

Predicting Spring Green-up Across Diverse North American Grasslands

Alison K. Post^{a*}, Koen Hufkens^b, Andrew D. Richardson^a

^aCenter for Ecosystem Science and Society and School of Informatics, Computing, and Cyber Systems, Northern Arizona University, Flagstaff, AZ 86011, USA

^bBlueGreen Labs, Melsele, Belgium

Correspondence to Alison Post:
alison.post@nau.edu

*ORCID iD: 0000-0003-2931-6490

Running title: Predicting grassland spring green-up

Abstract

Vegetation phenology influences many ecosystem and climate processes, such as carbon uptake and energy and water cycles. Thus, understanding drivers of vegetation phenology is crucial for predicting current and future impacts of climate change on ecological systems. Existing models can accurately predict the date of spring green-up in temperate forests but tend to perform poorly in grassland systems. We hypothesize this is because most do not incorporate water availability, a primary limiting factor for grassland plants. In this study, we used long-term datasets of digital imagery from the PhenoCam Network of 43 diverse North American grassland sites (195 site-years) to test existing spring phenology models, as well as develop several new models that incorporate precipitation or soil moisture (53 models). Most of the new models performed substantially better, with the best model requiring sufficient accumulated precipitation followed by warm temperatures to trigger spring onset (root mean square error, RMSE, between predicted and observed dates = 16.0 days). Importantly, the best model performed well across all grassland types using a single set of parameters, from temperate to arid grasslands. Since plants are adapted to their local climates, model performance was further improved when parameters were independently optimized for four separate climate regions (RMSE = 10.4 days). Therefore, both sufficient precipitation and temperature are required for grassland green-up, but optimal thresholds vary by region. Running the top model with projected climate data (representative concentration pathway 8.5) suggests that, depending on the climate region, spring onset will occur up to 12 days earlier within 100 years in temperature-limited sites, but the trend is unclear for precipitation-limited sites (3.5 ± 8.0 days later). This new phenology model improves our ability to understand and predict grassland dynamics, with implications for both current and future ecosystem processes related to carbon and water cycling.

Keywords: climate change, grasslands, model, precipitation, phenology, spring

1. Introduction

Phenology studies the timing of recurring biological events and their interactions with abiotic and biotic factors (Lieth, 1974). For instance, vegetation seasonality, such as the timing and length of the growing season, is influenced by climate patterns, but at the same time, also heavily influences many climate processes, such as water, energy, and carbon dioxide fluxes, as well as biological interactions (Richardson et al., 2013; Caparros-Santiago et al., 2021). In this way, vegetation phenology is a robust indicator of the influence of climate change on biological systems. Thus, understanding the drivers of vegetation phenology is crucial for predicting the current and future impacts of climate change on ecological systems.

There have been numerous attempts to model vegetation phenology using standard climate variables, represented by two main approaches. The first method attempts to predict vegetation dynamics over the entire growing season (i.e., Jolly et al., 2005; Choler et al., 2010; Koen et al., 2016, Tian et al., 2019), while the other method predicts seasonal transition dates, such as the start and end of the growing season. For this study, we focused on this second modeling approach, with the goal of determining environmental triggers that initiate the start of the growing season (SOS). These models tend to be simpler, and thus, can be more easily incorporated into larger earth system models. For example, the Community Land Model (CLM) uses pre-determined thresholds of key environmental variables to initiate spring leaf-out (CLM5 Documentation).

The oldest and simplest spring onset model is the Growing Degree Day (GDD) model (also known as the Thermal Time (TT) model), which states that SOS occurs when enough heat

(degree-days) above a given base temperature has accumulated (Reaumur, 1735). More recent models have added requirements for chilling and photoperiod (daylength) to the GDD model in various configurations (Basler, 2016). These models can accurately predict the date of spring onset in temperate forests (Fu et al., 2012; Basler, 2016), but tend to perform poorly in drier ecosystems, such as grasslands (Fu et al., 2014; Xin et al., 2015; Cao et al., 2018). For example, Liu et al. (2018) used satellite imagery data to test the ability of six common spring phenology models to predict SOS across the Northern Hemisphere and found they worked relatively well for forested ecosystems, but not grasslands. This is likely because traditional forest models do not include water availability. Unlike most forest ecosystems, grasslands are inherently water-limited, so spring green-up depends, not only on temperature, but also on sufficient precipitation inputs (Moore et al., 2015; Zhu et al., 2015). Given that grasslands account for 30% of the Earth's ice-free land surface and are significant carbon sinks (Gao et al., 2016), understanding the drivers of grassland phenology represents a critical research priority.

In an attempt to better capture grassland dynamics, several spring phenology models have been developed that include a metric of water availability. For example, White et al. (1997) developed a grassland model that, in addition to temperature, requires a portion of mean annual precipitation for SOS to occur. Jolly et al. (2005) created the Growing Season Index model that uses minimum temperature, photoperiod, and vapor pressure deficit (VPD) to predict SOS. Xin et al. (2015) later adapted this model to create the Accumulated Growing Season Index (AGSI) model specifically for grasslands, as well as a variation in which they replaced VPD with soil moisture. Similarly, Chen et al. (2014) added precipitation to the traditional GDD model both in sequence (precipitation threshold must be met before temperature accumulation starts) and in parallel (precipitation and temperature accumulate at the same time). In all cases, the addition of

water availability significantly improved model performance in predicting SOS in grassland ecosystems. However, despite these advances, most large earth system models, such as CLM, are still unable to accurately capture grassland phenological transitions (Zhang et al., 2019; Li et al., 2022), resulting in large uncertainties in estimating critical ecosystem processes and fluxes.

The availability of high-resolution, long-term phenology datasets is a limitation for modeling grassland dynamics. Many existing models have only been tested with either spatially limited field data or with coarse resolution satellite data. Remote sensing products have inherent limitations, including coarse temporal and spatial resolution, resulting in multiple plant functional groups being included in a single pixel, thereby causing uncertainty in seasonal trends (Chen et al., 2018; Cui et al., 2019; Taylor et al., 2021). In addition, satellite data in locations with sparse vegetation cover, such as drylands, have high background noise due to bare soil reflectance. This results in low-amplitude seasonal signals (Taylor et al., 2021), so many remote sensing studies exclude dryland ecosystems, such as arid grasslands, from their analyses (e.g., Jeong et al., 2011; Xin et al., 2015; Ren et al., 2018, 2022; Liu et al., 2020). Likewise, Mediterranean annual grasslands (such as California grasslands) are often excluded due to inconsistent annual green-up patterns (Xin et al., 2015; Ren et al., 2018). Thus, the broader applicability of these models across diverse grassland types is still unknown.

Repeat digital photography provides an alternative data source for assessing the viability of these models. PhenoCams are cameras mounted above ecosystem canopies that capture digital images at high temporal resolution (usually every 30 minutes). There are currently over 700 cameras located within diverse ecosystems around the world, creating the PhenoCam Network (phenocam.nau.edu), which provides powerful, high resolution, long-term phenology data. By using the red-green-blue (RGB) color channels of the digital images, a robust measure of

vegetation greenness can be extracted, called the Green Chromatic Coordinate (GCC), which is highly correlated with other common vegetation indices, such as the Normalized Difference Vegetation Index (NDVI) and the Enhanced Vegetation Index (EVI), as well as Gross Primary Production (GPP) (Migliavacca et al., 2011; Klosterman et al., 2014; Browning et al., 2017; Cui et al., 2019). This is an effective method for tracking seasonal changes in vegetation phenology (Richardson et al., 2013). Additionally, PhenoCam imagery is unique in that a designated portion of the image, referred to as the region of interest (ROI), can be identified for analysis. Thus, GCC can be extracted separately for distinct functional groups in an image and, by only focusing on vegetated areas, the influence of bare soil can be minimized (Browning et al., 2017).

Given that grassland ecosystems are water-limited, we hypothesized that incorporating a precipitation or soil moisture requirement into spring phenology models would improve model performance. In this study, we used PhenoCam imagery from 43 diverse grassland sites across North America (195 site-years) to improve upon existing models. We addressed the following questions: 1) Which model and model parameters best predict SOS across varied grassland ecosystems? 2) Does the use of region-specific parameters improve model performance? 3) According to the best-fit model, how will future climate change impact the timing of SOS?

2. Methods

2.1 Data Compilation

We downloaded metadata for all the PhenoCam sites, including location, climate, and primary vegetation type using the “phenocamapi” R package (R 4.1.0 Core Team, 2021; Seyednasrollah, 2018). We sorted the sites by vegetation type to include only those with a grassland region of interest (ROI) and excluded sites which were visually identified as mowed

fields, forest understory, or experimental plots. We geographically restricted our study area to sites within North America; this corresponds with the highest density of PhenoCam sites and coverage by a high-resolution climate dataset (Daymet, Thornton et al., 2021). In total, we identified 43 North American PhenoCam grassland sites to include in our analyses, totaling 195 site-years of data and representing a spread of ~250 days in spring onset date (Fig 1; Table A.1).

For each site, the green chromatic coordinate (GCC) is calculated by the PhenoCam website (phenocam.nau.edu) for an identified ROI using the red-green-blue color channels of the digital images. GCC is a ratio of green intensity relative to the total brightness of each pixel and is robust to changes in image lighting (Richardson et al., 2018). We downloaded and processed the GCC dataset through 2020 for each grassland site using the “phenocamr” R package (Hufkens et al., 2018). The full procedure is explained in Hufkens et al. (2018), but briefly, this package removes outliers, smooths the data, and identifies spring transition dates for each site. To minimize day-to-day noise in the data due to changes in weather and illumination geometry, we used the 90th percentile GCC values from the 3-day data product (Richardson et al., 2018). We extracted the 50% spring green-up date (SOS) for each data-year, corresponding to the day of the year in which a site has reached 50% of its maximum annual greenness (GCC). We tried other common thresholds (10% and 25%), but found the models fit best with the 50% threshold. Because the GCC curve is rising steeply at the 50% threshold, these dates have lower uncertainties than those extracted from the shallower portion of the curve (lower thresholds). This allows us to better constrain model parameters while also giving us more power to reject models that are not well-supported by the data.

California grasslands, located on the West Coast of the United States, required additional processing. Due to the Mediterranean climate, these grasslands often have two distinct green-up

peaks – one in the fall/early winter if there is sufficient moisture and then a larger one in the late winter/spring (Liu et al., 2021). However, the amplitude of the fall green-up was not always large enough to be detected, and no model could fit dates that included a mix of late fall and early spring dates. Thus, for sites with two separate greenness peaks, when the earlier green-up peak was identified (5 site-years), we instead estimated the later transition date using the same criteria (date when 50% of maximum greenness was reached) to use in model analyses.

Model development and data compositing was supported by the “phenor” R package (Hufkens et al., 2018). Daymet climate data (1 km x 1km resolution; Thornton et al., 2021) was used for temperature, precipitation, and vapor pressure deficit, and photoperiod was based on site location. The climate data were included for the “water-year,” from September 21 of the prior year to September 20 of the year in which the transition date occurs. Interpolated datasets, such as Daymet, often have higher precipitation inaccuracies in the North American Southwest due to low gauge density and high spatial and temporal variability in rainfall (Jing et al., 2017; Henn et al., 2018). Therefore, we found better model accuracy using on-the-ground precipitation data for this region when available. For nine Southwest sites (Table A.1), we replaced Daymet precipitation values with site-level precipitation data collected by Ameriflux (<https://ameriflux.lbl.gov/>) or the Southwest Experimental Garden Array (SEGA, <https://sega.nau.edu/>).

Soil moisture data was downloaded from the North American Land Data Assimilation System (NLDAS-2; <https://ldas.gsfc.nasa.gov/nldas>) due to limited and inconsistent *in-situ* measurements. NLDAS-2 provides three different gridded soil moisture datasets (Noah, Mosaic, VIC) at varying depths that are estimated using different methods. We found that the Mosaic

dataset at 0-10 cm soil depth resulted in the best model fits and retained these data in our final analysis.

