6-1997

A continental phenology model for monitoring vegetation responses to interannual climatic variability

Michael A. White

Peter Edmond Thornton
The University of Montana

Steven W. Running
University of Montana - Missoula

Recommended Citation
White, M. A., P. E. Thornton, and S. W. Running (1997), A continental phenology model for monitoring vegetation responses to interannual climatic variability, Global Biogeochem. Cycles, 11(2), 217–234, doi:10.1029/97GB00330

GLOBAL BIOGEOCHEMICAL CYCLES, VOL. 11, NO. 2, PAGES 217-234, JUNE 1997

A continental phenology model for monitoring vegetation responses to interannual climatic variability

Michael A. White, Peter E. Thornton, and Steven W. Running

Numerical Terradynamic Simulation Group, School of Forestry, University of Montana, Missoula

Abstract. Regional phenology is important in ecosystem simulation models and coupled biosphere/atmosphere models. In the continental United States, the timing of the onset of greenness in the spring (leaf expansion, grass green-up) and offset of greenness in the fall (leaf abscission, cessation of height growth, grass brown-off) are strongly influenced by meteorological and climatological conditions. We developed predictive phenology models based on traditional phenology research using commonly available meteorological and climatological data. Predictions were compared with satellite phenology observations at numerous 20 km x 20 km contiguous landcover sites. Onset mean absolute error was 7.2 days in the deciduous broadleaf forest (DBF) biome and 6.1 days in the grassland biome. Offset mean absolute error was 5.3 days in the DBF biome and 6.3 days in the grassland biome. Maximum expected errors at a 95% probability level ranged from 10 to 14 days. Onset was strongly associated with temperature summations in both grassland and DBF biomes; DBF offset was best predicted with a photoperiod function, while grassland offset required a combination of precipitation and temperature controls. A long-term regional test of the DBF onset model captured field-measured interannual variability trends in lilac phenology. Continental application of the phenology models for 1990-1992 revealed extensive interannual variability in onset and offset. Median continental growing season length ranged from a low of 129 days in 1991 to a high of 146 days in 1992. Potential uses of the models include regulation of the timing and length of the growing season in large-scale biogeochemical models and monitoring vegetation response to interannual climatic variability.

Introduction and Background

Vegetation phenology, the study of recurring vegetation cycles and their connection to climate, is an important variable in a wide variety of Earth and atmospheric science applications. In particular, as interest in global change research grows, accurate phenology models will become increasingly vital tools, enabling researchers to monitor and predict vegetation responses to interannual climatic variability. The presence or absence of a photosynthetically active canopy has dramatic effects on regional to global ecosystem simulation models [*Running and Nemani*, 1991; *Goetz and Prince*, 1996], coupled biosphere/atmosphere general circulation models (GCMs) [*Sellers et al.*, 1996], and land surface parameterization schemes [*Henderson-Sellers et al.*, 1993]. In these applications, the timing and the length of the growing season control the spatiotemporal dynamics of crucial carbon and water cycles and strongly influence latent/sensible heat transport [*Schwartz*, 1992].

Phenology is highly variable [*Schmidt and Lotan*, 1980] and responsive to long-term variation in climate [*Sparks and Carey*, 1995]. Our research goal was to develop a simple model capable of representing climatically induced phenological variability at a regional level in the continental United States. Central features of the methodology include biome-specific models and simple data requirements consisting only of commonly available meteorological and climatological variables. We faced a fundamental scale discontinuity. GCMs and ecosystem simulation models are incapable of capturing species level detail, while most phenological research has focused on a single species or geographic area. In spite of this narrow focus, consistent literature findings indicate starting points from which to construct regional models.

In trees, the initiation of the growing season, or onset of greenness, has been successfully modeled using a cumulative thermal summation [*Hickin and Vittum*, 1976; *Thomson and Moncrieff*, 1981; *Cannel and Smith*, 1986; *Murray et al.*, 1989; *Hari and Häkkinen*, 1991; *Hunter and Lechowicz*, 1992; *Caprio*, 1993; *Hänninen* et al., 1993]. The technique dates to *Réaumur* [1735] and is extremely simple. After an arbitrary start date (usually January 1), mean air temperature or soil temperature above an arbitrary threshold (usually 0°C or 5°C) is summed until a critical value is exceeded, at which point the prescribed phenological event is predicted to occur. Most trees must fulfill a chilling requirement before warmer temperatures begin to affect springtime growth, and some models include this parameter [*Lavender*, 1981; *Kramer*, 1994]. Many, though, including *Lindsay and Newman* [1956]

Paper number 97GB00330.
0886-6236/97/97GB-00330$12.00

and *Valentine* [1983], use thermal summation models which ignore chilling requirements.

