

Supporting Information Appendix

MODIS phenology product

The MODIS Land Cover Dynamics (MCD12Q2) Product (informally called the MODIS Global Vegetation Phenology Product) provides estimates of continuous variation in vegetation phenology at global scales. The raw data are satellite derived, Enhanced Vegetation Index (EVI) data from MODIS satellite surface-reflectance data obtained at 500m spatial resolution in each pixel of the earth globally. These data are summarized as 8-day composite EVI reflectance for each MODIS pixel as part of the MODIS Global Vegetation Phenology Product. Using the methodology developed by Zhang and colleagues (1) four key transition dates are estimated from the maximum rates of change in the curvature in EVI over time; these define the key phenological phases of vegetation dynamics at the landscape scale and annual time scales on the earth's surface: green-up (greenness starts to increase rapidly), maturity (greenness reaches maximum value), senescence (greenness starts to decrease rapidly), and dormancy (greenness reaches minimum value). See more details in the User Guide of this data product, http://www.bu.edu/lcsc/files/2012/08/MCD12Q2_UserGuide.pdf.

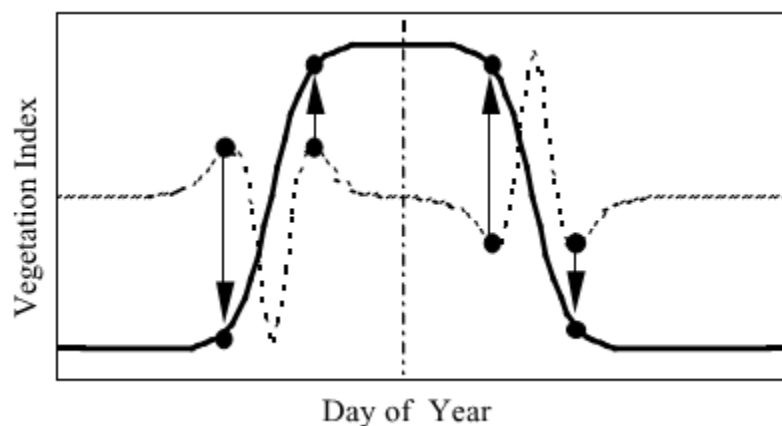


Figure S1 (Fig. 2 from ref. 1) A schematic showing how transition dates are calculated using minimum and maximum values in the rate of change in curvature. The solid line is an idealized time series of vegetation index data, and the dashed line is the rate of change in curvature from the VI data. The circles indicate transition dates. The extreme values located between each circle indicate the point at which the rate of change in curvature changes sign.

Dormancy dates derived from MODIS product are proxies of the timing of plant development stage (i.e. phenology) termed dormancy. Through comparisons among ground-based phenology observation, time series of color index from camera images, and remote sensing phenology (shown in Figure S2), dormancy dates indicate full plant dormancy, in which the greenness reaches the minimum values, shown as fully changed leaf colors to brown and leaf drop from trees. Besides Zhang's papers (1-4), many other published papers using remote sensing phenology data (also called as Land Surface Phenology) (5-6) also use the term 'dormancy' to define the final autumn phenological phase as derived from satellite data. This also corresponds to autumn dormancy in wood plants as defined by plant physiologists (8-10).

Table S1 Unstandardized coefficients of variables in eight models on deciduous forest fall dormancy dates in Northeastern Highlands from 2001 to 2010. AIC, BIC, and RMSE were calculated for 2001-2010 and 2011-2012, and yielded consistent results, so AIC, BIC, and RMSE only for 2011-2012 are shown. Smallest AIC, BIC, and RMSE are bold.

Variables	MLR	Ridge	Bayesian LASSO	Elastic Net	PACS	Spike&Slab	BMA	Posterior Median Model
Latitude	-4.206	-4.208	-4.215	-4.232	-4.211	-4.322	-4.238	-4.255
Elevation	-0.007	-0.007	-0.007	-0.007	-0.007	-0.006	-0.007	-0.007
CDD20 _(Aug.1-Nov.15)	-0.022	-0.022	-0.022	-0.022	-0.022	-0.027	-0.022	-0.022
FD _(Sep.1-Nov.15)	-0.102	-0.100	-0.073	-0.030	-0.056	0	-0.024	0
FD _(Sep.1-Nov.15) ²	0.006	0.006	0.006	0.005	0.005	0.004	0.0004	0.004
FD _(Apr.1-Jun.30)	-0.052	-0.052	-0.052	-0.051	-0.052	0	-0.057	-0.059
HD32 _(Jul.1-Aug.31)	1.124	1.113	1.085	0.912	1.025	1.250	0.111	1.111
HD35 _(Jul.1-Aug.31)	-0.833	-0.830	-0.799	-0.762	-0.732	0	-0.824	-0.821
GDR _(May.1-Jun.30)	-0.058	-0.058	-0.055	-0.061	-0.045	0	-0.022	0
GDR _(Jul.1-Aug.31)	0.092	0.092	0.088	0.088	0.083	0	0.087	0.088
GDR _(Sep.1-Nov.15)	0.746	0.745	0.745	0.735	0.751	0.740	0.075	0.755
GDR _(Sep.1-Nov.15) ²	-0.028	-0.028	-0.028	-0.027	-0.028	-0.027	-0.029	-0.029
RD _(May.1-Jun.30)	-0.207	-0.207	-0.206	-0.208	-0.199	-0.177	-0.197	-0.192
RD _(Jul.1-Aug.31)	0.596	0.584	0.557	0.369	0.558	0.574	0.582	0.579
RD _(Jul.1-Aug.31) ²	-0.012	-0.012	-0.011	-0.007	-0.011	-0.011	-0.012	-0.012
RD _(Sep.1-Nov.15)	-0.065	-0.065	-0.062	-0.058	-0.062	0	-0.066	-0.067
ECA _(May.1-Jun.30)	-0.203	-0.203	-0.202	-0.208	-0.207	0	-0.204	-0.206
ECA _(Jul.1-Aug.31)	0.130	0.130	0.130	0.138	0.125	0	0.126	0.124
ECA _(Sep.1-Nov.15)	0.274	0.274	0.272	0.269	0.268	0	0.281	0.286
HD32 _(Jul.1-Aug.31)	-0.057	-0.057	-0.055	-0.044	-0.051	-0.076	-0.057	-0.058
*RD _(Jul.1-Aug.31)								
AIC(2011-2012)	49838	49841	49839	49886	49856	50026	51741	49820
BIC(2011-2012)	49981	49983	49983	50029	49999	50105	51884	49949
RMSE(2011-2012)	14.234	14.236	14.235	14.270	14.247	14.391	15.754	14.223