2.2 Model Structures

We altered existing spring phenology models to include either a precipitation (“W”) or soil moisture (“SM”) component, as well as developed several novel model structures. The “phenor” R package provides a framework for easily fitting and comparing common phenology models (Hufkens et al., 2018). We fit 20 existing (“original”) spring phenology models and compared them to 33 new models (Tables 1 and A.2, Hufkens et al. 2018). The Water Time (WT) model is the simplest new model, which states that a site must accumulate enough precipitation for SOS to occur. We also tested a couple variations that include a photoperiod requirement, the Photo Water Time (PWT) and M1 (M1W) models. Unlike forests, grasslands have not been found to require chilling (Cao et al., 2018, Wang et al., 2022), so following Chen et al. (2014), we replaced the chilling requirement with a precipitation requirement in both the Sequential (SQW) and Parallel (PAW) models. These state that sufficient precipitation must occur either before (SQW) or concurrently (PAW) with temperature accumulation. We also tried a sequential model with the reverse order (SQWr), in which the temperature requirement must be met before precipitation accumulation begins. For all the models, we also included a model variant in which precipitation accumulates in a sigmoidal fashion, so that each precipitation event is given a weight between 0-1 (lowercase “s” added to model code, Table 1). Finally, soil moisture (SM) is often a better indicator of plant water availability than precipitation, which is subject to losses via evaporation and runoff (Liu et al., 2013; Tao et al., 2020), so we developed a final class of models that replaced precipitation with soil moisture (“SM” models; Tables 1 and A.2).

For the best-fit model structure (SQW), we tested several additional model variants. We found that the optimized value for one parameter was usually close to zero (P-base: minimum precipitation event size added to the accumulated precipitation total), so we included a model variation without this parameter (SQW_NoPbase) and excluded it from all subsequent model variants. The other variants introduced a “reset” parameter, so that either precipitation or temperature accumulation restarted if a certain threshold was surpassed. Our thought was that certain extreme events might interrupt or delay plant development. For example, the SQW_cdd model states that precipitation accumulation restarts if a certain number of consecutive dry days (cdd) occur. Likewise, in the SQW_Tmin model, temperature accumulation restarts if the daily minimum temperature (Tmin) drops below a certain threshold. SQW_cdd_Tmin combines both reset thresholds, and SQW_Pi_Tmin states that precipitation accumulation resets if a minimum temperature threshold is reached. As with the other models, we also tried all the variants with a sigmoidal precipitation accumulation structure (Table 1).

2.3 Fitting Models

We used generalized simulated annealing to optimize model parameters using the “GenSA” R package (Xiang et al., 2013). This process is commonly used for fitting phenological models (Chaine, 2000). Parameter ranges were selected to be wide enough to cover the full range of plausible values but narrow enough for an efficient parameter search (Appendix B). To identify the best-fit parameters, we ran the optimization algorithm 25 times for each model, with 100,000 iterations each time. We selected parameters from the model run with the lowest RMSE for each model. Performance between models was compared using the Akaike Information Criterion (AIC). To visualize model performance, we used the site-year residuals (observed-

predicted dates) for each model to build a hierarchical tree that grouped models based on similar performance (“hclust” function in R with the clustering method set to “average”).

For each of the top models, we validated the results using a leave-one-site-out method. We fit the model (100,000 iterations) excluding all data for one PhenoCam site (N = 1-16 years per site), and then used the resulting optimized parameters to predict the SOS dates for the excluded site based on its climate data. We repeated this process for each site and combined all the predicted SOS dates for the excluded sites into a single dataset, which we used to calculate the RMSE of the validation run. As above, this process was repeated 25 times to ensure the best-fit parameters were identified.

2.4 Dividing Sites by Climate

Given that grassland plants in different climates likely have different temperature and precipitation requirements (White et al., 1997; Xin et al., 2015; Ren et al., 2018), we also divided the sites by climate and re-ran the models for each subset of sites to obtain region-specific parameters. We divided the sites based on their Köppen-Geiger (KG) climate classification (Peel et al., 2007), but to limit the number of groups and ensure a sufficient sample size for each, we only used the first two of the three nested climate criteria (general climate and seasonal precipitation timing, indicated by letters). The dry sites were an exception, which we combined into a single group because there were fewer of them. Therefore, the sites fell into four separate KG climate groups: B: arid and semi-arid (12 sites, 46 site-years), Cf: temperate and humid subtropical (8 sites, 32 site-years), Cs: Mediterranean/dry summer (10 sites, 43 site-years), and Df: humid continental (13 sites, 74 site-years).

We ran the models separately for each KG group using the same methods described above. Then, for each of the top three regional models (and several extra), we combined the

predicted SOS dates for all four KG groups into one dataset and calculated the overall RMSE and AIC. Since AIC balances model fit with complexity, using regionally optimized parameters added extra complexity “cost” to the AIC calculation. For example, if a model had five parameters that were optimized separately for each of the four KG groups, when the results were combined, 20 parameters (5 parameters x 4 groups) were utilized in total (see Table A.5). This allowed for a direct comparison between model runs using a single parameter set (All Sites) and those using four separate climate-specific parameter sets (KG groups). For the best overall model, we performed separate leave-one-site-out validations for each KG group using the same methods as described above. We then combined those results into a single dataset to calculate the validation RMSE across all sites when each region was separately parameterized.

2.5 Future Projections

To determine how the timing of spring onset is predicted to shift as a result of climate change, we used the top model to predict future SOS dates across several grassland sites. Given the uncertainty of future climate projections, especially for precipitation patterns (Mishra et al., 2012; Polley et al., 2013), we included an ensemble of projected climate scenarios. We chose a representative site from each KG climate zone that contained the most typical grassland type for the region and downloaded projected daily precipitation and temperature data (2007-2100) for each from the NA-CORDEX data collection (<https://www.earthsystemgrid.org/search/cordexsearch.html>) from several different regional climate models (RCM) under a “business as usual” (representative concentration pathway “RCP” 8.5) scenario. Each RCM can be driven by various global climate models, and we chose to only use model output that had been bias-corrected against Daymet climate data (see <https://na-cordex.org/dataset-description>) to better match the data used in our analyses. This resulted in a

total of 12 future climate scenarios. For each, we formatted the data for use in “phenor” and ran the top model using the optimized model parameters for each site’s respective KG climate zone to obtain predicted yearly SOS dates. To estimate the rate of change in SOS dates through time, we fit a Sen’s slope (SOS vs year, “zyp” package: Bronaugh & Werner, 2019) for each scenario and calculated the mean and standard deviation across all the scenarios for each location. Then, to evaluate the statistical significance of the trend, we used the median SOS dates across all the scenarios to perform a Mann-Kendall test for each site (“Kendall” package: McLeod, 2022).

3. Results

3.1 Model Performance

The original spring phenology models (cells with black text, Table 1) fit poorly across the diverse grassland sites, with the RMSE between predicted and observed SOS dates ranging from 26.9-40.3 days (Fig 2 and Table 2). The models were particularly poor at predicting the earliest and latest transition dates (Model II regression slopes = 0.22-0.76, Fig 2). Although, models that included VPD (SGSI, AGSI) fit better (both RMSE ~27 days) than those with only chilling and/or degree-days as drivers. Interestingly, the only model that included precipitation (GRP) had the worst overall performance (RMSE = 40 days). All model fits are reported in Table 2.

Of the new models tested, those with only precipitation (with or without photoperiod) fit poorly (RMSE = 26-37 days). Instead, the best models included both precipitation and temperature (RMSE values: SQW = 18.4, SQWr = 21.5, PAW = 22.8 days). These performed better than any of the original models (Table 2). Models that used soil moisture instead of precipitation had similar patterns, but generally performed slightly worse (RMSE = 19.4 – 31.9 days; Table A.3). Additionally, the sigmoidal precipitation accumulation structure, in which very

275 small and large events are given less influence, usually improved model fit (RMSE values:
276 SQWs = 16.2, SQWr_s = 20.5, PAWs = 22.0 days). Of these models, the sigmoidal sequential
277 model (SQWs) was the best based on both RMSE and AIC (Table 2). Figure 3a shows how
278 model predictions for individual site-years were altered by using a sigmoidal precipitation
279 accumulation structure (SQWs) rather than a simple summation (SQW). Early transitioning sites
280 were especially improved (closer to 1:1 line) largely due to differences in the t_0 parameter,
281 which determines the number of days after the start of the water year (Sept 21) that precipitation
282 can start accumulating. SQWs started precipitation accumulation much earlier than SQW (t_0 =
283 47 vs 138 days), allowing the predicted SOS dates to be much closer to the observed dates for
284 early transitioning sites. In contrast, the larger t_0 value in SQW delayed precipitation
285 accumulation until February, after the observed SOS date for some sites had already passed.

286 The sigmoidal precipitation accumulation structure in SQWs is key to accommodating
287 this earlier start date (t_0). In this model structure, each precipitation event is assigned a value
288 between 0-1 based on the optimized model parameters that determine the shape of the sigmoidal
289 curve. For example, Figure 4a shows the sigmoidal relationship between rain event size and its
290 assigned weight based on the best-fit “b” and “c” parameters from the model (Table 3). For the
291 model fit with All Sites (grey line), any daily event less than ~3 mm is assigned a value less than
292 1, and any event greater than ~3 mm is assigned a value of 1. Thus, small events do not influence
293 the precipitation total as much as large events, but at the same time, all events > 3 mm are given
294 equal weight. Large rain events do not cause the site to reach the required precipitation threshold
295 (P-req) sooner because proportionally more water is usually lost to run-off (Fig 4b). This
296 dampened impact of individual rainfall events allowed for the utilization of a larger collection

window starting earlier in the water-year, which consequently improved predictions for sites with early transition dates.

Based on AIC comparisons, all the SQW model variants were better than any other model type, and those that included Tmin (SQWs_Tmin and SQWs_cdd_Tmin) performed better than SQWs (Table 2). This suggests that the addition of a temperature reset parameter, in which temperature accumulation resets when the minimum temperature drops below a certain threshold, such as during a spring cold snap, is beneficial for at least some sites by delaying their predicted SOS date (Fig 4c). Figure 3b shows how individual site-years were altered by the addition of this parameter; most were only slightly affected, but some individual points were significantly improved. However, the addition of a parameter that reset precipitation accumulation (SQW_cdd and SQW_Pi_Tmin) did not improve model performance. Thus, the best overall model for predicting SOS across diverse grasslands was SQWs_Tmin (RMSE = 16.0 days, $R^2 = 0.85$, Fig 5a). The RMSE of the best leave-one-site-out validation run was 19.1 days, and the average across all 25 validation runs was 21.04 ± 1.48 days.

3.2 Models by Climate Group

Unsurprisingly, model performance was further improved by separately optimizing model parameters for different climate regions – the model parameters can thus be interpreted as representing the phenological adaptation of populations or species to local conditions. When all the models and model variations were run separately for each KG group, SQWs_Tmin was consistently one of the top regional models. Based on RMSE, it was the top model for B (RMSE = 7.9 days), Cs (RMSE = 16.3 days), and Df (RMSE = 7.7 days), and within 0.5 days of the best model for Cf (RMSE = 8.4 days; Table A.4). However, based on AIC, slightly less-complex models (fewer parameters) performed marginally better for several of the KG regions, though

they all had the same basic model structure, with precipitation accumulating before temperature (Table A.4).

When the results of the top regional models for each KG group were combined and analyzed, SQWs_Tmin was once again the best overall model, with a substantially lower AIC value than the same model using only a single set of parameters (AIC = 968.7 vs 1094.7; Table A.5). Using regionally optimized parameters also reduced the RMSE from 16.0 to 10.4 days ($R^2 = 0.93$, Fig 5b). The best-fit parameters and RMSE values for All Sites and each KG group for the SQWs_Tmin model are listed in Table 3. The combined RMSE for the leave-one-site-out regionally parameterized validation of SQWs_Tmin was 16.0 days.