Extensive literature reviews by *Vegis* [1964] and *Nooden and Weber* [1978] strongly suggest that short days induce dormancy, or offset of greenness, (cessation of height growth, development of cold hardiness, abscission) in woody plants. It appears that the critical daylength varies with latitude such that northern populations become dormant while days are still relatively long but that southern populations continue growth well into short daylengths [*Heide*, 1974; *Hänninen* et al., 1990; *Oleksyn et al.*, 1992]. Further, the growing season can be prolonged by warm temperatures and curtailed by cold temperatures [*Heide*, 1974; *Koski and Selkäinaho*, 1982; *Smit-Spinks et al.*, 1985].

Other research takes a completely different approach to modeling forest phenology. *Kikuzawa* [1995] accurately predicts patterns of leaf phenology and longevity with a theory based on a carbon cost/benefit analysis and nutrient availability. Instead of specifically predicting phenophases, *Janecek et al.* [1989] regulate photosynthetic activity with a combination of climatic and internal ecosystem carbon assimilation/respiration.

In other biomes, growth is influenced by a variety of factors. Grasslands are sensitive to temperature, precipitation [*French and Sauer*, 1974; *Mueggler*, 1983], and soil moisture [*Dickinson and Dodd*, 1976] . Growth in dry grasslands is controlled largely by moisture [*Kemp*, 1983] while grasslands in cooler, moister areas are limited primarily by temperature [*Ram et al.*, 1988]. Grassland ecosystems will remain green until the growing season is curtailed by heat, drought, or cold [*French and Sauer*, 1974]. Growth in desert shrub grasslands is largely controlled by precipitation and drought stress [*Burk*, 1982; *Kemp*, 1983; *Sharifi et al.*, 1988]. Crop phenology is uniquely independent of an immediate tie to natural phenological influences. Knowledge of factors such as genotype, planting time, fertilization, irrigation, and harvest time can lead to accurate phenological models for specific crops [*Doraiswamy and Thompson*, 1982; *Hodges and French*, 1985; *Lomas and Herrera*, 1985; *O'Leary et al.*, 1985; *Brown*, 1986; *Undersander and Christiansen*, 1986]. For many crops, phenology models based solely on natural climatic cues are unlikely to be adequate predictors.

Since many models are driven by meteorological information, either from actual daily records or stochastically generated daily records from monthly means, meteorology is a viable driver for phenology models. However, to develop and test regional models, observations of phenology are required. At large scales, it is extremely difficult to obtain consistent field phenology observations across landcovers which represent ecosystem activity rather than species-level phenology. To overcome this difficulty, numerous studies have used high-frequency coverage of the terrestrial biosphere by the National Oceanic and Atmospheric Administration (NOAA) advanced very high resolution radiometer (AVHRR) to quantify ecosystem vegetation phenology. The normalized difference vegetation index (NDVI), calculated as (N-R) / (N+R), where N is the near-infrared reflectance and R is the red reflectance, has been related to several biophysical parameters including chlorophyll density [*Tucker et al.*, 1985], percent canopy cover [*Yoder and Waring*, 1994], absorbed photosynthetically active radiation [*Myneni and Williams*, 1994], leaf area index [*Spanner et al.*, 1990b], and productivity [*Prince et al.*, 1995]. Potentially, NDVI ranges from -1 to 1, but Earth surfaces are usually limited to -0.1-0.7 NDVI.

Early in the history of satellite phenology research, *Justice et al.* [1985] used the NDVI to qualitatively assess the global phenology of numerous landcover types. *Goward et al.* [1985] demonstrated that the NDVI corresponds to known seasonality in the continental United States. Satellites were later used to interpret phenology as an indicator of landcover changes in South America [*Stone et al.*, 1994] and to detect phenological dynamics in shrublands [*Duncan et al.*, 1993]. Quantitatively, a variety of methods have been used to identify dates of onset and offset from satellite data (Table 1). Clearly, there are numerous quantitative methodologies in use, each suited to a specific research question.