Table S2 Unstandardized coefficients of variables in eight models on deciduous forest fall dormancy dates in Northeastern Coastal Zone from 2001 to 2010. AIC, BIC, and RMSE were calculated for 2001-2010 and 2011-2012, and yielded consistent results, so AIC, BIC, and RMSE only for 2011-2012 are shown. Smallest AIC, BIC, and RMSE are bold.

Variables	MLR	Ridge	Bayesian LASSO	Elastic Net	PACS	Spike&Slab	BMA	Posterior Median Model
Latitude	-3.955	-3.956	-3.951	-3.925	-3.999	-4.054	-3.956	-3.953
Elevation	-0.017	-0.017	-0.017	-0.017	-0.017	-0.018	-0.017	-0.017
CDD ₂₀ (Aug.1-Nov.15)	-0.029	-0.029	-0.029	-0.029	-0.028	-0.027	-0.029	-0.029
FD _(Sep.1-Nov.15)	-0.269	-0.268	-0.267	-0.266	-0.269	-0.274	-0.268	-0.269
HD ₃₂ (Jul.1-Aug.31)	1.611	1.608	1.614	1.586	1.617	1.646	1.612	1.612
HD ₃₅ (Jul.1-Aug.31)	-1.706	-1.706	-1.708	-1.725	-1.720	-1.748	-1.706	-1.706
GDR _(May.1-Jun.30)	-0.079	-0.079	-0.079	-0.089	-0.084	0	-0.079	-0.078
GDR _(Jul.1-Aug.31)	0.050	0.050	0.050	0.055	0.048	0	0.049	0.050
GDR _(Sep.1-Nov.15)	0.111	0.111	0.111	0.112	0.111	0.103	0.111	0.111
RD _(Jul.1-Aug.31)	-0.099	-0.100	-0.099	-0.108	-0.101	-0.100	-0.099	-0.099
RD _(Sep.1-Nov.15)	0.743	0.737	0.722	0.529	0.587	1.002	0.744	0.742
RD _(Sep.1-Nov.15) ²	-0.019	-0.019	-0.019	-0.014	-0.015	-0.026	-0.019	-0.019
ECA _(May.1-Jun.30)	-0.243	-0.243	-0.242	-0.256	-0.255	-0.268	-0.243	-0.242
ECA _(Jul.1-Aug.31)	0.372	0.373	0.373	0.383	0.375	0.374	0.372	0.372
ECA _(Sep.1-Nov.15)	-0.795	-0.793	-0.779	-0.725	-0.768	-0.752	-0.795	-0.793
ECA _(Sep.1-Nov.15) ²	0.083	0.082	0.081	0.076	0.080	0.078	0.008	0.082
FD _(Sep.1-Nov.15) *elevation	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002
HD ₃₂ (Jul.1-Aug.31)								
*RD _(Jul.1-Aug.31)	-0.093	-0.092	-0.093	-0.090	-0.093	-0.095	-0.093	-0.093
AIC(2011-2012)	36904	36939	37198	36939	36927	36958	39493	36907
BIC(2011-2012)	37034	37068	37327	37068	37057	37073	39622	37036
RMSE(2011-2012)	6.735	6.747	6.838	6.747	6.743	6.755	7.700	6.736

Table S3 Mean relative biomass of top eight genera of woody plant in the two eco-regions of the study area. Values of mean relative biomass were calculated using data of 1103 plots from the Forest Inventory and Analysis (FIA) Program (<http://fia.fs.fed.us/>) during 2003-2010 for our spatial domain. Relative biomass of each genera of woody plant was calculated in each plot and then averaged for the two eco-regions.

Genus	Northeastern Highlands	Northeastern Coastal Zone
<i>Acer</i>	0.321	0.253
<i>Betula</i>	0.128	0.090
<i>Carya</i>	0.004	0.048
<i>Fagus</i>	0.034	0.013
<i>Fraxinus</i>	0.045	0.037
<i>Populus</i>	0.027	0.013
<i>Prunus</i>	0.021	0.026
<i>Quercus</i>	0.092	0.353

Figure S2 Growing season phenology of forest canopy trees at 3 spatial scales from observations in 2014: a) Observed percentage of leaf unfolding through to leaf drop among individuals of 6 deciduous forest species at one site (25m×50m) in one growing season (data from Xie); b) Time series of green color index extracted from digital time lapse cameras of the same site and two individual trees therein (11); c) Enhanced Vegetation Index (EVI) from MODIS (M*D13Q1) showing the 2014 canopy level phenology for the same location (250m resolution) as a & b (black points with spline curve) and 2000-2014 inter-annual variability (grey points and splines). These figures show the full annual phenological cycle for forest canopy trees (individuals to stands) at one site in a northeastern North American forest tract from leaf bud-break beginning around day 120 to fully expanded leaves around day 150, and leaf senescence beginning around day 250 and extending to leaf drop and dormancy by around day 300; compare these with idealized, yearly phenological schema shown in Figure S1.