Running the model separately for each KG climate group improved performance because it allowed parameters to be optimized for region-specific attributes (Table 3). For example, Cs had a much lower t_0 (34 days; number of days after the start of the water year before precipitation can start accumulating) than Cf ($t_0 = 198$ days). This means that Mediterranean sites (Cs) started accumulating precipitation in October, which makes sense considering California grasslands can green-up as early as November. In contrast, temperate grasslands (Cf) become active later in the spring, so did not start accumulating precipitation until April since earlier precipitation (such as October rainfall) has little influence on their spring physiological activity. However, colder Northern Great Plains grasslands (Df) started accumulating precipitation early ($t_0 = 46$) despite not greening up until later in the spring, but this was accounted for by requiring greater total accumulated precipitation.

Precipitation requirements also differed by climate. Based on the optimal parameters for each KG group, precipitation accumulated differently (“b” and “c” parameters determine shape of sigmoid curve), as well as the required precipitation amount (P-req; Table 3, Fig 4a). For

example, in arid grasslands (B), the precipitation accumulation curve reaches a horizontal asymptote at a rain even size of ~ 1 mm, so all larger events provide an equal amount of forcing (Fig 4a). Therefore, based on the required precipitation parameter ($P\text{-req} = 22.6$; Table 3), these sites must receive ~ 22 total rain events of at least 1 mm to surpass the precipitation threshold. In contrast, wetter regions require larger precipitation events to initiate spring green-up. Temperate grasslands (Cf) require a single ($P\text{-req} = 0.9$), but much larger (asymptote ~ 10 mm; Fig 4a) rain event, and Mediterranean grasslands (Cs) require about 3 events ($P\text{-req} = 2.8$) greater than ~ 8 mm or several smaller events of equivalent size (shallower curve; Table 3, Fig 4a). Interestingly, because the y-intercept of the sigmoidal curve is 0.4 for cooler grasslands (Df), these sites accumulate “precipitation” even on days with no rain (i.e., days with 0mm of rainfall contribute 0.4 in forcing; Fig 4a). Thus, since no-rain days advance the “state of water” in this region, $P\text{-req}$ is a function of both precipitation and the amount of time that has passed since accumulation began.

3.3 Visualizing Model Performance

The hierarchical tree constructed using model residuals (observed-predicted SOS dates) provides a visual representation of the similarity between model performances (Fig 6). Models with more similar predicted SOS dates are grouped closer together and those with more different dates are farther apart. For ease of interpretation, the clustering algorithm divided the models into nine groups (salmon-colored boxes, numbered 1-9) based on similarity. From this, it is evident that the original models grouped separately (black text; Groups 7-9) from the new models that included precipitation or soil moisture as a driver (Groups 1-6), suggesting the addition of water availability altered model performance in a consistent fashion. Interestingly, the three original models that included a measure of water availability (GRP, SGSI, AGSI) grouped separately

(Groups 7 & 9) from the rest of the original models, but also did not cluster with the new models due to their unique model structures (Hufkens et al., 2018). Within the new models, those with both water availability (precipitation or soil moisture) and temperature had more similar predicted SOS dates (Groups 4-6) than those with only water availability (with or without photoperiod; Groups 1-3). Also, the soil moisture models (green text; Groups 2 & 5) clustered separately from the precipitation models (blue text; Groups 1, 3, 4, 6), and the precipitation models with a sigmoidal accumulation structure (lowercase “s” in model code) generally grouped separately from those with a simple precipitation summation (Groups 1 vs 3 and Groups 4 vs 6). When the top three models based on AIC were identified for All Sites and each KG group (colored stars), with one exception, the top three models were all sequential with precipitation and temperature as drivers (Groups 4 & 6, SQW model structure). This confirms that, despite differences in the best SQW model variant for each KG group, all the top models perform similarly due to underlying commonalities in model structure, and the general precipitation-temperature sequential model structure is best for predicting SOS across diverse grassland types.

3.4 Future Projections

Forward projections for a representative site for each KG group using the top model (SQWs_Tmin) suggest that SOS will occur earlier across much of the North American grassland region by the end of the century. Ensemble means predicted that SOS in temperate grasslands will occur earlier, with the cooler grassland site (Df, North Dakota: 11.1 ± 2.7 days over 100 years, $p < 0.001$) experiencing a greater shift than the warmer grassland site (Cf, Kansas; 4.7 ± 2.3 days over 100 years, $p < 0.001$; Fig 7). Although, the warmer site was constrained by a much larger t_0 value (198 days), which means precipitation cannot begin accumulating until early

April, limiting the amount that SOS can advance. Mediterranean grasslands (Cs, California) also exhibited a shift to earlier spring onset (12.3 ± 6.0 days over 100 years, $p < 0.001$), but the signal for arid grasslands was inconsistent (B, New Mexico; SOS delayed 3.5 ± 8.0 days over 100 years, $p = 0.776$; Fig 7). This was because some of the climate scenarios suggested a slight delay in SOS for this region, and others a slight advancement, resulting in a large standard deviation and no clear directional shift.

4. Discussion

Grasslands cover 30% of North America and heavily influence both regional and global water and energy fluxes, as well as represent a significant carbon sink (Abberton et al., 2009; Pendall et al., 2018). Therefore, determining how various climate factors influence grassland plant phenology is important for understanding current ecological processes as well as future impacts of climate change. However, many established (“original”) phenology models were made for forested ecosystems and perform poorly in grasslands (Xin et al., 2015; Liu et al., 2018), likely because most do not include water availability. We used PhenoCam images from 43 diverse North American grassland sites to test numerous existing spring phenology models as well as several new model variations that incorporate precipitation or soil moisture. Our best new model suggests that grasslands require sufficient precipitation followed by warm temperatures to initiate spring green-up.

4.1 Original Models

In agreement with past studies (Fu et al., 2014; Xin et al., 2015; Cao et al., 2018; Liu et al., 2018; Ren et al., 2022), the original models poorly predicted the start of the growing season (SOS) in grassland ecosystems, with the difference between predicted and observed dates

(RMSE) ranging between 4-6 weeks (Fig 2 & Table 2). Similar to the findings of Liu et al. (2018), the original models were especially poor at predicting early and late SOS dates – they tended to predict SOS too late in early-transitioning Mediterranean grasslands and too early in late-transitioning arid grasslands (Fig 2). This is because these ecosystems need more than just temperature and/or daylength for spring initiation, but also sufficient precipitation inputs. California grasslands typically green-up early in the year, coinciding with a period of available moisture before the start of the summer dry season (Huenneke & Mooney, 1989; Liu et al., 2021). Likewise, many Southwestern arid grasslands depend on monsoonal rains that occur later in the summer (Jul – Sept) to become active (Adams & Comrie, 1997), so their green-up dates occur much later in the year. Thus, the absence of a precipitation driver in most of the original models did not allow for this flexibility in timing and resulted in inaccurate predicted SOS dates.

Several of the original models did incorporate a measure of water availability. AGSI and SGSI include vapor pressure deficit (VPD) as a driver (Xin et al., 2015), but while they did perform best of the original models, both still had errors of about four weeks (RMSE ~27 days; Fig 2 & Table 2). This suggests that either the model drivers (photoperiod, VPD, temperature) were not sufficient predictors of SOS or that the model structure was not appropriate.

Interestingly, the only model that explicitly included precipitation as a parameter (GRP) performed the worst, with a RMSE of ~40 days (Fig 2). This could be because the GRP model only considers precipitation accumulated the week prior to the predicted SOS date and, as evidenced by the much longer precipitation accumulation period in our new best-fit model, this is not long enough to fully trigger spring green-up in most grassland ecosystems.

4.2 New Models

Most of the new spring phenology models performed significantly better than the original ones, and of those, precipitation models were generally better than soil moisture models. Past modeling studies have found mixed results using soil moisture in place of precipitation or VPD (Liu et al., 2013; Xin et al., 2015; Tao et al., 2020). This could be attributed to inaccurate soil moisture estimates from gridded data products (Xia et al., 2014), so with more accurate data, soil moisture could be a better driver of grassland phenology. However, in this study, we found the best new models included precipitation and temperature (SQW, SQWr, PAW); models with only one or the other performed poorly. Thus, in agreement with other studies (Chen et al., 2014; Moore et al., 2015; Ren et al., 2018, 2022; Fan et al., 2020; Fu et al., 2021), both sufficient temperature and precipitation are required for grasslands to become active in the spring. However, the order in which they accumulate is important. Our results suggest that a site must first accumulate enough precipitation, followed by temperature (sequential, SQW). This was better than the opposite order (temperature before precipitation; reverse sequential, SQWr; RMSE ~3 days greater) or accumulating both at the same time (parallel, PAW; RMSE ~4.5 days greater; Table 2). Several recent studies have found similar results (Chen et al., 2014; Ren et al., 2022). For example, Ren et al. (2022) used satellite data to fit several spring phenology models across all Northern Hemisphere grasslands and found the sequential (SQW) model was best for a majority of the study region. This is likely due to the physiological requirements for grass germination and/or tiller emergence. As any gardener knows, seeds placed in dry soil will not grow, even with ample sunlight and warm temperatures. Likewise, grass seeds require moist soil conditions to germinate (Qi & Redmann, 1993; Abbott & Roundy, 2003; Springer, 2005; Durr et al., 2015), as do the belowground buds of perennial grasses to produce tillers (Korte & Chu,

1983; Russell et al., 2017). Thus, in water-limited environments, sufficient precipitation is a prerequisite to warm temperatures for the initiation of plant physiological processes.

We also found that using a sigmoidal accumulation structure for precipitation generally improved model fit (SQWs). This introduces a minimum and maximum effective precipitation size, thereby altering the way that individual precipitation events contribute to the required precipitation accumulation. In agreement with prior studies, which have found small events to be less effective at stimulating plant growth in grasslands (Huxman et al., 2004; Heisler-White et al., 2009; Post & Knapp, 2021), events that are too small are not counted. Likewise, excess rainfall from large rain events can be lost via runoff or deep drainage, so does not promote greater plant available moisture (Sala et al., 1988; Ye et al., 2016). Thus, all events over a certain size threshold contribute equally to the precipitation accumulation. In this way, the sigmoidal accumulation structure improved model accuracy by accounting for ecologically relevant rainfall.

Likewise, the addition of a temperature accumulation reset parameter improved model performance, making SQWs_Tmin the best overall model. In this model, temperature accumulation resets if the daily minimum temperature drops below a certain threshold (T-thres; Fig. 4c). This causes a delay in reaching the temperature accumulation threshold (F-crit), resulting in a later predicted SOS date. Thus, the temperature reset parameter improves model performance when spring temperatures are abnormally low. A late cold snap during the spring season can hinder the initiation of plant physiological processes and growth (Allen & Ort, 2001; Kovi et al., 2016; Kong & Henry, 2019), though more research is needed on the direct impacts on grassland systems (Kral-O'Brien et al., 2019).

Importantly, using a single set of parameters, the SQWs_Tmin model was able to accurately predict SOS across diverse grassland sites within about two weeks (RMSE = 16.0 days; Fig 5a). This represents a significant improvement over the results from past studies, especially for warm, water-limited grasslands. Using remote sensing data, Liu et al. (2018) tested six traditional spring models across the Northern Hemisphere and found RMSE values of 15.7-18.5 days for C₃-dominated (cool) grasslands, and 25.8-27.3 days for C₄-dominated (warm) grasslands. In a similar remote sensing study, Fu et al. (2014) found RMSE values of 18-37 days over North American grasslands for both a temperature-only (TT) model, as well as a model with both precipitation and soil temperature that is used in several common earth system models (Biome-Biogeochemical Cycle model; White et al., 1997; Zhang et al., 2019). Other studies have tested additional spring phenology models in various grassland regions, but their results are hard to compare with ours because they either fit models separately for each satellite pixel (Ren et al., 2022) or cover a different and/or smaller geographic region (White et al., 1997; Chen et al., 2014; Cao et al., 2018; Fan et al., 2020). For example, our model is about a day better than the AGSI model when fit across the Western US by Xin et al. (2015), but their study excluded regions with unique phenological signals (i.e., California & Southwest). Our inclusion of these regions likely led to slightly higher RMSE values in our analyses. Nonetheless, the SQWs_Tmin model represents a clear improvement for predicting SOS in grassland ecosystems, which can be attributed to two factors: 1) the use of fine-scale, near-surface remote sensing data (PhenoCam) instead of coarse resolution satellite data and 2) a new flexible model structure that better captures the influence of multiple climate variables across diverse grassland systems.

