Outline: 7/14/17

Abstract

# Introduction

Dynamic Global Vegetation Models (DGVMs) are an important tool widely used for understanding ecosystem dynamics and predicting climate change impacts. Their ability to replicate spatial and temporal trends in ecosystem productivity has been demonstrated through comparison to remotely sensed vegetation data (e.g. MODIS, cite) and site-level measurements of annual flux derived from eddy covariance towers (cite). Several recent studies, however, have shown that DGVMs often fail to replicate seasonal patterns of leaf phenology and productivity (e.g. (Kucharik *et al.*, 2006; Richardson *et al.*, 2011; Schaefer *et al.*, 2012); these errors can also contribute to bias at annual or longer time scales (Richardson *et al.*, 2011). Leaf phenology, or the annual cycle of leaf growth and senescence, is closely linked to climate (Cleland *et al.*, 2007; Zhao *et al.*, 2013), and climate change is already beginning to drive shifts in the timing of these key events (Walther *et al.*, 2002; Parmesan & Yohe, 2003; Walther, 2003). Failure to capture current patterns of leaf phenology and their dependence on climate will affect our ability to predict how carbon fluxes may change in the future. ((Because the phenology of many plant species is linked to their spatial distribution (Chuine & Beaubien, 2001; Chuine, 2010; Chapman *et al.*, 2014) and competitive interactions (cite), errors in how process-based models represent leaf phenology may also hinder our ability to predict climate change impacts on the composition of vegetation communities, exacerbating biases in ecosystem-level patterns of carbon flux. (CUT?))

The ability of DGVMs to capture seasonal patterns of carbon flux can be limited by both problems with model parameterization (cite) and the structure of specific phenology routines (Kucharik *et al.*, 2006). Cite several lit examples of each problem and note the ecosystem- this is my lit review paragraph. two primary reasons for this problem. Parameter uncertainty in things like temperature response etc. Extra sentence of explanation and particular parameters identified by authors as problematic. Structural problems in how phenology is represented. Most models have simplistic representation of phen: summergreen or evergreen, with some also including raingreen.

These issues may be particularly troublesome in semi-arid ecosystems, where many species have developed unusual patterns of leaf phenology as an adaptation to seasonality in both temperature and moisture availability. One study in Austrailia found that xx% of woody plants were either brevi-deciduous, meaning they… or semi-deciduous, dropping only a portion of their leaves in response to moisture stress rather than cold. None of the phenology routines typical in DGVMs accurately represent these phenology types. Because semi-arid ecosystems are relatively under-studied relative to the temperate ecosystems for which DGVMs have been more thoroughly tested and developed, parameter uncertainy may also contribute to difficulties in estimating seasonal productivity. Universal parameters developed and tested in temperate ecosystems may not be appropriate for semi-arid species, and many PFT-specific parameters are poorly constrained by data. The relative paucity of field and modeling studies is compounded by the fact that few flux towers have been located in semi-arid ecosystems (cite, maybe get stats from ameriflux), leaving large gaps in our knowledge about how these ecosystems contribute to the global carbon cycle.

Here, we investigate how problems with both model parameterization and the structure of phenology routines affect modeled patterns of GPP and LAI in a semi-arid shrubland. Specifically, we ask: 1) How well do the existing methods of representing phenology in the LPJ-GUESS DGVM capture seasonal patterns of LAI and GPP? 2) How do parameter uncertainty and structural simplicity contribute to bias in seasonal and annual GPP? and 3) Can we improve the representation of seasonal patterns by a) optimizing existing parameters, or b) developing a new phenology routine? We address these questions using a combination of field and remotely-sensed data for four sites located in the sagebrush steppe, a widespread semi-arid ecosystem type that is of particular conservation concern due to its importance to a variety of different wildlife species (cite). By comparing model outputs to these independent data sets and using that data to inform further model development, we demonstrate how model-data assimilation can improve seasonal flux predictions.