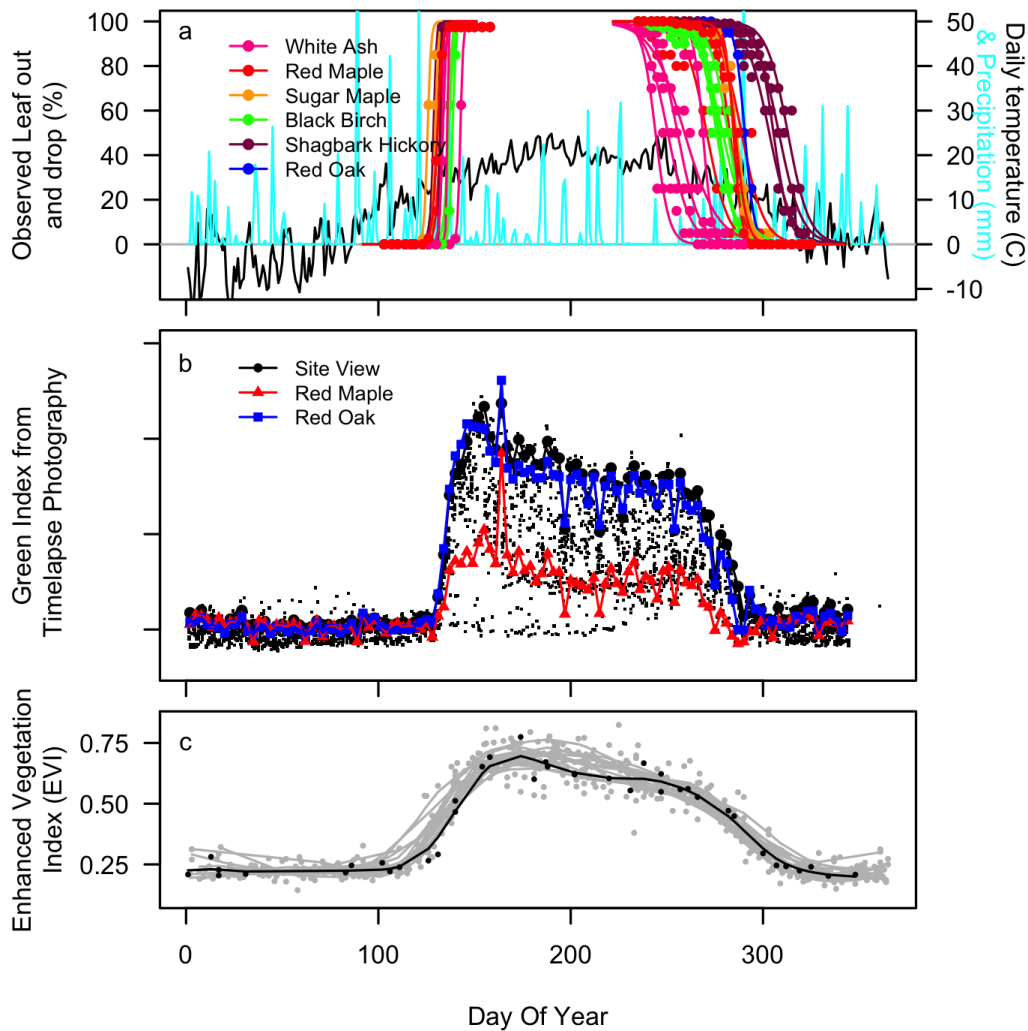


Figure S3 Added variable plots (partial regression) for four variables in the best models of two eco-regions. Variables are: latitude (lat), CDD_{20(Aug.1-Nov.15)} (CDD820), HD_{32(Jul.1-Aug.31)} (HD7832), and GDR_(Sep.1-Nov.15) (GDR9d). Solid red lines are regression lines (p-value<0.001 for all lines).