4.3 Regional models

The SQWs_Tmin model worked well for all grasslands using a single set of parameters, but model performance was further improved by using region-specific parameters. Most regional fits (KG groups B, Cf, Df) predicted SOS within about a week of observed dates (Table 3; RMSE = 7.8-8.4 days). The exception was Cs, which had a larger error (RMSE = 16.4 days) because this group included two very different grassland types, despite both having extended dry periods (Fig 1). California grasslands green-up in late winter/early spring (DOY 1-100), whereas high elevation Southwestern grasslands green-up in response to late-summer monsoons (DOY 150-200). Given this ~200-day spread in green-up dates, it is impressive that the model was accurate to within about two weeks. Notably, all the top models for each KG group had the same model structure with precipitation accumulating before temperature (Fig 6), and when regional results were combined, SQWs_Tmin was the best overall model with predictions within ~10 days of observed transition dates (Fig 5b). Thus, using region-specific parameters improved model accuracy by about six days.

Due to differences in species composition, it follows that grasslands in different climate zones have different temperature and precipitation requirements. For example, the Northern Great Plains consists of predominantly cool-adapted C₃ grasses, whereas the Southern Great Plains and desert grasslands contain mostly warm-adapted C₄ grasses (Paruelo & Lauenroth, 1996; Wang et al., 2013). The KG climate zones Df (cool C₃) and Cf (warm C₄) roughly align with this functional group transition (Fig 1; Paruelo & Lauenroth, 1996; von Fischer et al., 2008), which likely accounts for the difference in optimized parameters between these groups. For instance, Df has a much lower minimum temperature reset threshold ($T_{\text{thres}} = -12.9\text{ }^{\circ}\text{C}$) than Cf ($T_{\text{thres}} = -0.3\text{ }^{\circ}\text{C}$; Table 3) because plants in that region can tolerate colder temperatures

(Epstein et al., 1998; Liu & Osborne, 2008). Likewise, precipitation requirements differ substantially between climate groups. In agreement with prior studies that have found dryland plants to respond to small precipitation pulses (Sala & Lauenroth, 1982; Liang et al., 2021), arid grasslands (B) can utilize small rainfall events (~ 1 mm, Fig 4a). In contrast, mesic grasslands (Cf) require a larger precipitation pulse to activate spring growth (~ 10 mm, Fig 4a) because the plants are adapted to wetter soil conditions (Knapp et al., 2008). Interestingly, cooler grasslands (Df) continuously accumulated “precipitation” even when no rain occurred, suggesting a diminished importance of precipitation inputs for this system (Fig 4a). This is because the Northern Great Plains (Df) are more limited by temperature than by precipitation (Jolly et al., 2005) since they can utilize moisture stored in the soil from snow melt and/or abundant spring rainfall (Ren et al., 2018, Mohammed et al., 2019). In fact, temperature-only models worked decently well in this region (RMSE ~ 9 days, Table A.4), and Ren et al. (2022) found that a temperature-only model (TT) was best for cool, wet grassland regions. Nonetheless, we found that precipitation-temperature models consistently out-performed temperature-only models in this region because they better accommodate years that are abnormally dry.

Given the unique climatological and phenological patterns of California grasslands (within KG group “Cs”), our models performed remarkably well for this region. California grasslands tend to green-up much earlier than other North American grasslands, and depending on precipitation patterns, they can have two distinct seasonal green-up cycles – one in the late fall/early winter and another in the late winter/early spring (Huenneke & Mooney, 1989; Liu et al., 2021). They are also dominated by annual grasses, so species composition can change drastically from one year to the next depending on seed dispersal and recruitment (Huenneke & Mooney, 1989; Bart et al., 2017). For these reasons, many large-scale phenology studies have

excluded this region from analysis (e.g., Xin et al., 2015; Ren et al., 2018). Bart et al. (2017) is one of the few studies that has attempted to model the phenological drivers of California grasslands. They used MODIS satellite data at two different sites in California to develop a SOS grassland model using soil water potential, temperature, and VPD. However, even when the model was parameterized separately for each site, RMSE was still 20-22 days. Thus, our new model (SQWs_Tmin) is a significant improvement, with a RMSE of 16.4 days across all Mediterranean (Cs) grassland sites, including California grasslands. Although, it is important to note that our model is currently able to predict only a single annual green-up event, so there is opportunity for improvement with a model that could simulate multiple green-up events.

4.4 Forward Projections

Using projected climate data and regionally optimized parameters for the top model (SQWs_Tmin), we found that spring onset in most North American grasslands is predicted to occur about 5-12 days earlier within the next 100 years (Fig 7). This aligns with previous studies that have forecasted earlier spring onset in grassland systems due to warming air temperatures (Zhang et al., 2013; Li et al., 2016; Chang et al., 2017; Wang et al., 2020). However, our observed shift is less drastic than the findings of a recent paper, which predicted spring to occur 21-25 ($\pm$ 3-4) days earlier by 2100 across most the Great Plains region (Hufkens et al. 2016). Similar to other studies (Polley et al., 2013; Hufkens et al., 2016), the largest shifts were in grassland systems that are primarily temperature limited at the beginning of the growing season, such as the Northern Great Plains (~11 days) and California grasslands (~12 days). In contrast, the arid grassland site did not exhibit a clear pattern (Fig 7). This is likely because, in arid grasslands, green-up is more sensitive to changes in precipitation than temperature (Currier & Sala, 2022), and highly variable SOS dates make it difficult to detect a clear trend in spring

onset. Overall, our results suggest that climate change will generally cause earlier spring onset across most North American grasslands, but the magnitude of that advancement will be highly region-specific.

5. Conclusion

Knowledge of grassland plant phenology and its drivers is crucial to developing a more complete understanding of both regional and global ecosystem processes. Using PhenoCam data, we found a new model (SQWs_Tmin) that accurately (within ~2 weeks) predicts SOS across diverse grassland types better than any previous model, from the Great Plains to Mediterranean annual grasslands to arid grasslands. This new model emphasizes the influence of precipitation on grassland phenology, in conjunction with temperature. While SQWs_Tmin works well across all grasslands with a single parameter set, our results highlight the advantages of using region-specific parameters for improving overall model performance (from 16 to 10 days). This is because plants in different climates differ in their temperature and moisture requirements to initiate spring growth. Using this new model and projected climate data (RCP 8.5) for a representative site within each KG climate zone, we determined that spring onset is expected to occur 5-12 days earlier across most North American grasslands within 100 years. Arid grasslands were the exception, which exhibited no substantial shift because they are more limited by precipitation than temperature. By integrating SQWs_Tmin into larger earth system models, predictions of spring onset and associated ecosystem processes could be greatly improved. Our new grassland phenology model represents a significant development in understanding and predicting the drivers of spring onset across diverse grassland types, thereby improving estimates of carbon, nutrient, and water cycling across these prolific systems, both now and in the future.

Acknowledgements

We thank our many collaborators, including site PIs and technicians, for their efforts 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 also thank two anonymous reviewers for their thoughtful suggestions, which greatly improved the manuscript.

Data Availability

Imagery data are available from the PhenoCam website (phenocam.sr.unh.edu) and climate data are available from Daymet (daymet.ornl.gov), Ameriflux (ameriflux.lbl.gov), SEGA (sega.nau.edu), and NLDAS-2 (<https://ldas.gsfc.nasa.gov/nldas>). Full model equations are listed in Table A.2 and parameter ranges are include in Appendix B. All code is available on [Github \(\)](#).

References

- Abberton, M., R. Conant, and C. Batello, editors. 2009. Grassland carbon sequestration: management, policy, and economics. Food and Agriculture Organization of the United Nations, Rome.
- Abbott, L. B., and B. A. Roundy. 2003. Available water influences field germination and recruitment of seeded grasses. *Journal of Range Management* 56: 56–64.
- Adams, D. K., and A. C. Comrie. 1997. The North American Monsoon. *Bulletin of the American Meteorological Society* 78: 2197–2213.

613 Allen, D. J., and D. R. Ort. 2001. Impacts of chilling temperatures on photosynthesis in warm-
614 climate plants. *Trends in Plant Science* 6: 36–42.

615 Bart, R. R., C. L. Tague, and P. E. Dennison. 2017. Modeling annual grassland phenology along
616 the central coast of California. *Ecosphere* 8(7): e01875.

617 Basler, D. 2016. Evaluating phenological models for the prediction of leaf-out dates in six
618 temperate tree species across central Europe. *Agricultural and Forest Meteorology* 217:
619 10–21.

620 Bronaugh, D. and A. Werner. 2019. Pacific Climate Impacts Consortium. zyp: Zhang + Yue-
621 Pilon. Trends Package. R package version 0.10-1.1. [https://CRAN.R-](https://CRAN.R-project.org/package=zyp)
622 [project.org/package=zyp](https://CRAN.R-project.org/package=zyp).

623 Browning, D. M., J. W. Karl, D. Morin, A. D. Richardson, and C. E. Tweedie. 2017. Phenocams
624 bridge the gap between field and satellite observations in an arid grassland ecosystem.
625 *Remote Sensing* 9: 10971.

626 Cao, R., M. Shen, J. Zhou, and J. Chen. 2018. Modeling vegetation green-up dates across the
627 Tibetan Plateau by including both seasonal and daily temperature and precipitation.
628 *Agricultural and Forest Meteorology* 249: 176–186.

629 Caparros-Santiago, J. A., V. Rodriguez-Galiano, and J. Dash. 2021. Land surface phenology as
630 indicator of global terrestrial ecosystem dynamics: A systematic review. *ISPRS Journal*
631 *of Photogrammetry and Remote Sensing* 171: 330–347.

632 Chang, J., P. Ciais, N. Viovy, J. F. Soussana, K. Klumpp, and B. Sultan. 2017. Future
633 productivity and phenology changes in European grasslands for different warming levels:
634 Implications for grassland management and carbon balance. *Carbon Balance and*
635 *Management* 12(11).

636 Chen, X., J. Li, L. Xu, L. Liu, and D. Ding. 2014. Modeling greenup date of dominant grass
637 species in the Inner Mongolian Grassland using air temperature and precipitation data.
638 International Journal of Biometeorology 58: 463–471.

639 Chen, X., D. Wang, J. Chen, C. Wang, and M. Shen. 2018. The mixed pixel effect in land
640 surface phenology: A simulation study. Remote Sensing of Environment 211: 338–344.

641 Choler, P., W. Sea, P. Briggs, M. Raupach, and R. Leuning. 2010. A simple ecohydrological
642 model captures essentials of seasonal leaf dynamics in semi-arid tropical grasslands.
643 Biogeosciences, 7(3), 907–920.

644 Chuine, I. 2000. A unified model for budburst of trees. Journal of Theoretical Biology 207: 337–
645 347.

646 CLM5 Documentation. Technical Description of version 5.0 of the Community Land Model
647 (CLM). 2020. https://www.cesm.ucar.edu/models/cesm2/land/CLM50_Tech_Note.pdf.

648 Cui, T., L. Martz, E. G. Lamb, L. Zhao, and X. Guo. 2019. Comparison of grassland phenology
649 derived from MODIS satellite and PhenoCam near-surface remote sensing in North
650 America. Canadian Journal of Remote Sensing 45: 707–722.

651 Currier, C. M., and O. E. Sala. 2022. Precipitation versus temperature as phenology controls in
652 drylands. Ecology: e3793.

653 Dürr, C., J. B. Dickie, X. Y. Yang, and H. W. Pritchard. 2015. Ranges of critical temperature and
654 water potential values for the germination of species worldwide: Contribution to a seed
655 trait database. Agricultural and Forest Meteorology 200: 222–232.

656 Epstein, H. E., W. K. Lauenroth, I. C. Burke, and D. P. Coffin. 1998. Regional productivities of
657 plant species in the Great Plains of the United States. Plant Ecology 134: 173–195.