In this paper, we integrate the basic concepts of traditional meteorologically based phenology modeling with intensive satellite phenology observations to produce biome-specific ecosystem phenology models. We generate observations of onset and offset of greenness for 1990-1992 from high-frequency satellite measurements of contiguous landcover study sites distributed across climatic zones in the continental United States. Through a cross-validation procedure, models are developed which minimize the differences between the predicted and observed dates of onset and offset. Using a spatially distributed meteorology database, final models are applied to the continental United States, and the effects of interannual climatic variability are assessed.

Methodology

Satellite Observations

We obtained a United States Geological Survey (USGS) Earth Resource Observation Systems (EROS) data set of daily

Table Survey of Satellite Detection Methodologies

Method	Reference
0.17 NDVI threshold	*Fischer* [1994]
0.09 NDVI threshold	*Markon et al.* [1995]
0.099 NDVI threshold	*Lloyd* [1990]
Divergence of smoothed curve from autoregressive moving average	*Reed et al.* [1994]
Inflection points on fitted, bell-shaped curve	*Badhwar* [1984]
Largest NDVI increase after air temperature exceeds 5°C	*Kaduk and Heimann* [1996]

NDVI, normalized difference vegetation index

NOAA-11 AVHRR observations for 56 20 km x 20 km study sites within the continental United States and southern Canada [*Hood*, 1993]. Each site, located in an area of general scientific interest and relatively contiguous landcover, consists of 400 1 km AVHRR pixels. From 1990-1992, the five AVHRR channels, satellite zenith angle, solar zenith angle, relative azimuth angle, date, and time of acquisition were recorded for all sites. Because of difficulties in geometric registration, satellite overpasses with extensive cloud cover were not archived. Consequently, at a given site, it is possible to have two overpasses on the same day, or to have no observations for several days. If two overpasses occurred on the same day, we selected the higher NDVI value. Archived data were corrected by EROS for sensor degradation using radiometric calibrations based on preflight gain coefficients and for solar illumination variability using the cosine of the zenith angle [*Hood*, 1993]. Images were registered to a Lambert's Azimuthal Equal Area projection. On the basis of findings by *Goward et al.* [1991] suggesting that extremely off-nadir view angles result in exponentially increasing footprints and severe reflectance and anisoptropy difficulties, we screened the data to remove observations greater than 50° off-nadir.

Studies of vegetation dynamics usually process AVHRR daily images using a maximum value compositing technique [*Holben*, 1986]. Within a compositing period, usually 2 weeks but often longer in chronically cloud-covered areas, the maximum NDVI is selected on a pixel-by-pixel basis, resulting in a complete, hopefully cloud-free image pieced together from multiple overpasses. Pixels with extremely off-nadir satellite view angles are often eliminated. The main assumption of the maximum value compositing technique is that nonoptimal atmospheric, soil, view, and illumination angle conditions depress the NDVI and that the maximum NDVI in the composite period best represents vegetation status. Despite the considerable advantages of this methodology (reduction in cloud contamination and data volume), there is an inevitable loss in temporal resolution.

A smoother of some kind was required for analysis. Since maintaining the temporal detail of the original data set was a priority, we used an alternative to the compositing procedure. The Best Index Slope Extraction (BISE), suggested by *Reed et al.* [1994], provides a methodology which preserves the unique temporal resolution of the data set while reducing the effects of cloud contamination, atmospheric interference, and bidirectional reflectance [*Viovy and Arino*, 1992]. The BISE algorithm contains two main assumptions: (1) NDVI is depressed by cloud and atmospheric contamination, and (2) rapid, nonpersistent increases or decreases in NDVI are inconsistent with natural vegetation growth.