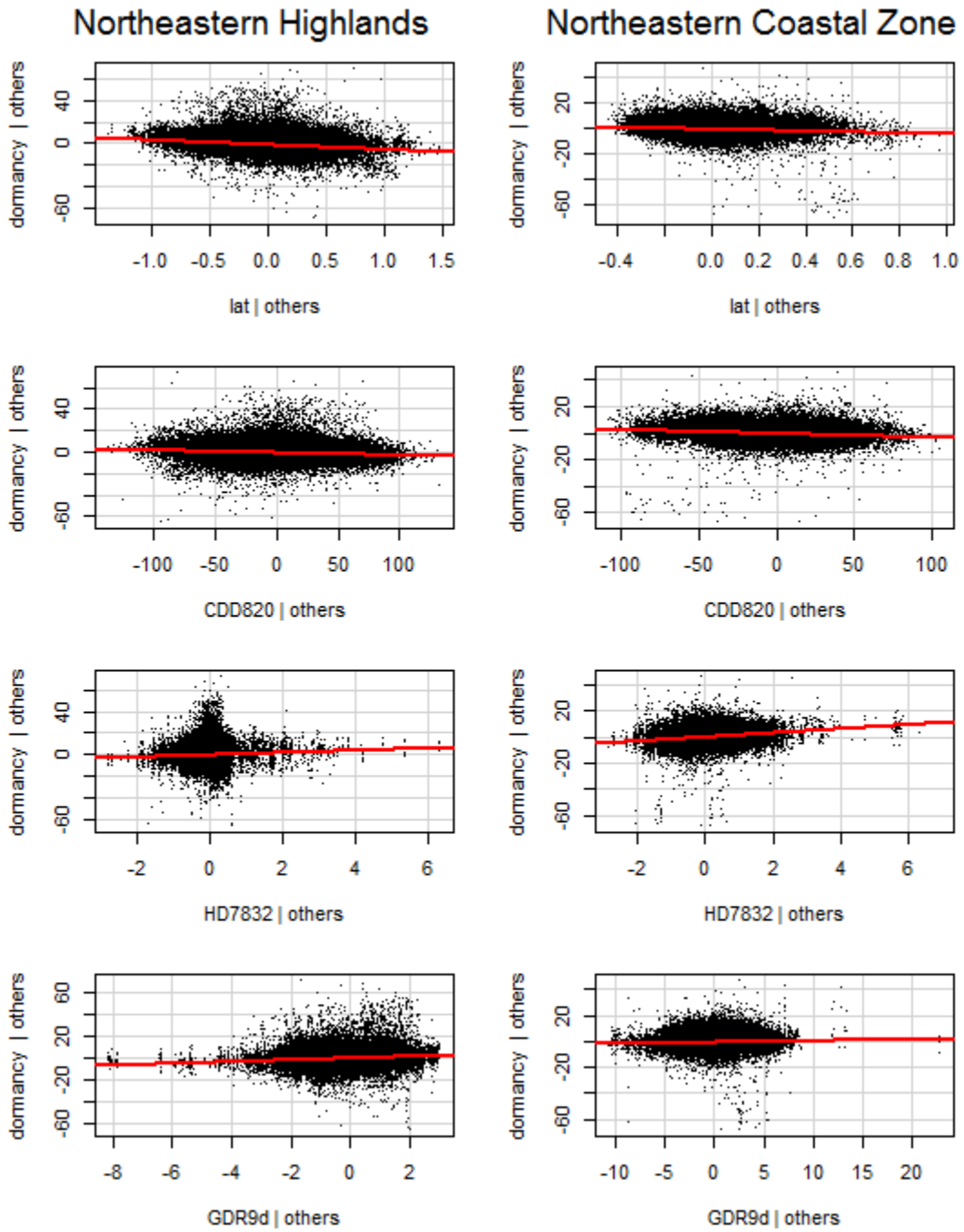


Figure S4 Boxplots of 10-year average values of climatic variables in Northeastern Highlands. The x-axis indicates different 10-year periods with climate change scenarios. L0: 2001-2010; L1: 2041-2050, RCP4.5; L2: 2090-2099, RCP 4.5; L3: 2041-2050, RCP8.5; L4: 2090-2099, RCP 8.5. Variables are: $CDD_{20(Aug.1-Nov.15)}$ (CDD820), $ECA_{(May.1-Jun.30)}$ (ECA56), $ECA_{(Jul.1-Aug.31)}$ (ECA78), $ECA_{(Sep.1-Nov.15)}$ (ECA9d), $FD_{(Sep.1-Nov.15)}^2$ (Fdf_2), $FD_{(Apr.1-Jun.30)}$ (FDs), $GDR_{(Jul.1-Aug.31)}$ (GDR78), $GDR_{(Sep.1-Nov.15)}$ (GDR9d), $GDR_{(Sep.1-Nov.15)}^2$ (GDR9d_2), $HD_{32(Jul.1-Aug.31)}$ (HD7832), $HD_{32(Jul.1-Aug.31)} * RD_{(Jul.1-Aug.31)}$ (HD7832_RD78), $HD_{35(Jul.1-Aug.31)}$ (HD7835), $RD_{(May.1-Jun.30)}$ (RD56), $RD_{(Jul.1-Aug.31)}$ (RD78), $RD_{(Jul.1-Aug.31)}^2$ (RD78_2), and $RD_{(Sep.1-Nov.15)}$ (RD9d).