658 Fan, D., X. Zhao, W. Zhu, W. Sun, and Y. Qiu. 2020. An improved phenology model for
 659 monitoring green-up date variation in *Leymus chinensis* steppe in Inner Mongolia during
 660 1962–2017. *Agricultural and Forest Meteorology* 291: 108091.

661 von Fischer, J. C., L. L. Tieszen, and D. S. Schimel. 2008. Climate controls on C3 vs. C4
 662 productivity in North American grasslands from carbon isotope composition of soil
 663 organic matter. *Global Change Biology* 14: 1141–1155.

664 Fu, Y. H., M. Campioli, M. Van Oijen, G. Deckmyn, and I. A. Janssens. 2012. Bayesian
 665 comparison of six different temperature-based budburst models for four temperate tree
 666 species. *Ecological Modelling* 230: 92–100.

667 Fu, Y. H., X. Zhou, X. Li, Y. Zhang, X. Geng, F. Hao, X. Zhang, H. Hanninen, Y. Guo, H. J. De
 668 Boeck, and H. J. De Boeck. 2021. Decreasing control of precipitation on grassland spring
 669 phenology in temperate China. *Global Ecology and Biogeography* 30: 490–499.

670 Fu, Y., H. Zhang, W. Dong, and W. Yuan. 2014. Comparison of phenology models for
 671 predicting the onset of growing season over the Northern Hemisphere. *PLoS ONE* 9(10):
 672 e109544.

673 Gao, Q., W. Zhu, M. W. Schwartz, H. Ganjurjav, Y. Wan, X. Qin, X. Ma, M. A. Williamson,
 674 and Y. Li. 2016. Climatic change controls productivity variation in global grasslands.
 675 *Scientific Reports* 6: 26958.

676 García-Mozo, H., C. Galán, J. Belmonte, D. Bermejo, P. Candau, C. Díaz de la Guardia, B.
 677 Elvira, M. Gutiérrez, V. Jato, I. Silva, M. M. Trigo, R. Valencia, and I. Chuine. 2009.
 678 Predicting the start and peak dates of the *Poaceae* pollen season in Spain using process-
 679 based models. *Agricultural and Forest Meteorology* 149: 256–262.

680 Heisler-White, J. L., J. M. Blair, E. F. Kelly, K. Harmoney, and A. K. Knapp. 2009. Contingent
681 productivity responses to more extreme rainfall regimes across a grassland biome. *Global*
682 *Change Biology* 15: 2894–2904.

683 Henn, B., A. J. Newman, B. Livneh, C. Daly, and J. D. Lundquist. 2018. An assessment of
684 differences in gridded precipitation datasets in complex terrain. *Journal of Hydrology*
685 556: 1205–1219.

686 Huenneke, L. F., and H. A. Mooney. 1989. *Grassland Structure and Function: California Annual*
687 *Grassland*. Kluwer Academic Publishers, Dordrecht, Netherlands.

688 Hufkens, K., D. Basler, T. Milliman, E. K. Melaas, and A. D. Richardson. 2018. An integrated
689 phenology modelling framework in R. *Methods in Ecology and Evolution* 9: 1276–1285.

690 Hufkens, K., T. F. Keenan, L. B. Flanagan, R. L. Scott, C. J. Bernacchi, E. Joo, N. A. Brunsell, J.
691 Verfaillie, and A. D. Richardson. 2016. Productivity of North American grasslands is
692 increased under future climate scenarios despite rising aridity. *Nature Climate Change* 6:
693 710–714.

694 Huxman, T. E., K. A. Snyder, D. Tissue, A. J. Leffler, K. Ogle, W. T. Pockman, D. R. Sandquist,
695 D. L. Potts, and S. Schwinning. 2004. Precipitation pulses and carbon fluxes in semiarid
696 and arid ecosystems. *Oecologia* 141: 254–268.

697 Jeong, S. J., C. H. Ho, H. J. Gim, and M. E. Brown. 2011. Phenology shifts at start vs. end of
698 growing season in temperate vegetation over the Northern Hemisphere for the period
699 1982–2008. *Global Change Biology* 17: 2385–2399.

700 Jing, X., B. Geerts, Y. Wang, and C. Liu. 2017. Evaluating seasonal orographic precipitation in
701 the interior western United States using gauge data, gridded precipitation estimates, and a
702 regional climate simulation. *Journal of Hydrometeorology* 18: 2541–2558.

703 Jolly, W. M., R. Nemani, and S. W. Running. 2005. A generalized, bioclimatic index to predict
704 foliar phenology in response to climate. *Global Change Biology* 11: 619–632.

705 Klosterman, S. T., K. Hufkens, J. M. Gray, E. Melaas, O. Sonnentag, I. Lavine, L. Mitchell, R.
706 Norman, M. A. Friedl, and A. D. Richardson. 2014. Evaluating remote sensing of
707 deciduous forest phenology at multiple spatial scales using PhenoCam imagery.
708 *Biogeosciences* 11: 4305–4320.

709 Knapp, A. K., C. Beier, D. D. Briske, A. T. Classen, L. Yiqi, M. Reichstein, M. D. Smith, S. D.
710 Smith, J. E. Bell, P. A. Fay, J. L. Heisler, S. W. Leavitt, R. Sherry, B. Smith, and E.
711 Weng. 2008. Consequences of more extreme precipitation regimes for terrestrial
712 ecosystems. *BioScience* 58: 811–821.

713 Kong, R. S., and H. A. L. Henry. 2019. Interactions of plant growth responses to spring freezing
714 and summer drought: a multispecies comparison. *American Journal of Botany* 106: 531–
715 539.

716 Korte, C. J., and A. C. P. Chu. 1983. Some Effects of Drought on Perennial Ryegrass Swards.
717 *Proceedings of the New Zealand Grassland Association* 6: 211–216.

718 Kovi, M. R., Å. Ergon, and O. A. Rognli. 2016. Freezing tolerance revisited — effects of
719 variable temperatures on gene regulation in temperate grasses and legumes. *Current*
720 *Opinion in Plant Biology* 33: 140–146.

721 Kral-O’Brien, K. C., P. L. O’Brien, and J. P. Harmon. 2019. Need for false spring research in the
722 Northern Great Plains, USA. *Agricultural and Environmental Letters* 4: 190025.

723 Li, Q., L. Xu, X. Pan, L. Zhang, C. Li, N. Yang, and J. Qi. 2016. Modeling phenological
724 responses of Inner Mongolia grassland species to regional climate change. *Environmental*
725 *Research Letters* 11: 015002.

726 Li, X., E. Melaas, C. M. Carrillo, T. Ault, A. D. Richardson, P. Lawrence, M. A. Friedl, B.
 727 Seyednasrollah, D. M. Lawrence, and A. M. Young. 2022. A Comparison of Land
 728 Surface Phenology in the Northern Hemisphere Derived from Satellite Remote Sensing
 729 and the Community Land Model. *Journal of Hydrometeorology* 23:859–873.

730 Liang, M., X. Feng, and E. S. Gornish. 2021. Rainfall pulses mediate long-term plant community
 731 compositional dynamics in a semi-arid rangeland. *Journal of Applied Ecology* 58: 708–
 732 717.

733 Lieth, H., 1974. Phenology and seasonality modeling. Springer, Berlin, Heidelberg.

734 Liu, H., Y. Jin, L. M. Roche, A. T. O’Geen, and R. A. Dahlgren. 2021. Understanding spatial
 735 variability of forage production in California grasslands: Delineating climate, topography
 736 and soil controls. *Environmental Research Letters* 16: 014043.

737 Liu, H., F. Tian, H. C. Hu, H. P. Hu, and M. Sivapalan. 2013. Soil moisture controls on patterns
 738 of grass green-up in Inner Mongolia: An index based approach. *Hydrology and Earth
 739 System Sciences* 17: 805–815.

740 Liu, L., and X. Zhang. 2020. Effects of temperature variability and extremes on spring
 741 phenology across the contiguous United States from 1982 to 2016. *Scientific Reports* 10:
 742 17952.

743 Liu, M. Z., and C. P. Osborne. 2008. Leaf cold acclimation and freezing injury in C3 and C4
 744 grasses of the Mongolian Plateau. *Journal of Experimental Botany* 59: 4161–4170.

745 Liu, Q., Y. H. Fu, Y. Liu, I. A. Janssens, and S. Piao. 2018. Simulating the onset of spring
 746 vegetation growth across the Northern Hemisphere. *Global Change Biology* 24: 1342–
 747 1356.

748 McLeod, A.I. (2022). Kendall: Kendall Rank Correlation and Mann-Kendall Trend Test. R
 749 package version 2.2.1. <https://CRAN.R-project.org/package=Kendall>.

750 Migliavacca, M., M. Galvagno, E. Cremonese, M. Rossini, M. Meroni, O. Sonnentag, S.
 751 Cogliati, G. Manca, F. Diotri, L. Busetto, A. Cescatti, R. Colombo, F. Fava, U. Morra di
 752 Cella, E. Pari, C. Siniscalco, and A. D. Richardson. 2011. Using digital repeat
 753 photography and eddy covariance data to model grassland phenology and photosynthetic
 754 CO₂ uptake. *Agricultural and Forest Meteorology* 151: 1325–1337.

755 Mishra, V., F. Dominguez, and D. P. Lettenmaier. 2012. Urban precipitation extremes: How
 756 reliable are regional climate models? *Geophysical Research Letters* 39: L03407.

757 Mohammed, A. A., I. Pavlovskii, E. E. Cey, and M. Hayashi. 2019. Effects of preferential flow
 758 on snowmelt partitioning and groundwater recharge in frozen soils. *Hydrology and Earth
 759 System Sciences* 23: 5017–5031.

760 Moore, L. M., W. K. Lauenroth, D. M. Bell, and D. R. Schlaepfer. 2015. Soil water and
 761 temperature explain canopy phenology and onset of spring in a semiarid steppe. *Great
 762 Plains Research* 25: 121–138.

763 Paruelo, J. M., and W. K. Lauenroth. 1996. Relative abundance of plant functional types in
 764 grasslands and shrublands of north America. *Ecological Applications* 6: 1212–1224.

765 Peel, M. C., B. L. Finlayson, and T. A. McMahon. 2007. Updated world map of the Köppen-
 766 Geiger climate classification. *Hydrology and Earth System Sciences* 11: 1633–1644.

767 Pendall, E., D. Bachelet, R. Conant, B. El Masri, L. B. Flanagan, A. K. Knapp, J. Liu, S. Liu, and
 768 S. M. Schaeffer. 2018. State of Air, Land, and Water: Grasslands. Pages 399–427.
 769 Second State of the Carbon Cycle Report: A Sustained Assessment Report.

770 Polley, H. W., D. D. Briske, J. A. Morgan, K. Wolter, D. W. Bailey, and J. R. Brown. 2013.
 771 Climate change and North American rangelands: Trends, projections, and implications.
 772 Rangeland Ecology and Management 66: 493–511.

773 Post, A. K., and A. K. Knapp. 2021. How big is big enough? Surprising responses of a semiarid
 774 grassland to increasing deluge size. Global Change Biology 27: 1157–1169.

775 Qi, M. Q., and R. E. Redmann. 1993. Seed germination and seedling survival of C3 and C4
 776 grasses under water stress. Journal of Arid Environments 24: 277–285.

777 R Core Team. 2021. R: A language and environment for statistical computing. R Foundation for
 778 Statistical Computing, Vienna, Austria. <https://www.R-project.org>.

779 Reaumur, R.D., 1735. Observation du thermometer, faites à Paris pendant l'année 1735,
 780 compares avec celles qui ont été faites sous la ligne, à l'Isle de France, à Alger et en
 781 quelques-unes de nos isles de l'Amérique. Paris: Mémoires de l'Académie des Sciences.

782 Ren, S., X. Chen, W. Lang, and M. D. Schwartz. 2018. Climatic controls of the spatial patterns
 783 of vegetation phenology in midlatitude grasslands of the Northern Hemisphere. Journal of
 784 Geophysical Research: Biogeosciences 123: 2323–2336.

785 Ren, S., X. Chen, and C. Pan. 2022. Temperature-precipitation background affects spatial
 786 heterogeneity of spring phenology responses to climate change in northern grasslands
 787 (30°N-55°N). Agricultural and Forest Meteorology 315: 108816.