# Methods

## Study area (DONE)

We focus our analysis on four eddy covariance tower sites located within the Reynold’s Creek Critical Zone Observatory in western Idaho, USA. These sites represent a large gradient in elevation, temperature, and precipitation (Figure S1), with mean annual temperature ranging from 9.2°C at the hottest site to 5.4°C at the coolest site, and mean annual precipitation ranging from 298-803mm. Productivity varies predictably along this gradient, with the coolest site having over twice the annual GPP of the hottest site (Flerchinger et al in review?). Vegetation consists of low-lying shrubs dominated by sagebrush (*Artemisia sp.*) mixed with perennial grasses and forbs. The dominant shrub species, however, varies across sites (Flerchinger et al in review?). From hot and dry to cooler and wetter, the sites are characterized by Wyoming big sagebrush (*Artemisia tridentata ssp. wyomingensis*), low sagebrush (*Artemisia arbuscula*), recently-burned (2007) mountain big sagebrush (*Artemisia tridentata ssp. vaseyana*), and undisturbed mountain big sagebrush. All of these *Artemisia* species are semi-deciduous and grow two sets of leaves each year: ephemeral leaves, which are shed by mid-summer, and persistent leaves, which remain on the plant throughout the winter.

## Model description

The LPJ-GUESS DGVM (Smith *et al.*, 2001) simulates establishment, growth, competition, and mortality for any number of plant functional types (PFTs) defined by a set of physiological parameters. For this study we used two PFTs: a generic C3 grass/forb and a generic shrub. PFT parameters determine patterns of carbon acquisition and allocation, and consequently affect simulated carbon fluxes at the ecosystem level. Carbon moves from vegetation through several soil carbon pools, and is released back to the atmosphere by respiration and decomposition. LPJ-GUESS can also simulate carbon fluxes due to fire, but we ran the model with this option turned off. Initial parameter values for both PFTs follow Smith *et al.* (2001).

(LPJ-GUESS runs on a daily time step and is driven by data on temperature, precipitation, solar radiation, soil texture, and CO2 concentration. We used Daymet data (Thornton *et al.*, 2017) for the climate variables, STATSGO data (Miller & White, 1998) for soil texture, and CO2 concentration data developed for use with the TRENDY project (<http://dgvm.ceh.ac.uk/node/9/index.html>). We ran the model at 1-km resolution. All model runs began with a 1000-year spin-up to allow all carbon pools to reach equilibrium. The spin-up period was driven by de-trended Daymet data and the pre-industrial CO2 concentration.) Appendix?

Woody plant leaf phenology is represented by three broad phenology types: evergreen, raingreen, and summergreen. Evergreen PFTs maintain a constant leaf area throughout the year, but productivity is limited in colder months when temperatures fall below the minimum temperature for photosynthesis. Raingreen PFTs shed their leaves when available soil water drops below a PFT-specific threshold, but re-grow them once water stress is reduced. Summergreen PFTs produce leaves each spring once temperatures reach 5°C, then shed those leaves each fall. Winter temperatures must be low enough to satisfy a chilling requirement in order for budburst to occur in the spring. PFT-specific parameters determine the budburst chilling requirement, timing of budburst, and length of time from budburst to full leaf cover. Leaf senescence for summergreen PFTs occurs when either the total number of days with full leaf cover exceeds 210, the maximum for all summergreen PFTs, or the temperature drops below 5°C. This typically causes leaf senescence to occur in October, which while appropriate for many deciduous trees is several months too late for sagebrush.

Grass phenology can be set to summergreen, raingreen, or a hybrid that follows spring cues for leaf-out but allows for leaf senescence in response to drought. We used the summergreen type for all simulations because the raingreen phenology routine caused grass to grow a second cohort of leaves in the fall once soil moisture increased, and this wouldn’t happen in the sagebrush steppe. Summergreen grasses start to leaf out when the daily temperature exceeds 5°C, and as with woody PFTs, the length of time to full leaf cover is determined by a PFT-specific parameter. Grasses have no maximum number of leaf-on days, and instead senesce only once the temperature drops below 5°C. (Grasses and forbs account for 38-75% of LAI across the four sites (cite Gerald personal communication? or published somewhere?), so improving our representation of seasonal patterns required making improvements to grass phenology as well. cut?)