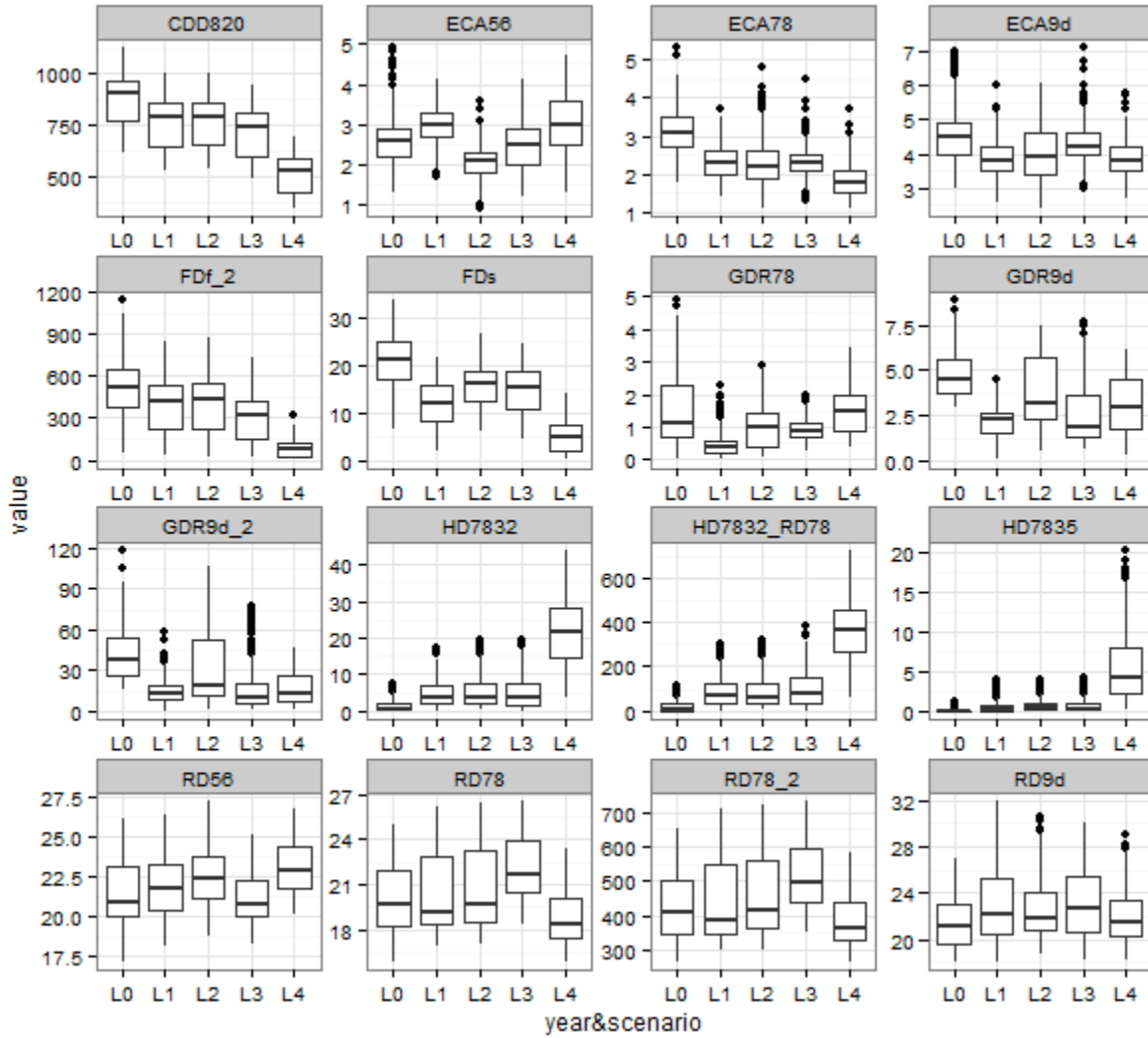


Figure S5 Boxplots of 10-year average values of climatic variables in Northeastern Coastal Zone. The x-axis indicates different 10-year periods with climate change scenarios. L0: 2001-2010; L1: 2041-2050, RCP4.5; L2: 2090-2099, RCP 4.5; L3: 2041-2050, RCP8.5; L4: 2090-2099, RCP 8.5. Variables are: $CDD_{20(Aug.1-Nov.15)}$ (CDD820), $ECA_{(May.1-Jun.30)}$ (ECA56), $ECA_{(Jul.1-Aug.31)}$ (ECA78), $ECA_{(Sep.1-Nov.15)}$ (ECA9d), $ECA_{(Sep.1-Nov.15)}^2$ (ECA9d_2), $FD_{(Sep.1-Nov.15)}$ (Fdf), $FD_{(Sep.1-Nov.15)} * elevation$ (Fdf_elev), $GDR_{(May.1-Jun.30)}$ (GDR56), $GDR_{(Jul.1-Aug.31)}$ (GDR78), $GDR_{(Sep.1-Nov.15)}$ (GDR9d), $HD_{32(Jul.1-Aug.31)}$ (HD7832), $HD_{32(Jul.1-Aug.31)} * RD_{(Jul.1-Aug.31)}$ (HD7832_RD78), $HD_{35(Jul.1-Aug.31)}$ (HD7835), $RD_{(Jul.1-Aug.31)}$ (RD78), $RD_{(Sep.1-Nov.15)}$ (RD9d), and $RD_{(Sep.1-Nov.15)}^2$ (RD9d_2).