788 Richardson, A. D., K. Hufkens, T. Milliman, D. M. Aubrecht, M. Chen, J. M. Gray, M. R.
 789 Johnston, T. F. Keenan, S. T. Klosterman, M. Kosmala, E. K. Melaas, M. A. Friedl, and
 790 S. Frolking. 2018. Tracking vegetation phenology across diverse North American biomes
 791 using PhenoCam imagery. Scientific Data 5: 180028.

792 Richardson, A. D., T. F. Keenan, M. Migliavacca, Y. Ryu, O. Sonnentag, and M. Toomey. 2013.
 793 Climate change, phenology, and phenological control of vegetation feedbacks to the
 794 climate system. *Agricultural and Forest Meteorology* 169: 156–173.

795 Russell, M. L., L. T. Vermeire, A. C. Ganguli, and J. R. Hendrickson. 2017. Phenology of
 796 perennial, native grass, belowground axillary buds in the northern mixed-grass prairie.
 797 *American Journal of Botany* 104: 915–923.

798 Sala, O. E., and W. K. Lauenroth. 1982. Small rainfall events: An ecological role in semiarid
 799 regions. *Oecologia* 53: 301–304.

800 Sala, O. E., W. J. Parton, L. A. Joyce, and W. K. Lauenroth. 1988. Primary production of the
 801 central grassland region of the United States. *Ecology* 69: 40–45.

802 Seyednasrollah, B. 2018. Phenocamapi R Package: Interacting with the PhenoCam server.
 803 <https://doi.org/10.5281/zenodo.1464365>.

804 Springer, T. L. 2005. Germination and early seedling growth of chaffy-seeded grasses at
 805 negative water potentials. *Crop Science* 45: 2075–2080.

806 Tao, Z., W. Huang, and H. Wang. 2020. Soil moisture outweighs temperature for triggering the
 807 green-up date in temperate grasslands. *Theoretical and Applied Climatology* 140: 1093–
 808 1105.

809 Taylor, S. D., D. M. Browning, R. A. Baca, and F. Gao. 2021. Constraints and opportunities for
 810 detecting land surface phenology in drylands. *Journal of Remote Sensing* 2021: 9859103.

811 Thornton, P. E., R. Shrestha, M. Thornton, S. C. Kao, Y. Wei, and B. E. Wilson. 2021. Gridded
 812 daily weather data for North America with comprehensive uncertainty quantification.
 813 *Scientific Data* 8: 190.

814 Tian, S., A.I.J.M Van Dijk, P. Tregoning, and L.J. Renzullo. 2019. Forecasting dryland
815 vegetation condition months in advance through satellite data assimilation. *Nature*
816 *Communications*, 10:469.

817 Wang, C., E. R. Hunt, L. Zhang, and H. Guo. 2013. Phenology-assisted classification of C3 and
818 C4 grasses in the U.S. Great Plains and their climate dependency with MODIS time
819 series. *Remote Sensing of Environment* 138: 90–101.

820 Wang, G., M. Xiao, X. Xia, Y. Huang, Z. Luo, Y. Wei, and W. Zhang. 2022. Chilling
821 accumulation is not an effective predictor of vegetation green-up. *Geophysical Research*
822 *Letters* 49: e2021GL096558.

823 Wang, H., H. Liu, G. Cao, Z. Ma, Y. Li, F. Zhang, X. Zhao, X. Zhao, L. Jiang, N. J. Sanders, A.
824 T. Classen, and J. S. He. 2020. Alpine grassland plants grow earlier and faster but
825 biomass remains unchanged over 35 years of climate change. *Ecology Letters* 23: 701–
826 710.

827 White, M. A., P. E. Thornton, and S. W. Running. 1997. A continental phenology model for
828 monitoring vegetation responses to interannual climatic variability. *Global*
829 *Biogeochemical Cycles* 11: 217–234.

830 Xia, Y., J. Sheffield, M. B. Ek, J. Dong, N. Chaney, H. Wei, J. Meng, and E. F. Wood. 2014.
831 Evaluation of multi-model simulated soil moisture in NLDAS-2. *Journal of Hydrology*
832 512: 107–125.

833 Xiang, Y., S. Gubian, B. Suomela, and J. Hoeng. 2013. Generalized simulated annealing for
834 global optimization: The GenSA package. *R Journal* 5: 13–28.

835 Xin, Q., M. Broich, P. Zhu, and P. Gong. 2015. Modeling grassland spring onset across the
836 Western United States using climate variables and MODIS-derived phenology metrics.
837 Remote Sensing of Environment 161: 63–77.

838 Ye, J. S., J. F. Reynolds, F. T. Maestre, and F. M. Li. 2016. Hydrological and ecological
839 responses of ecosystems to extreme precipitation regimes: A test of empirical-based
840 hypotheses with an ecosystem model. Perspectives in Plant Ecology, Evolution and
841 Systematics 22: 36–46.

842 Zhang, G., Y. Zhang, J. Dong, and X. Xiao. 2013. Green-up dates in the Tibetan Plateau have
843 continuously advanced from 1982 to 2011. Proceedings of the National Academy of
844 Sciences of the United States of America 110: 4309–4314.

845 Zhang, L., H. Lei, H. Shen, Z. Cong, D. Yang, and T. Liu. 2019. Evaluating the representation of
846 vegetation phenology in the Community Land Model 4.5 in a temperate grassland.
847 Journal of Geophysical Research: Biogeosciences 124: 187–210.

848 Zhu, L., and J. Meng. 2015. Determining the relative importance of climatic drivers on spring
849 phenology in grassland ecosystems of semi-arid areas. International Journal of
850 Biometeorology 59: 237–248.

Model Type	Model Code	Drivers	# Param
Linear	LIN	T	2
Thermal time	TT (TTs)	T	3 (4)
Photo thermal time	PTT (PTTs)	T, L	3 (4)
M1 model	M1 (M1s)	T, L	4 (5)
Sequential	SQ (SQb)	T, C	8
Sequential M1	SM1 (SM1b)	T, C, L	8
Parallel	PA (Pab)	T, C	9
Parallel M1	PM1 (PM1b)	T, C, L	8
Alternating	AT	T, C	5
Unified M1	UM1	T, C, L	8
Grassland pollen model	GRP	T, L, P	5
Growing season index	SGSI, AGSI	T, L, V	7
Water time	WT (WTs)	P	3 (4)
Photo water time	PWT (PWTs)	P, L	3 (4)
M1 model	M1W (M1Ws)	P, L	4 (5)
Sequential	SQW (SQWs)	P, T	5 (6)
Sequential reverse	SQWr (SQWrS)	P, T	5 (6)
Parallel	PAW (PAWs)	P, T	6 (7)
Sequential variation	SQW_NoPbase	P, T	4
Sequential variation	SQW_cdd (SQWs_cdd)	P, T, cdd	5 (7)
Sequential variation	SQW_Tmin (SQWs_Tmin)	P, T, Tmin	5 (7)
Sequential variation	SQW_cdd_Tmin (SQWs_cdd_Tmin)	P, T, cdd, Tmin	6 (8)
Sequential variation	SQW_Pi_Tmin (SQWs_Pi_Tmin)	P, T, Tmin	5 (7)
Water time	WT_SM (WTs_SM)	SM	3 (4)
Photo water time	PWT_SM (PWTs_SM)	SM, L	3 (4)
M1 model	M1W_SM (M1Ws_SM)	SM, L	4 (5)
Sequential	SQW_SM (SQWs_SM)	SM, T	5 (6)
Sequential reverse	SQWr_SM (SQWrS_SM)	SM, T	5 (6)
Parallel	PAW_SM (PAWs_SM)	SM, T	6 (7)

Table 1. List of all 53 models included in the study along with their acronyms (Model Code), climate drivers (Drivers), and number of parameters (# Param). Models in parentheses are sigmoidal variations (lowercase “s” in Model Code). Black text represents the “original” models, blue text the new precipitation models, and green text the new soil moisture models. Drivers: T = temperature, P = precipitation, L = light (photoperiod), C = chilling, cdd = consecutive dry days, Tmin = minimum temperature, V = vapor pressure deficit, SM = soil moisture.

Model	AIC	RMSE
SQWs_Tmin	1094.691	15.975
SQWs_cdd_Tmin	1099.191	16.077
SQWs	1099.268	16.246
SQWs_cdd	1100.391	16.210
SQW_cdd_Tmin	1107.048	16.574
SQW_Tmin	1114.479	16.979
SQW_cdd	1137.597	18.016
SQW_NoPbase	1143.222	18.372
SQW	1145.316	18.376
SQW_Pi_Tmin	1148.591	18.531
SQWs_Pi_Tmin	1154.578	18.626
SQWrs	1189.103	20.455
SQWr	1205.958	21.468
PAWs	1219.343	21.991
PAW	1230.799	22.763
M1Ws	1277.534	25.793
PWTs	1277.819	25.944
WTs	1281.842	26.213
AGSI	1298.034	26.907
SGSI	1300.073	27.048
SM1b	1343.235	30.059
SM1	1347.969	30.426
PA	1366.084	31.710
PAb	1367.015	31.786
M1s	1391.955	34.587
TTs	1393.778	34.928
M1	1394.079	34.955
PTT	1394.255	35.150
PTTs	1395.109	35.047
TT	1395.690	35.280
UM1	1397.946	34.586
LIN	1400.853	35.934
AT	1400.895	35.389
PM1	1404.309	35.155
PM1b	1404.343	35.158
SQ	1405.646	35.276
SQb	1405.740	35.284
WT	1416.319	37.196
M1W	1418.319	37.196
PWT	1418.696	37.423
GRP	1451.200	40.261

Table 2. Model performance for all original and new precipitation models using All Sites, listed from best to worst performance (lowest – highest AIC). Black text represents the “original” models and blue text the new precipitation models. Results for the soil moisture models are listed in Table A.3.

Sites	t0	T-base	b	c	F-crit	P-req	T-thres	RMSE
All Sites	47	4.03	3.57	1.38	192.00	20.17	-10.14	15.98
B	80	6.68	13.24	0.44	59.20	22.59	-3.19	7.86
Cf	198	-3.62	125.63	10.26	134.42	0.86	-0.27	8.40
Cs	34	-8.65	1.06	4.29	1958.77	2.83	-17.61	16.35
Df	46	-0.58	0.92	0.45	321.51	77.12	-12.87	7.75

Table 3. The optimized parameters for the best model (SQWs_Tmin) run with All Sites and each of the four Köppen-Geiger (KG) climate groups, as well as the root mean square error (RMSE, in days) for each model fit. The text color matches Figures 2, 5, and 6 to distinguish the different KG groups. The parameters are as follows: t0 = the number of days after Sept 21 before precipitation can begin accumulating, T-base = the minimum average daily temperature (°C) for a day to be added to the temperature accumulation, b and c = values that determine the shape of the sigmoidal precipitation accumulation curve, F-crit = the required temperature summation (°C) for spring onset to occur, P-req = the required precipitation summation before temperature accumulation can begin, T-thres = the minimum temperature (°C) re-set threshold (temperature accumulation resets to zero if daily minimum temperature drops below this value).