Figure 1a demonstrates an application of the BISE algorithm to 1 year of raw AVHRR observations for a single 1 km AVHRR pixel from a deciduous broadleaf forest site in New England. Searching forward from the first date, an ascending curve is fitted to any high points which do not exceed a truncation threshold, and a descending curve is fitted to any points which are not isolated, large decreases in NDVI. Point-to-point increases in NDVI greater than the truncation threshold are ignored. If there is a decrease in NDVI from one value to the next, it is accepted only if there is no point within a user-defined sliding period which is greater than

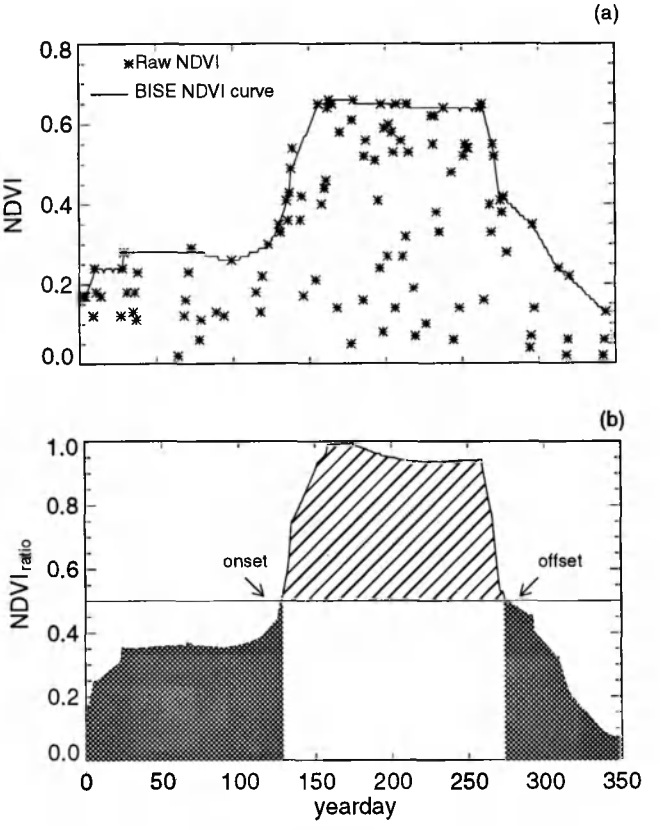

Figure 1. Satellite phenology: processing and detection. (a) Best Index Slope Extraction from raw AVHRR data. Points represent 1 year of daily AVHRR observations from a 1 km northeastern deciduous broadleaf forest pixel. The solid line is the fitted BISE curve. (b) All 400 NDVI curves from the site are averaged and transformed into a ratio from 0 to 1. Onset is observed at the yearday when $NDVI_{ratio}$ exceeds 0.5; offset is observed when $NDVI_{ratio}$ falls below 0.5. Striped area is the growing season; shaded area is the nongrowing season.

20% of the difference between the previous high and current low point. If such a value does occur, it is selected and the low point is ignored. Note that rapid decreases in NDVI will be accepted provided they are persistent. Once BISE has selected the data points for a given time series, a curve is fitted to the time series using a linear extrapolation.

BISE is an appealing method because, unlike compositing, it is independent of a specific time resolution. Rapid increases in NDVI are detected as soon as they occur. Likewise, drastic, persistent decreases in NDVI, as may be caused by fire or defoliation, are captured as soon as they occur. BISE is sensitive to the length of the sliding period and to the magnitude of the truncation threshold. Too long a sliding period may miss natural vegetation changes, while too short a sliding period will result in extremely noisy NDVI curves. Truncation thresholds must allow for rapid growth while rejecting abnormally large jumps. *Viovy and Arino* [1992] found that a 30 day sliding period works best but that a 40 day sliding period produced no apparent adverse consequences and that a truncation threshold of 0.1 removed aberrant NDVI jumps unrelated to normal vegetation activity.

Since the goal of our research is to analyze regional phenological trends, not to investigate within site phenological variability or spatial autocorrelation, we ran BISE for the 400 pixels per site and averaged the curves, producing one NDVI curve for each site. We assume that the single NDVI curve represents the general trend of ecosystem greenness at each site. In a similar technique, *Fuller and Prince* [1996] reduced the effects of NDVI spatial variation by averaging values from 520 km². In their study, though, the sites contained large urban areas; here, contiguous landcover from 400 km² is averaged.