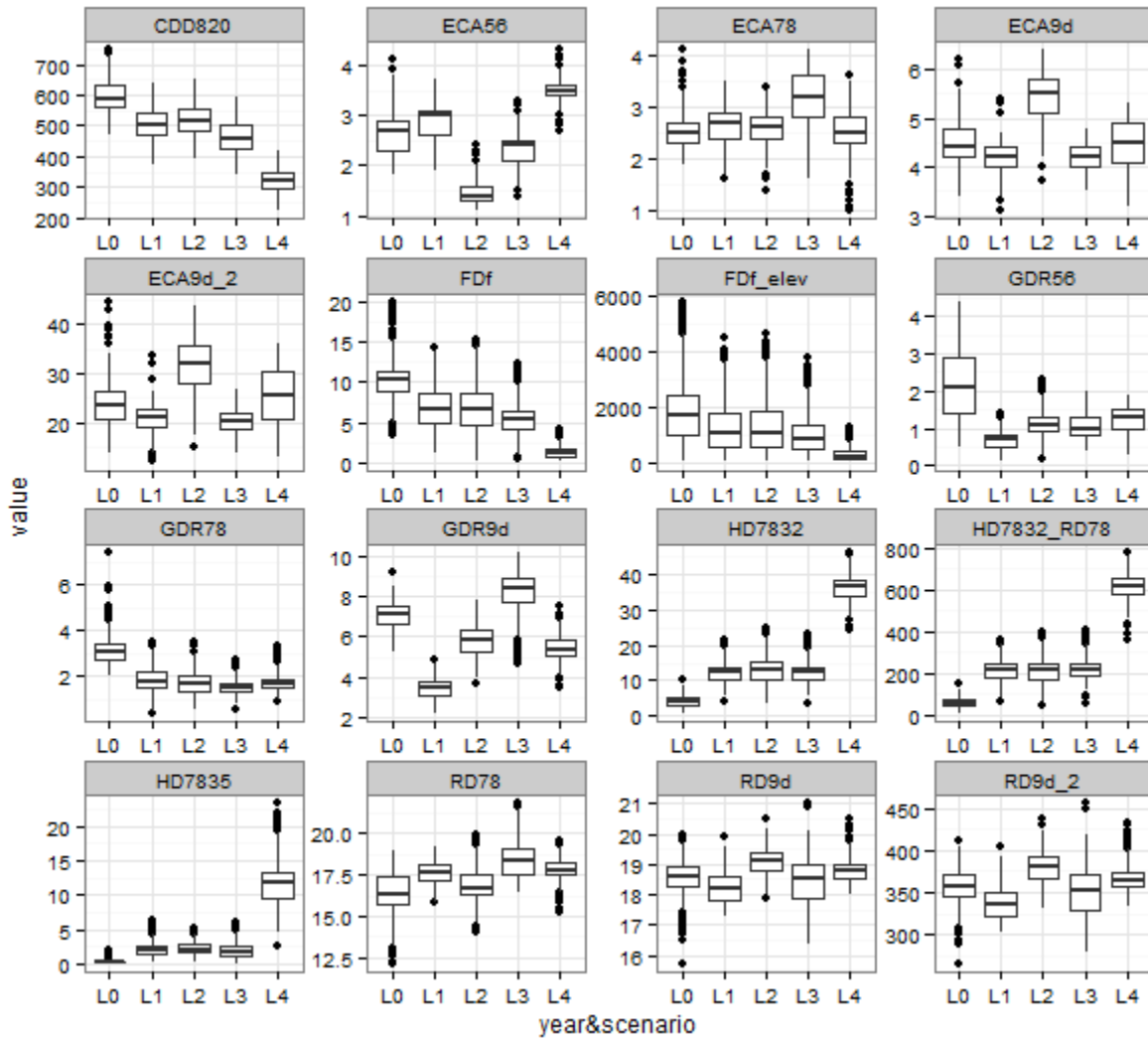
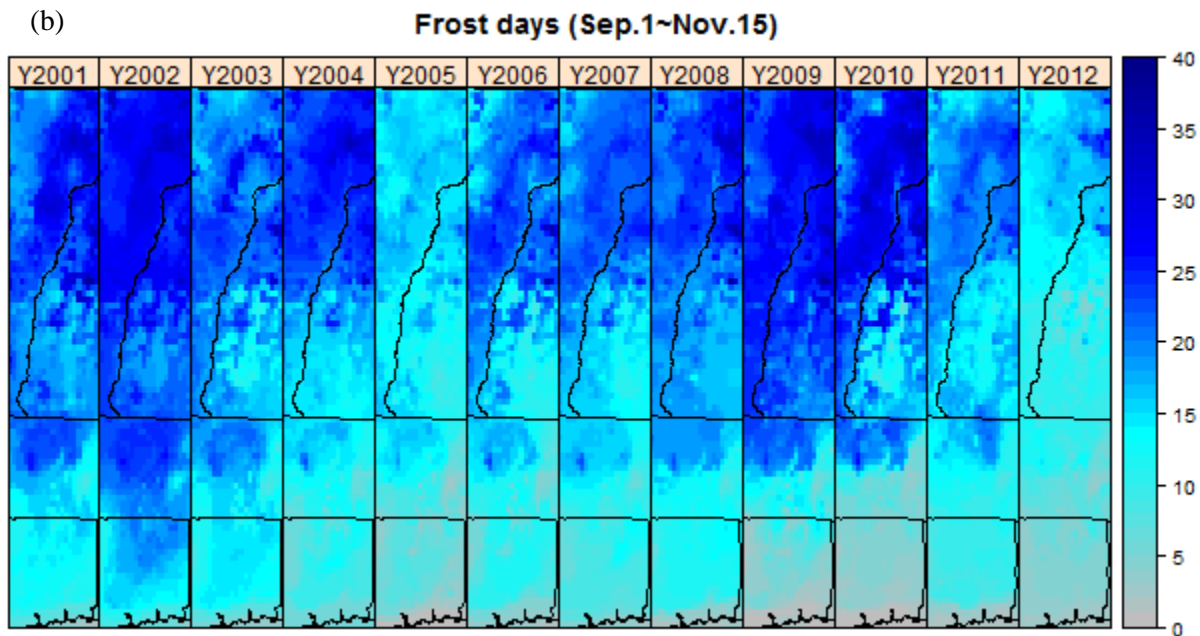
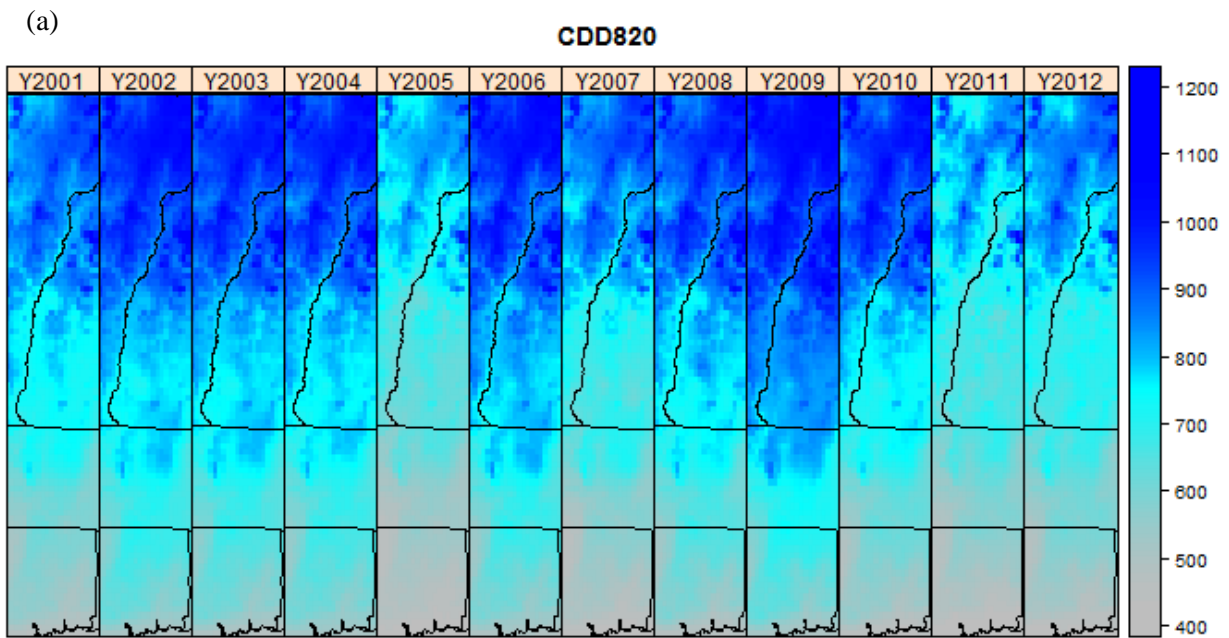