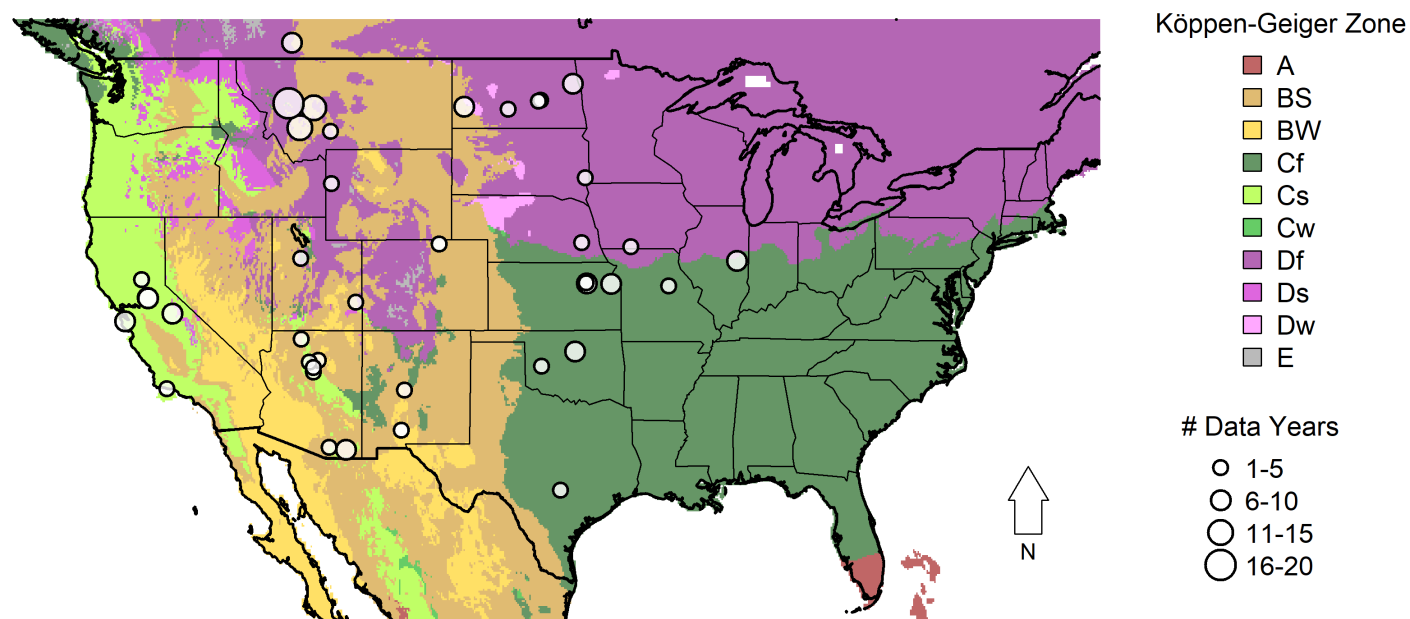


Fig 1. Locations of PhenoCam grassland sites included in the study. Marker sizes indicate the number of available data years. Background colors designate the 2-letter Köppen-Geiger zones used to divide the sites by climate.

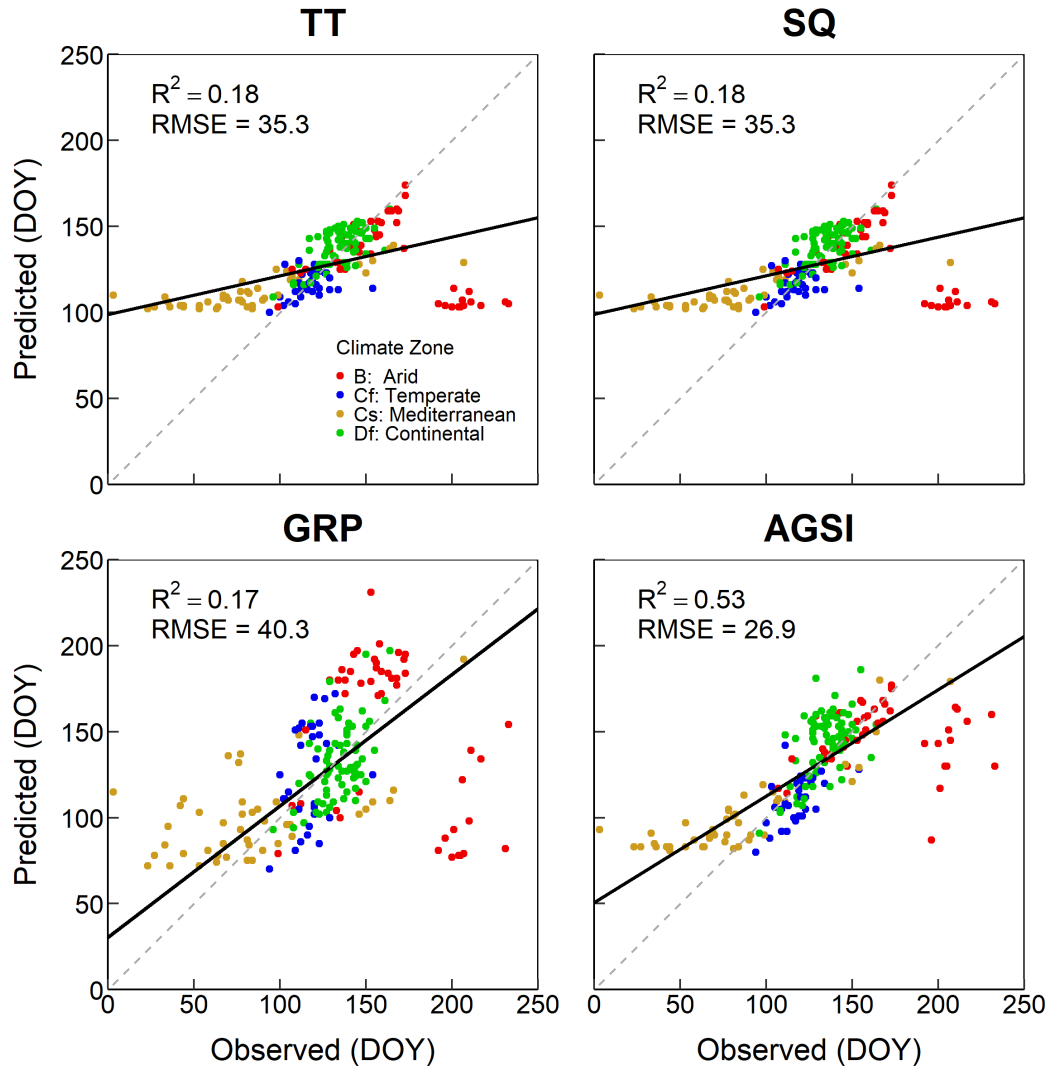


Fig 2. Scatterplots between predicted and observed spring onset dates (day of year, DOY) across all grassland sites for four “original” models. The R^2 value and root mean square error (RMSE, in days) are reported for each. Black lines are Model II regression fits, and grey dotted lines are 1:1 lines. Included models are Thermal Time (TT), Sequential (SQ), Grassland Pollen (GRP), and Accumulated Growing Season Index (AGSI). See Table 1 for model descriptions. Colors indicate the Köppen-Geiger climate zone for each data point.

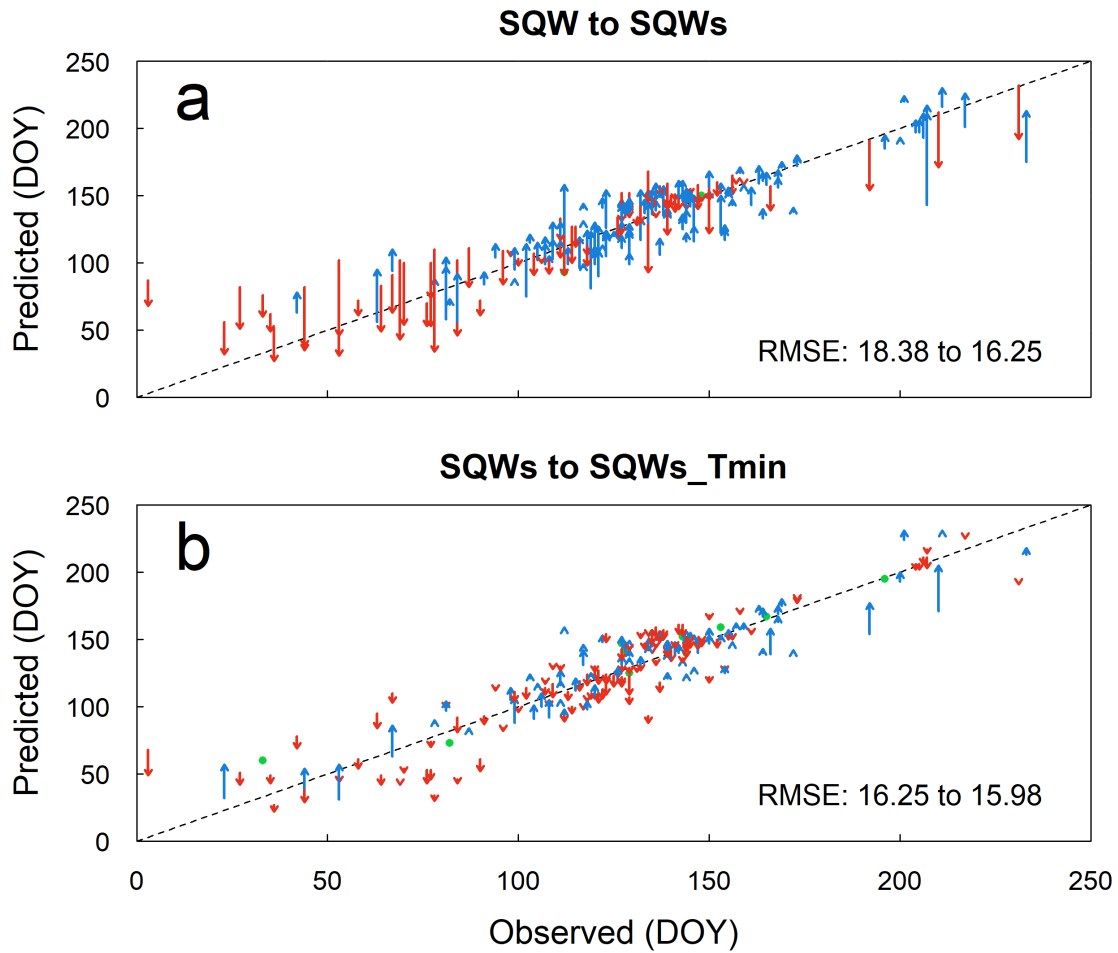


Fig 3. Scatterplots of predicted versus observed spring onset dates to compare model fits. The color and length of the arrows indicate the direction and magnitude, respectively, of shifts in predicted spring onset dates between different model structures. Red arrows indicate earlier dates, blue arrows later dates, and green points no change. The black dashed line is the 1:1 line. Panel a shows shifts in predicted dates between a model with a simple precipitation summation (SQW) and one with a sigmoidal precipitation accumulation structure (SQWs). Panel b shows how the addition of a minimum temperature reset parameter (Tmin) to SQWs influences predicted dates. The shift in RMSE (in days) between the models is included within each panel.

