011.02562.x, 2012.

Richardson, A. D., Keenan, T. F., Migliavacca, M., Ryu, Y., Sonnentag, O., and Toomey, M.: Climate change, phenology, and phenological control of vegetation feedbacks to the climate system, Agricultural and Forest Meteorology, 169, 156–173, https://doi.org/10.1016/j.agrformet.2012.09.012, 2013.

Richardson, A. D., Hufkens, K., Milliman, T., Aubrecht, D. M., Chen, M., Gray, J. M., Johnston, M. R., Keenan, T. F., Klosterman, S. T., Kosmala, M., Melaas, E. K., Friedl, M. A., and Frolking, S.: Tracking vegetation phenology across diverse North American biomes using PhenoCam imagery, Sci Data, 5, 1–24, https://doi.org/10.1038/sdata.2018.28, 2018.

Rodriguez-Galiano, V. F., Luque-Espinar, J. A., Chica-Olmo, M., and Mendes, M. P.: Feature selection approaches for predictive modelling of groundwater nitrate pollution: An evaluation of filters, embedded and wrapper methods, Science of The Total Environment, 624, 661–672, https://doi.org/10.1016/j.scitotenv.2017.12.152, 2018.

Seyednasrollah, B., Young, A. M., Hufkens, K., Milliman, T., Friedl, M. A., Frolking, S., and Richardson, A. D.: Tracking vegetation phenology across diverse biomes using Version 2.0 of the PhenoCam Dataset, Sci Data, 6, 222, https://doi.org/10.1038/s41597-019-0229-9, 2019.

Song, G., Wu, S., Lee, C. K. F., Serbin, S. P., Wolfe, B. T., Ng, M. K., Ely, K. S., Bogonovich, M., Wang, J., Lin, Z., Saleska, S., Nelson, B. W., Rogers, A., and Wu, J.: Monitoring leaf phenology in moist tropical forests by applying a superpixel-based deep learning method to time-series images of tree canopies, ISPRS Journal of Photogrammetry and Remote Sensing, 183, 19–33, https://doi.org/10.1016/j.isprsjprs.2021.10.023, 2022.

Sparks, T. H. and Carey, P. D.: The Responses of Species to Climate Over Two Centuries: An Analysis of the Marsham Phenological Record, 1736-1947, Journal of Ecology, 83, 321–329, https://doi.org/10.2307/2261570, 1995.

Sutskever, I., Vinyals, O., and Le, Q. V.: Sequence to Sequence Learning with Neural Networks, arXiv, https://doi.org/10.48550/arXiv.1409.3215, 2014.

Thornton, P. E. and Rosenbloom, N. A.: Ecosystem model spin-up: Estimating steady state conditions in a coupled terrestrial carbon and nitrogen cycle model, Ecological Modelling, 189, 25–48, https://doi.org/10.1016/j.ecolmodel.2005.04.008, 2005.

Thornton, P. E., Law, B. E., Gholz, H. L., Clark, K. L., Falge, E., Ellsworth, D. S., Goldstein, A. H., Monson, R. K., Hollinger, D., Falk, M., Chen, J., and Sparks, J. P.: Modeling and measuring the effects of disturbance history and climate on carbon and water budgets in evergreen needleleaf forests, Agricultural and Forest Meteorology, 113, 185–222, https://doi.org/10.1016/S0168-1923(02)00108-9, 2002.

Walter, J., Nagy, L., Hein, R., Rascher, U., Beierkuhnlein, C., Willner, E., and Jentsch, A.: Do plants remember drought? Hints towards a drought-memory in grasses, Environmental and Experimental Botany, 71, 34–40, https://doi.org/10.1016/j.envexpbot.2010.10.020, 2011.

White, M. A., Thornton, P. E., and Running, S. W.: A continental phenology model for monitoring vegetation responses to interannual climatic variability, Global Biogeochemical Cycles, 11, 217–234, https://doi.org/10.1029/97GB00330, 1997.

Wingate, L., Ogée, J., Cremonese, E., Filippa, G., Mizunuma, T., Migliavacca, M., Moisy, C., Wilkinson, M., Moureaux, C., Wohlfahrt, G., Hammerle, A., Hörtnagl, L., Gimeno, C., Porcar-Castell, A., Galvagno, M., Nakaji, T., Morison, J., Kolle, O., Knohl, A., Kutsch, W., Kolari, P., Nikinmaa, E., Ibrom, A., Gielen, B., Eugster, W., Balzarolo, M., Papale, D., Klumpp, K., Köstner, B., Grünwald, T., Joffre, R., Ourcival, J.-M., Hellstrom, M., Lindroth, A., George, C., Longdoz, B., Genty, B., Levula, J., Heinesch, B., Sprintsin, M., Yakir, D., Manise, T., Guyon, D., Ahrends, H., Plaza-Aguilar, A., Guan, J. H., and Grace, J.: Interpreting canopy development and physiology using a European phenology camera network at flux sites, Biogeosciences, 12, 5995–6015, https://doi.org/10.5194/bg-12-5995-2015, 2015.

Wolkovich, E. M., Cook, B. I., Allen, J. M., Crimmins, T. M., Betancourt, J. L., Travers, S. E., Pau, S., Regetz, J., Davies, T. J., Kraft, N. J. B., Ault, T. R., Bolmgren, K., Mazer, S. J., McCabe, G. J., McGill, B. J., Parmesan, C., Salamin, N., Schwartz, M. D., and Cleland, E. E.: Warming experiments underpredict plant phenological responses to climate change, Nature, 485, 494–497, https://doi.org/10.1038/nature11014, 2012.

Wu, C.-Y., Ahmed, A., Beutel, A., Smola, A. J., and Jing, H.: Recurrent Recommender Networks, in: Proceedings of the Tenth ACM International Conference on Web Search and Data Mining, New York, NY, USA, 495–503, https://doi.org/10.1145/3018661.3018689, 2017.

Wu, J., Wang, D., Li, L. Z. X., and Zeng, Z.: Hydrological feedback from projected Earth greening in the 21st century, Sustainable Horizons, 1, 100007, https://doi.org/10.1016/j.horiz.2022.100007, 2022.

Zhang, H., Chuine, I., Regnier, P., Ciais, P., and Yuan, W.: Deciphering the multiple effects of climate warming on the temporal shift of leaf unfolding, Nat. Clim. Chang., 12, 193–199, https://doi.org/10.1038/s41558-021-01261-w, 2022.

Zhao, Z., Chen, W., Wu, X., Chen, P. C. Y., and Liu, J.: LSTM network: a deep learning approach for short-term traffic forecast, IET Intelligent Transport Systems, 11, 68–75, https://doi.org/10.1049/iet-its.2016.0208, 2017.

Zhou, R., Liu, Y., Cui, M., Lu, J., Shi, H., Ren, H., Zhang, W., and Wen, Z.: Global Assessment of Cumulative and Time-Lag Effects of Drought on Land Surface Phenology, GIScience & Remote Sensing, 59, 1918–1937, https://doi.org/10.1080/15481603.2022.2143661, 2022.

Zhou, X., Xin, Q., Dai, Y., and Li, W.: A deep-learning-based experiment for benchmarking the performance of global terrestrial vegetation phenology models, Global Ecology and Biogeography, 30, 2178–2199, https://doi.org/10.1111/geb.13374, 2021.