Satellite Detection of Onset and Cessation

Satellite detection of phenological events is a subjective process. It is difficult if not impossible to objectively define an absolute beginning and end of the growing season from satellite observations. The terms onset and offset are popular in the satellite phenology field precisely because they avoid a connection to field measurements of specific phenological stages such as budburst or flowering. Rather than attempting to identify a landcover-specific developmental stage, we pursued a methodology designed to consistently detect onset and offset across landcovers. The state of the ecosystem is assessed with a transformation of the NDVI:

$$NDVI_{ratio} \quad \frac{NDVI - NDVI_{min}}{NDVI_{max} - NDVI_{min}} \qquad (1)$$

where $NDVI_{ratio}$ is the output ratio, ranging from 0 - 1, NDVI is the daily NDVI, $NDVI_{max}$ is the annual maximum NDVI, and $NDVI_{min}$ is the annual minimum NDVI. This method is similar to the Vegetation Critical Index of *Kogan* [1995] and the Relative Greenness of *Burgan and Hartford* [1993]. The ratio method, developed by *Kogan* [1995] and *Burgan and Hartford* [1993] with an ultimate goal of near real-time applicability, received criticism because of a reliance on long-term values for $NDVI_{min}$ and $NDVI_{max}$, which are likely to be unstable through time, especially in grass canopies. Since we create the $NDVI_{ratio}$ using annually redefined $NDVI_{min}$ and $NDVI_{max}$ for each year, the above concerns should not be significant. Figure 1b illustrates the $NDVI_{ratio}$ transformation of the average NDVI curve (created from 400 pixels) from the same site as in Figure 1a. An $NDVI_{ratio}$ of 0 corresponds to the annual minimum NDVI; an $NDVI_{ratio}$ of 1 represents the annual maximum NDVI. The transformation is attractive because it is consistent. An $NDVI_{ratio}$ of 0.25 states that a sight has attained 25% of its maximum greenness regardless of landcover. Thus a single $NDVI_{ratio}$ threshold may be used, obviating the need to establish absolute NDVI thresholds or landcover-specific thresholds.

We tested a range of thresholds to determine, for a given landcover, which threshold identified the periods of greatest increase (onset) and decrease (offset) in the $NDVI_{ratio}$ curve. Using a range of thresholds, we generated observation dates for onset and offset of greenness for the 56 sites for 1990-1992 and calculated total growing season length as offset-onset. Table 2 illustrates that for deciduous broadleaf forest (DBF) and grassland sites, decreases in the threshold lead to unsurprising increases in total growing season length. For both biomes, the change in growing season length is minimized at a threshold of 0.5, indicating that on average, a threshold of 0.5 tends to occur at the point of greatest slope. This in turn indicates that the increase and decrease in greenness is most rapid at a threshold of 0.5. We submit that for ecological models using only one date at which to trigger photosynthetic activity, the period of most rapid growth, not first leaf or budburst, is the most ecologically relevant trigger. Further, at low NDVI, vegetation signals are often confused with soil reflectance [*Huete et al.*, 1992], making accurate detection difficult. Thus we choose 0.5 as the most appropriate threshold.

Table 2. Satellite Phenology Observations: Effects of Varying $NDVI_{ratio}$ Threshold

	Tree				Grass			
T_n	On	Off	GS		On	Off	GS	
0.60	127	278	151		150	240	90	
0.55	125	284	159	8	145	247	102	12
0.50	122	287	165	6	141	253	112	10
0.45	118	290	172	7	137	260	123	11
0.40	111	292	181	9	132	266	134	11

GS, growing season; Δ, change in growing season length, T_n, $NDVI_{ratio}$ threshold

Other satellite phenology detection methodologies exist. The threshold methods in Table 1 [*Lloyd*, 1990; *Fischer*, 1994; *Markon et al.*, 1995] assume that a single threshold is applicable across landcovers. However, variation in background reflectances of different vegetation types makes this a tenuous assumption [*Huete et al.*, 1992]. The methods of *Badhwar* [1984] and *Reed et al.* [1994] employ smoothed curves which might mask high-frequency vegetation changes. The $NDVI_{ratio}$ method normalizes for differences between landcovers, retains temporal detail, and, similarly to *Kaduk and Heimann* [1996], identifies maximal changes in NDVI as times of onset and offset.