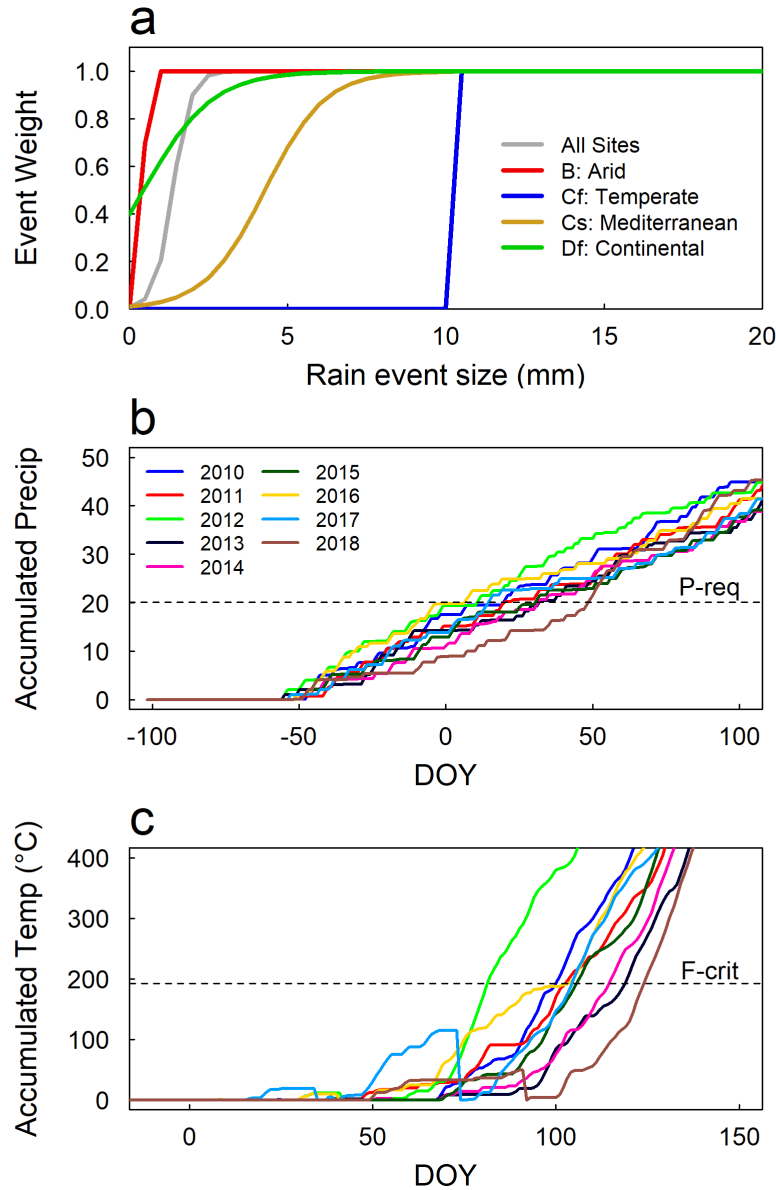


Fig 4. a) Sigmoidal precipitation accumulation curves (see Table A.2 for equation) for the best model (SQWs_Tmin) using the optimized parameters (see Table A.3) for All Sites and the four Köppen-Geiger climate groups (colored lines) over a range of precipitation event sizes. b) Example precipitation accumulation curve for a site in Illinois (uiefprairie) for SQWs_Tmin using the optimized parameters for All Sites. Colors indicate different years, and P-req (dashed line) is the required precipitation summation before temperature accumulation can begin. Precipitation is unitless since each event is assigned a weight (see panel a). c) Example temperature accumulation curve for the same model and site. Notice years 2017 and 2018 drop to zero (reset) because they exceeded the minimum temperature threshold. F-crit is the required temperature summation for spring onset to occur. Colors are the same as panel b.

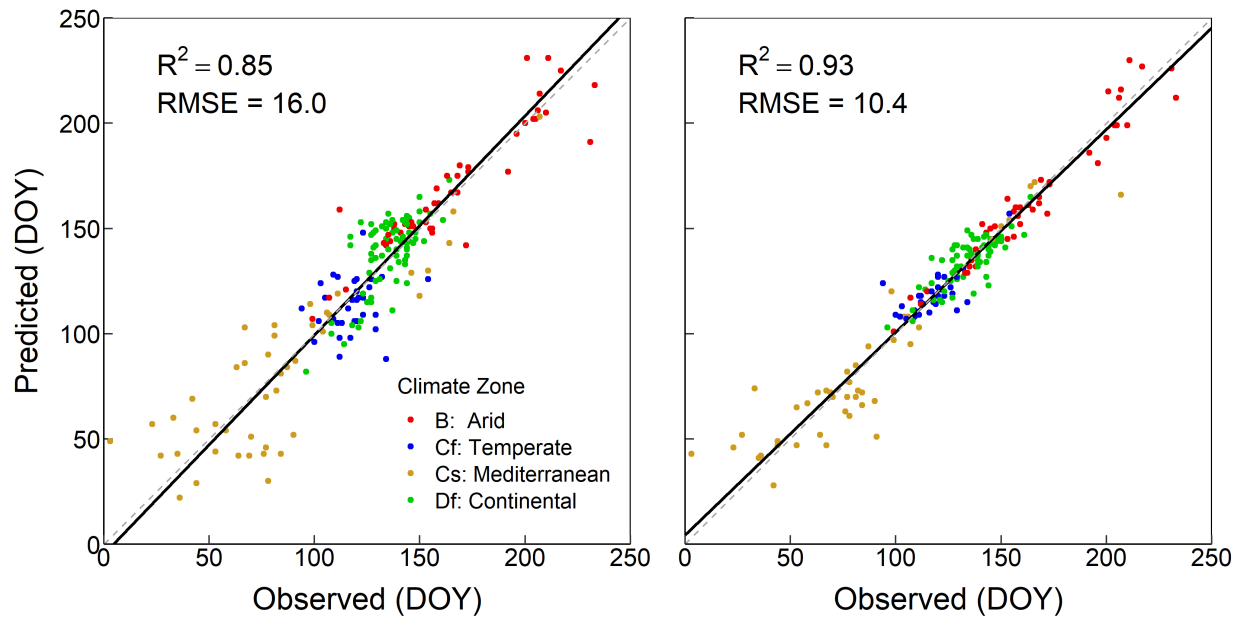
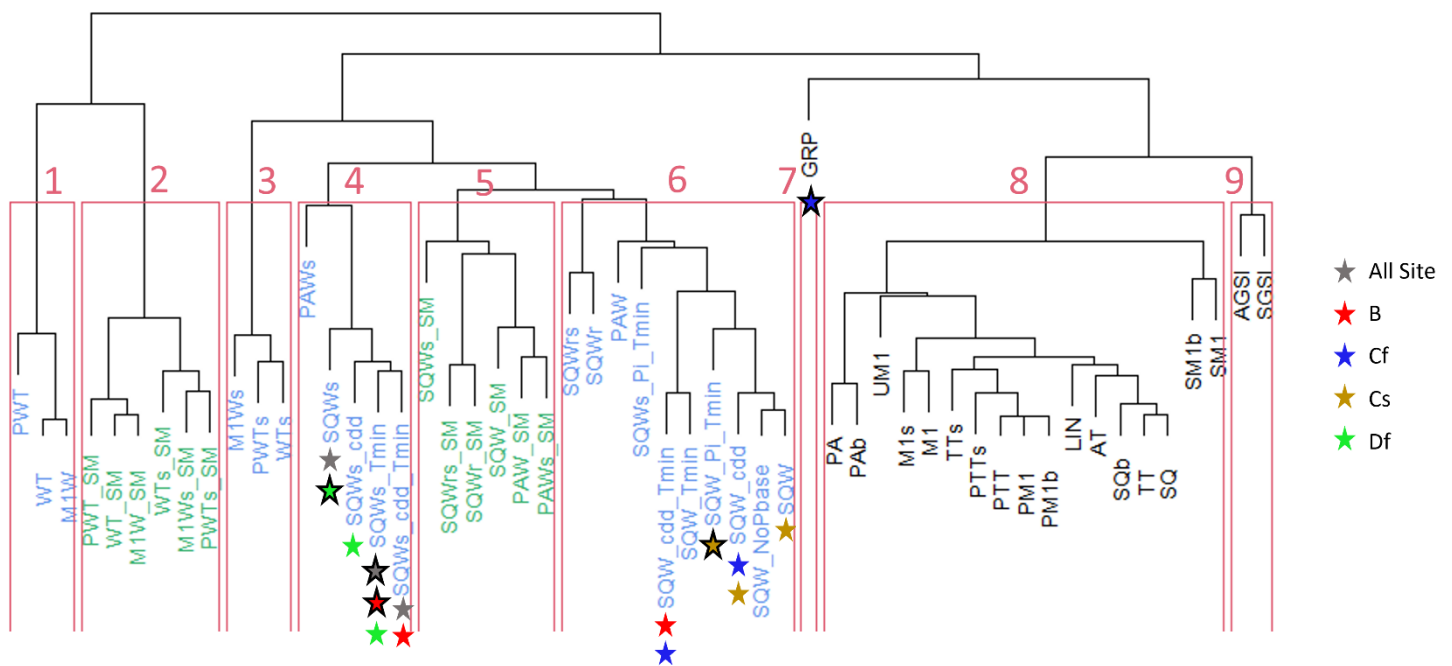


Fig 5. Scatterplots of predicted versus observed spring onset dates across all sites for the best new model (SQWs_Tmin) using either a single set of parameters (left) or separate regional parameter sets for each Köppen-Geiger climate group (right). The R^2 value and root mean square error (RMSE, in days) are reported for each. Black lines are Model II regression fits, and grey dotted lines are 1:1 lines. Colors indicate the Köppen-Geiger climate zone for each data point.



i. Dendrogram using model residuals (predicted-observed SOS dates) from models fit with All Sites. It depicts the similarity between model output – models with more similar predicted SOS dates are grouped closer together. Text color indicates model type: black = original model, blue = new precipitation model, green = soil moisture model. Colored stars designate the top 3 models for All Sites and each separate Köppen-Geiger group based on AIC; the best model for each is represented by a star with a black border. The salmon-colored boxes and numbers show model groupings for easier interpretation.

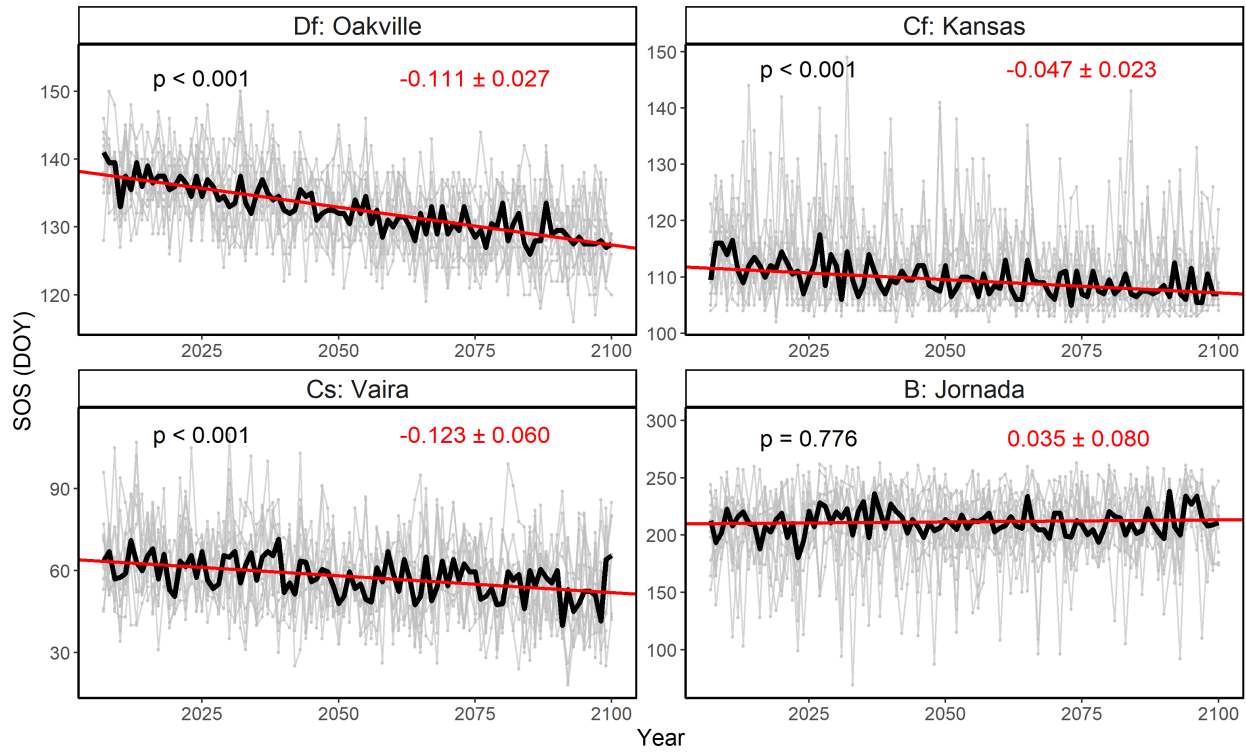


Fig 7. Predicted spring onset (SOS) dates for a representative site from each Köppen-Geiger (KG) group for the best model (SQWs_Tmin). The optimized model parameters for each site's respective KG climate zone were used to fit 12 future climate scenarios (RCP 8.5, 2007-2100). The predicted SOS dates from each scenario are shown by the grey lines. The black line represents the median, and the red line the average Sen's slope fit, across all 12 scenarios. The Sen's slope (mean $\pm$ sd) is listed in red text and the Mann-Kendall p-value in black text. The sites include oakville (North Dakota), kansas (Kansas), vaira (California), and Jornada ("ibp", New Mexico).