## Data sources

### Vegetation

Data on leaf area index (LAI) and percent cover for each functional type was measured in 2016 at three of our study sites as part of a different project (cite Pat Controy). P1: Description of vegetation data provided by Pat Conroy.

*(P2: Describe MODIS: brief description of how LAI is modeled based on remote sensing data, format/time step, and how/when it was acquired (downloaded from xxx on xx date).*

*Field estimates of LAI closely matched the MODIS estimates for the same dates (Fig. in appendix), so MODIS LAI was used for all analyses in order to have a complete monthly time series for water years 2015-2016. Bi-weekly values for LAI were converted to monthly by interpolating the point estimates to a daily time step then taking the average LAI for each month.)* Will cut all of this if I decide not to optimize based on monthly LAI.

### Gross Ecosystem Production

p1: eddy covariance- describe data and processing and cite others

p2: remote sensing and model: also compared site-level flux to GPP from MODIS. Describe MERRA2 and site runs.

## Sensitivity analysis (DONE)

We performed a sensitivity analysis to determine which parameters had the largest influence on seasonal and annual patterns of ecosystem production and narrow down the set of parameters optimized (Fig. 1). We began with a set of fourteen parameters. Eight of these were chosen because they were known to have a large influence on total carbon fluxes based on previous studies (Zaehle *et al.*, 2005; Wramneby *et al.*, 2008; Poulter *et al.*, 2010) or the appropriate value for sagebrush was uncertain due to a lack of published field measurements. The six additional parameters were included because they are directly related to how phenology is represented in LPJ-GUESS.

We used Latin hypercube (LHC) sampling (McKay *et al.*, 1979), a stratified random sampling approach, to generate test values for each of the parameters. LHC sampling provides an efficient way to sample the full parameter space by ensuring that test values are evenly distributed throughout the multi-dimensional hypervolume defined by the distributions of potential parameter values. We assumed a uniform distribution for each parameter, bounded by minimum and maximum potential values derived where possible from literature estimates (Table 1). We generated 640 unique sets of parameter values used to drive model runs for each of the four flux sites.

Sensitivity was assessed by examining how each parameter affected the 30-year mean (1986-2015) for several output variables. We considered four annual variables, two at the ecosystem level (NEE and GPP) and two at the PFT level (sagebrush foliar projective cover (FPC) and leaf area index (LAI). We also assessed how each parameter affected monthly GPP and LAI as well as the proportion of annual GPP that occurred in each season (spring, summer, and fall). For each parameter, we calculated the ranked partial correlation coefficient (RPCC) between the parameter and the value of each output variable. This was done separately for each of the four sites. We used |RPCC| > 0.2 for at least one output variable as the cut-off for including any given parameter in the next stage of analysis.

## Phenology models (DONE)

We compared three different methods of representing phenology in LPJ-GUESS. First, we ran the model using standard parameter values for the shrub and grass functional types (Smith *et al.*, 2001) to determine if the original model could replicate seasonal patterns of GPP and LAI in a semi-arid shrubland. We did this for each of the three phenology options available in LPJ-GUESS. Second, we optimized parameters that the sensitivity analysis identified as having a large effect on GPP and LAI to determine if parameter values specific to sagebrush could overcome any shortcomings in the original model. Third, we developed a new phenology type for semi-arid shrubs to determine if structural model changes would yield further improvements. We call this type “semi-deciduous” after Williams *et al.* (1997).