Ground Observations

Long-Term Ecological Research programs at two satellite study sites (Hubbard Brook Experimental Forest, site 7; Harvard Forest, site 8) collected detailed phenological measurements for 1990-1992. Researchers at Harvard Forest recorded dates of development at initial budburst; 75% of total leaf expansion; greater than 95% of total leaf expansion; onset of fall color; 10-25% leaf fall; and 95% leaf fall. Data were collected for 33 understory and overstory species at 3-7 day intervals from April through June. Using a representative sample of dominant overstory and understory tree species (O'Keefe, personal communication, 1996), we compiled average dates for each developmental stage. At Hubbard Brook, data were recorded in spring and fall for American beech (*Fagus grandifolia*), sugar maple (*Acer saccharum*), and yellow birch (*Betula alleghaniensis*) at eight plots (data available online at gopher://hbrook.unh.edu). Three dominant

or codominant individuals of each species were marked and sampled at each plot with a numerical scheme where spring conditions range from 0 (unexpanded buds) to 4 (full leaf expansion and summer canopy), and fall conditions range from 4 (full summer canopy) to 0 (winter conditions, all leaves fallen). We compiled average development dates for each species at each plot and averaged all eight plots. Finally, we determined if the dates of onset and offset of greenness extracted from the satellite methodology corresponded to field observations and, if so, to what developmental stage.

Meteorological Interpolation

We used Daymet [*Thornton et al.*, 1997], a spatial extrapolation of a mountain microclimate simulation model (MT-CLIM, [*Running et al.*, 1987]), to generate daily surfaces of temperature, precipitation, and radiation for the continental United States. As inputs, Daymet requires daily observations of temperature and precipitation, station locations and elevations, and a digital elevation model (DEM). The primary interpolation tool is a spatially variable truncated gaussian weighting filter. Interpolation of maximum and minimum temperature (TMAX and TMIN, degrees Celcius) and precipitation (PPT, centimeters) at any given grid cell is a function of the weighted station records and calculated elevational corrections. Elevational corrections are based on daily regression calculations using observed elevational weather variability within the pixel's gaussian filter radius. This is an especially important component of precipitation predictions in complex terrain. Incident short wave radiation (RAD, watts per square meter) is calculated using daylength, Earth-Sun geometry, assumptions regarding atmospheric transmissivity, optical air mass, and diurnal temperature range [*Bristow and Campbell*, 1984]. Diffuse and direct radiation are calculated at a minute time step, summed, and divided by daylength, producing daylight radiative flux density. See *Thornton et al.* [1997] for full details and *Running et al.* [1987], *Hungerford et al.* [1989], and *Glassy and Running* [1994] for further background and testing of many MT-CLIM functions implemented in Daymet.

We obtained 1168 daily TMAX, TMIN, and PPT records within the continental United States from the National Center for Atmospheric Research (Figure 2). Station distribution, correlated with population density, is high throughout much of the eastern United States, the Puget Sound area, California, and the Wasatch Front in central Utah. Station density is low in sparsely populated, desert areas such as southern Nevada, Maine, and Death Valley, California. We discarded station records which contained more than 25 missing days in 1 year or more than 5 consecutive missing days. For remaining records which still contained missing days, we used a linear interpolation to fill missing temperature data, set missing precipitation values to 0.0 cm, and set trace precipitation to 0.01 cm.

In order to keep file sizes and computational time at a manageable level while retaining a relatively fine level of topographic detail, we generated output surfaces at a 10 km resolution. Pixels are defined by a 1 km DEM resampled to a 10 km resolution. To avoid spatially discontinuous elevations created by nearest neighbor resampling, we averaged elevations. We constructed a single meteorological file for each satellite site as the area weighted average of the up to nine Daymet pixels which could overlap the study sites.

Model Development

We combined three tools to develop the phenological models: (1) satellite observations of onset and offset, (2) landcover, and (3) climatological and meteorological data. The basic methodology was to reclassify satellite observation sites into coarse landcover groups and to determine, based on traditional phenological functions, what combination of meteorological equations would best predict the satellite-observed dates of onset and offset for each landcover. *Hood* [1993] provides a detailed landcover classification based on *Loveland et al.* [1991]. Since there are not enough sites to construct highly specific models and we are concerned with the phenology of major biome types, we reclassified *Hood* [1993] into one of the six classes defined by *Nemani and Running* [1995]: (1) barren/sparse vegetation, (2) evergreen

Figure 2. Distribution of weather stations used in Daymet interpolation.