Figure S6 Values of four significant climatic factors (cf. Table 1 & 2) influencing autumn phenology across the landscape of the study domain from 2001 to 2012, showing temporal and spatial variation. (a) cumulative Cold Degree Day from Aug.1st to Nov. 15th, (b) number of frost days from Sep. 1st to Nov. 15th, (c) number of hot days with maximum temperature higher than 32°C from Jul. 1st to Aug. 31st, and (d) droughts from Sep. 1st to Nov. 15th. Data are calculated from PRISM daily weather data (<http://www.prism.oregonstate.edu/>). Black lines are New England states boundaries (cf. figure 1).



References cited:

1. Zhang X, Friedl MA, Schaaf CB, Strahler AH, Hodges JCF, Gao F, Reed BC, Huete A (2003) Monitoring vegetation phenology using MODIS. *Remote Sens Environ*, 84, 471–475.
2. Zhang X., Friedl M.A., Schaaf CB, Strahler AH, Hodges JCF, Gao F (2002) Using MODIS data to study the relation between climatic spatial variability and vegetation phenology in northern high latitudes. *International Geoscience and Remote Sensing Symposium (IGARSS)*, 2, 1149-1151.
3. Zhang X, Friedl MA, Schaaf CB, Strahler AH (2004) Climate controls on vegetation phenological patterns in northern mid- and high latitudes inferred from MODIS data. *Global Change Biology*, 10(7), 1133-1145.
4. Zhang X, Friedl MA, Schaaf CB (2006) Global vegetation phenology from Moderate Resolution Imaging Spectroradiometer (MODIS): Evaluation of global patterns and comparison with in situ measurements. *Journal of Geophysical Research: Biogeosciences*, 111(4), art. no. G04017.
5. Liu L, Liang L, Schwartz MD, Donnelly A, Wang Z, Schaaf CB, Liu L (2015) Evaluating the potential of MODIS satellite data to track temporal dynamics of autumn phenology in a temperate mixed forest. *Remote Sensing of Environment*, 160,156-165.
6. Tang Q, Vivoni ER, Muñoz-Arriola F, Lettenmaier DP (2012) Predictability of evapotranspiration patterns using remotely sensed vegetation dynamics during the North American Monsoon. *Journal of Hydrometeorology*, 13(1), 103-121.
7. Yang Y, Guan H, Shen M, Liang W, Jiang L (2015) Changes in autumn vegetation dormancy onset date and the climate controls across temperate ecosystems in China from 1982 to 2010. *Global Change Biology*, 21(2), 652-665.
8. Pallardy SG (2008) *Physiology of Woody Plants 3rd edition*. Academic Press – Elsevier. Burlington, MA.
9. Hanninen H, Tanino K (2011) Tree seasonality in a warming climate. *Trends in Plant Science* 16(8): 412-416.
10. Paul KP, Rinne PH, van der Schoot C (2014) Shoot meristems of deciduous woody perennials: self-organization and morphogenetic transitions. *Curr Opin Plant Biol* 17:86-95.
11. Xie Y, Silander JA Jr. (2015) Species-specific leaf phenology of deciduous trees captured by digital cameras. Abstract for the Ecological Society of America 100th annual meeting, 9-14 August, 2015, Baltimore, MA, USA.

```
#####  
##### R codes for MODIS dormancy dates modeling #####
```

```
setwd("C:/~/") # set path of data file  
load("modeling_dataset.Rdata") # all models use standardized dataset  
data1.1s=as.data.frame(scale(data1.1)) # eco-region: Northeastern Highlands  
data2.1s=as.data.frame(scale(data2.1)) # eco-region: Northeastern Coastal Zone
```

```
#### Multiple linear regression
```

```
lm1s=lm(dormancy~lat+elev+CDD820+FDf+FDf_2+FDs+HD7832+HD7835+GDR56+GDR78+GDR9d+GDR9  
d_2 +RD56+RD78+RD78_2+RD9d+ECA56+ECA78+ECA9d+HD7832_RD78,data1.1s)  
summary(lm1s)
```

```
lm2s=lm(dormancy~lat+CDD820+elev+FDf+HD7832+HD7835+GDR56+GDR78+GDR9d+RD78  
+RD9d+RD9d_2+ECA56+ECA78+ECA9d+ECA9d_2+FDf_elev+HD7832_RD78, data2.1s)  
Summary(lm2s)
```

```
#### Penalized regressions
```

```
# ridge  
library(MASS)  
rg1=lm.ridge(dormancy~.,data1.1s,lambda=seq(0,50,by=0.1))  
summary(rg1)  
str(rg1)  
select(lm.ridge(dormancy~.,data1.1s,lambda=seq(0,50,by=0.1))) # find lambda for smallest GCV  
rg1$coef[,51]
```

```
select(lm.ridge(dormancy~.,data2.1s,lambda=seq(0,50,by=0.1)))  
rg2=lm.ridge(dormancy~.,data2.1s,lambda=seq(0,50,by=0.1))  
rg2$coef[,19]
```

```
# Elastic Net
```

```
library(glmnet)  
y1=data1.1s[,1]  
X1=as.matrix(data1.1s[,-1])  
fit_enet1=glmnet(X1,y1,alpha=0.5)  
print(fit_enet1)  
plot(fit_enet1)  
coef(fit_enet1,s=0.00055)
```

```
y2=data2.1s[,1]  
X2=as.matrix(data2.1s[,-1])  
fit_enet2=glmnet(X2,y2,alpha=0.5)  
print(fit_enet2)  
plot(fit_enet2)
```

```

coef(fit_enet2,s=0.0004)

# Bayesian LASSO
library(lars)
library(monomvn)

bhs1 = blasso(X=data1.1s[,-1], y=data1.1s[,1], case="hs", RJ=FALSE, normalize=F) #already normalized
str(bhs1)
summary(bhs1)
beta1=apply(bhs1$beta[-1,], 2, mean)

bhs2 = blasso(X=data2.1s[,-1], y=data2.1s[,1], case="hs", RJ=FALSE, normalize=F) #already normalized
beta2=apply(bhs2$beta[-1,], 2, mean)

#### PACS
library(mvtnorm,MASS)
x=as.matrix(data1.1s[,-1])
y=data1.1s[,1]
betawt=lm(y~x-1)$coef
pacs1=PACS(y, x, lambda=1, betawt, type=3, rr=0.5)
rownames(pacs1)=colnames(data1.1)[-1]
str(pacs1)

x=as.matrix(data2.1s[,-1])
y=data2.1s[,1]
betawt=lm(y~x-1)$coef
pacs2=PACS(y, x, lambda=1, betawt, type=3, rr=0.5)
rownames(pacs2)=colnames(data2.1)[-1]
str(pacs2)

#### BMA
library(BAS)
M.ZSn1.s=bas.lm(dormancy~lat+elev+CDD820+FDf+FDf_2+FDs+HD7832+HD7835+GDR56+GDR78
+GDR9d+GDR9d_2+RD56+RD78+RD78_2+RD9d+ECA56+ECA78+ECA9d+HD7832_RD78,
data1.1s,prior="ZS-null", n.models=NULL,
modelprior=uniform(), initprobs="Uniform")
summary(M.ZSn1.s)
coef2=coef(M.ZSn1.s)

M.ZSn2.s=bas.lm(dormancy~ lat+CDD820+elev+FDf+HD7832+HD7835+GDR56+GDR78+GDR9d+RD78
+RD9d+RD9d_2+ECA56+ECA78+ECA9d+ECA9d_2+FDf_elev+HD7832_RD78,
data2.1s,prior="ZS-null", n.models=NULL,
modelprior=uniform(), initprobs="Uniform")
summary(M.ZSn2.s)