In the semi-deciduous phenology type, sagebrush develops two sets of leaves each spring: one that is shed by mid summer (ephemeral leaves), and one that persists throughout the winter (persistent leaves). The start and duration of leaf-out for both persistent and ephemeral leaves are controlled by the same equation and parameters that are used in the summergreen phenology type (Appendix with all phenology equations). To allow ephemeral leaves to senesce mid-summer, we created two new model parameters: aphen\_max and downramp. Aphen\_max allows the maximum number of days with full leaf cover to be set independently for semi-deciduous species, instead of defaulting to 210 days. Downramp represents the percentage of leaves lost each day once aphen\_max is exceeded, so that semi-deciduous species can lose their ephemeral leaves gradually instead of all at once. These two parameters were also be added to the grass PFT. A third new parameter for the shrub PFT, phen\_winter, determines the proportion of maximum leaf cover consisting of persistent leaves.

## Model-Data Assimilation

We compared three different data sources for optimization: monthly GPP, monthly LAI, or both. GPP is a more sensitive indicator of ecosystem processes and is commonly used for parameter optimization (cite, cite), but field or remotely-sensed LAI data is also used (cite) because it is easy to acquire and does not require and specialized equipment. If LAI performs adequately for parameter optimization, it could potentially be used to parameterize a wider range of species in areas where flux data is not readily available.

We optimized all model parameters simultaneously by minimizing the weighted sum of squared residuals, Σiwi2ri2, where ri represents the residual between model output in month i and either the flux tower estimate of GPP or MODIS estimate of LAI. Weights were determined such that GPP and LAI were weighted equally in the optimization runs using both data types despite the difference in units. This cost function was minimized using the differential evolution algorithm (Storn & Price, 1997) implemented in the DEoptim R package (Mullen *et al.*, 2011). Differential evolution is a global optimization routine appropriate for optimizing the parameters of a complex non-linear model like LPJ-GUESS where there are likely to be a number of local optima (Price *et al.*, 2005). Parameter values can be constrained by a user-specified minimum and maximum; we used the same upper and lower parameter limits that we applied in the sensitivity analysis (Table 1).

(At each iteration in the differential evolution algorithm we ran LPJ-GUESS for the full 1000-yr spinup followed by a period driven by historic climate data and extracted output for water years 2015 and 2016. We ran the model with just one patch to speed computation. Only parameters with an |RPCC| > .2 were optimized. For the semi-deciduous parameter optimization, we set parameters that were optimized for the summergreen model but failed to meet this criteria to their optimized value rather than the standard. – cut?)

## Model evaluation

Once optimal parameters were identified, the summergreen and semi-deciduous shrub models were each re-run using 100 patches. We assessed the ability of each model to capture seasonal patterns of GPP and LAI using several metrics. For both GPP and LAI, we looked at the SSR as a measure of absolute error and R2 as a measure of how well the model matched the monthly pattern. For GPP, we also evaluated how well each model replicated three indicators of seasonality: the month where GPP first exceeds 20% of the maximum, month of maximum GPP, and month where GPP first drops below 20% of the maximum. For each of the optimized models we calculated the percent improvement in SSR for GPP and LAI over the standard parameterization of LPJ-GUESS.

We tested the hypothesis that improvements in sagebrush phenology would affect estimates of total productivity and community composition by examining the impact of each model formulation on annual GPP and the ratio of shrub cover to grass and forb cover. This was done individually for each of our four study sites.

# Results

## Variability in seasonal patterns between sites

* briefly decide how seasonality differed if this isn’t stepping on the toes of anyone at RC
  + earlier green-up and lower peak GPP at warmer sites
  + refer to table with months of 20%, peak, 20%

## Seasonal patterns estimated with LPJ-GUESS and MODIS

Model runs using the standard parameterization with each of the three phenology types revealed substantial differences when compared to the seasonal patterns of GPP from the flux towers and LAI from MODIS (Fig. 1). LPJ-GUESS tended to over-estimate GPP in the fall regardless of which phenology type was used, and the evergreen and summergreen types also over-estimated spring GPP at the two higher-elevation sites. The summergreen phenology type tended to under-estimate spring GPP. All three phenology types resulted in a large over-estimation of LAI during the summer season. Model runs with the summergreen phenology type compared most favorably to the data (lowest root mean squared error for both GPP and LAI), so we used this phenology type for the sensitivity analysis.

Estimates of GPP from modis also differed substantially from the flux measurements. Modis web product using MERRA2 was low. Using site-specific climate data improved GPP estimates…

Paragraph about how total carbon flux from that 2-year period differed among models- did phenology biases affect estimates of total carbon uptake? Analysis currently incomplete.

## Parameter sensitivity

Of the fifteen variables examined in the sensitivity analysis, eight were found to have a substantial influence (|RPCC| > .2) on annual GPP, LAI, or FPC (Table 1). Seven of these parameters were also closely linked to seasonal patterns in relative GPP, and the order of importance was identical to that for the annual variables (Table 2). One additional variable (pstemp\_low) also had a strong influence on monthly GPP during certain seasons. Six parameters (gmin, greff\_min, pstemp\_min, kchill\_b, GDD5, and leaf longevity) did not have a substantial influence in any of our tests and were excluded from further analyses.

In the seasonal analysis, the sign of the relationship between each parameter and the proportion of annual GPP varied according to season and in some cases across sites. Across all sites, higher values of SLA, ltor\_max, and pstemp\_max were associated with a greater proportion of GPP occurring during the fall. Conversely, higher values of root\_up and pstemp\_high were associated with more productivity in the spring (discussion: makes sense. Can take advantage of water then water is not limiting). Higher values of latosa lead to a lower proportion of GPP occurring during the summer.

Of the twelve parameters considered in our sensitivity analysis using the new phenology routine, ten had a |RPCC| > .2 for at least one annual, seasonal, or monthly variable. The order of variable importance was similar to that from the initial sensitivity analysis. SLA… continued to.

## Parameter estimates

* for parameters estimated in standard model vs. semi-deciduous, how different are the optimal values?
* (if time: optimize for each site separately and see how parameter values compare).

## Model performance

* compare how 5 models do at matching pattern of monthly GPP: evergreen, raingreen, summergreen, optimized summergreen, optimized semi-deciduous. Focus on R2.
* now focus just on 3 models: summergreen, optimized summergreen, and optimized semi-deciduous.
  + How did they do at hitting key seasonality metrics: month where GPP first hits 20% of peak (spring), month of peak GPP (summer), last month before GPP drops below 20% of peak.
  + does model capture differences in phenology between the sites?
  + does model capture differences in phenology in the two years?

## Impact on model predictions

* For the three primary models:
  + compare total annual GPP in 2015 (add 2016? water year so only through Oct).
    - NewPhen is the ONLY model that gets the site-site order of productivity correct!
    - NewPhen is the WORST at predicting the annual total of GPP.
  + compare total, site pattern, and year-year difference when optimizing just based on GPP
  + compare predicted FPC and LAI: do models get the correct range? do they predict the correct ration of shrub:grass? Does optimizing based on GPP help with this prediction?

# Discussion

## Maybe start with a paragraph highlighting uncertainties in semi-arid systems, what they are, and where we are in addressing these…

## Parameter sensitivity

* note differences between the parameters that we identified as important and those that emerged as important in pervious sensitivity analyses where LPJ-GUESS was used.
* note that the model was very sensitive to two out of the three new phenology parameters that we added. Original model sensitive to phengdd5ramp. Both models sensitive to three of four photosynthesis temperature parameters. All of these are things for which we don’t have good data, so using flux data to estimate can be helpful.

Not sure where to put this, but I’d like to talk about how the optimal parameter values differed between runs. Based on preliminary results it seems like LAI is much higher in the summergreen model (higher than field estimates), and the model might be using that to compensate for issues with phenology, which could cause issues with prediction at different sites or in the future.

## Importance of phenology

* estimates of overall GPP (annual) improved with new phenology routine (I hope!)
* estimates of competitive balance improved (again, I hope ☺)
* big implications for prediction: errors in total flux multiplied when we predict over large spatial extents of long time scales. Also need to know how changing climate may alter leaf phenology and be able to predict that accurately

Comparison/testing of hypotheses from other papers:

(Schaefer *et al.*, 2012): improving ps temp parameters/process will reduce over-estimation of GPP in spring and fall. Optim1 does this: pstemp\_min actually goes down a bit, pstemp\_lo is similar or slight increase (depends which optim routine), phen5 ramp changed a lot. Interestingly though, sensitivity analysis revealed that other parameters (Sla, root\_up, ltor) had bigger impact on spring vs. summer. Root\_up makes sense- need deep roots in summer, want less shallow roots in spring when water abundant.