```

```
coef2=coef(M.ZSn2.s)
plot(coef(M.ZSn2.s))
```

```
#### Bayesian posterior median model
```

```
library(bayesm)
```

```
library(coda)
```

```
Data1=list(y= as.matrix(data1.1s[,1]),X=as.matrix(data1.1s[,c(2:4,6:9,11:21)]))
```

```
Data2=list(y=as.matrix(data2.1s[,1]),X=as.matrix(data2.1s[,-1]))
```

```
Prior1 = list(betabar=rep(0,18),A=.01*diag(18))
```

```
R1 = 5000 # This is how many iterations to run the Gibbs sampler
```

```
keep1 = 1 # This is the thinning parameter. 1 means no thinning.
```

```
MCMC1 = list(R=R1,keep=keep1)
```

```
BLR1 = runiregGibbs(Data1,Prior1,MCMC1)
```

```
BLR2 = runiregGibbs(Data2,Prior1,MCMC1)
```

```
sum.blr1=summary(BLR1$betadraw)
```

```
sum.blr2=summary(BLR2$betadraw)
```

```
#### Spike and Slab
```

```
library(spikeSlabGAM)
```

```
spslG1.1s <- spikeSlabGAM(dormancy ~
```

```
lin(lat)+lin(elev)+lin(CDD820)+lin(FDf)+lin(FDf_2)+lin(FDs)+lin(HD7832)+lin(HD7835)+lin(GDR56)+lin(GDR78)+lin(GDR9d)+lin(GDR9d_2)
```

```
+lin(RD56)+lin(RD78)+lin(RD78_2)+lin(RD9d)+lin(ECA56)+lin(ECA78)+lin(ECA9d)+lin(HD7832_RD78),
```

```
data=data1.1s)
```

```
summary(spslG1.1s)
```

```
spslG2.1s <- spikeSlabGAM(dormancy ~
```

```
lin(lat)+lin(elev)+lin(CDD820)+lin(FDf)+lin(HD7832)+lin(HD7835)+lin(GDR56)+lin(GDR78)+lin(GDR9d)
```

```
+lin(RD78)+lin(RD9d)+lin(RD9d_2)+lin(ECA56)+lin(ECA78)+lin(ECA9d)+lin(ECA9d_2)+lin(FDf_elev)+lin(HD7832_RD78), data=data2.1s)
```

```
summary(spslG2.1s)
```

```
lm.spsl1=lm(dormancy~lat+elev+CDD820+FDf_2+HD7832+GDR9d+GDR9d_2+RD56+RD78+RD78_2  
+HD7832_RD78, data1.1s)
```

```
summary(lm.spsl1)
```

```
lm.spsl2=lm(dormancy~ lat+elev+CDD820+FDf+HD7832+HD7835+GDR9d+RD78+RD9d+RD9d_2  
+ECA56+ECA78+ECA9d+ECA9d_2+FDf_elev+HD7832_RD78,
```

```
data2.1s)
```

```
summary(lm.spsl2)
```