## The problem with resolution

* modis GPP from MERRA2 is clearly way off- improved when used local climate data
* some of the remaining discrepancies between GPP predicted by LPJ-GUESS with the semi-deciduous phenology type and GPP measured at flux sites can also be attributed to issues with resolution- both spatial and temporal.
  + Soil and climate data is 1-km resolution, in reality the environment is much more heterogeneous. The MBS and burn sites are both in areas that benefit from uphill snow drifts that reduce water stress
  + Model output is at monthly resolution. … help with issues this causes?

## Additional model development

* our work has shown the potential to improve carbon flux estimates by improving phenology, but there is still room for improvement.
  + issue with over-estimate in fall… grass?.
  + flux data doesn’t provide info on GPP at the species level, so hard to improve one PFT if the other is still causing problems- and we know that grass does (cite PHACE paper).
  + winter GPP stays low even when phenwinter is high. Why?
  + still have an issue where if the peak is correct the spring and fall are too high- this issue was evident in the modis GPP as well. Why? seems an issue with most veg models (same in Richardson paper).
  + need to better capture year-year variability, as well as potential for adaptation. We know that the optimal temp for photosynthesis in sage shifts higher throughout the summer, so whole range could shift higher as the climate warms- it has that potential.
  + would be better if mechanistic- what cue causes greenup and browndown? unfortunately this is not well understood (cite) and may vary by sp (cite), so our more generalized approach may be the best we can do right now.
  + Despite using generalized parameters developed for temperate ecosystems, the original model provided relatively good estimates of annual GPP. This was possible due to the tendency to under-estimate spring fluxes and over-estimate fall fluxes, causing the annual total to balance out. This nonetheless masks a significant problem in capturing seasonal patterns, and also results in an unrealistically high LAI.
  + we over-estimate GPP in dry conditions- matches finding of (Schaefer *et al.*, 2012)
  + we under-estimate GPP peak- supports conclusion of Schaefer et al. 2012 that we need to fix LUE of vcmax etc. This is a much more fundamental issue but relates to parameters rather than processes.
  + A comparison of the modeled patterns of GPP and LAI suggests that the excess fall productivity noted in the original model can be attributed largely to the maintenance of high leaf area throughout the late summer and early fall.
  + Our under-estimation of GPP at the two coolest sites may be due to several factors. First, the two high-elev sites in locally productive microsites. Second, low productivity and the driest site necessitated parameters that resulted in an overall reduction in productivity. This may relate to known issues with soil depth and moisture (Cite) that can cause the model to underestimate the importance of moisture limitation. This issue is also reflected in how poorly the model fits the more complex pattern of GPP at the Wyoming sagebrush site. Finally, the over-estimation of LAI and under-estimation of GPP point to an issue with respiration or light use efficiency. The improvements in seasonal patterns allowed a much better correlation with data, but to accurately match the magnitude would require further model developments that address underlying problems unrelated to phenology.

# Conclusions

* phenology matters, and can be improved in DGVM via parameter optimization or even more so with some simple structural adjustments that allow the model to reflect the two leaf cohorts common in many semi-arid shrubs.

Random Notes-

Ways to get total GPP:

* remote sensing + model: but MODIS is way off
* remote sensing + model + local climate data: better, but requires local met data
  + Ultimately, both these methods are only useful for current time period (b/c need remote sensing data). To understand implications of climate change, we need something else.
* DGVM: has potential, but issues with phenology
* Why? most DGVMs have very simple representation: in LPJ-GUESS, it is 3 different types: describe
* fit isn’t that bad, but none of them do a great job of matching seasonal patterns
  + show R2 for each
  + show total annual GPP for each