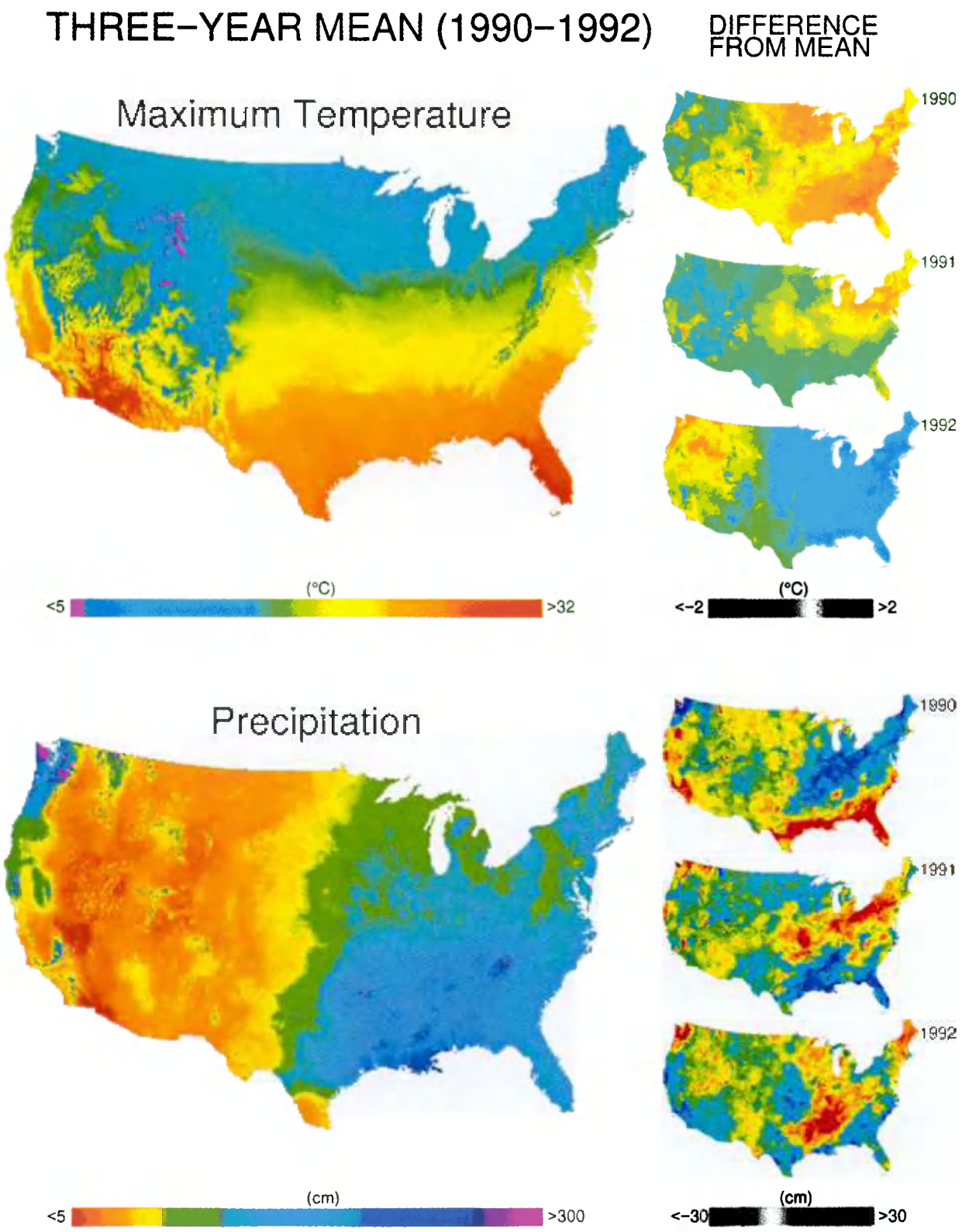

Plate 1. Three-year mean and difference from mean Daymet results. Large panels represent average conditions from 1990 to 1992. Small panels illustrate variation between years.

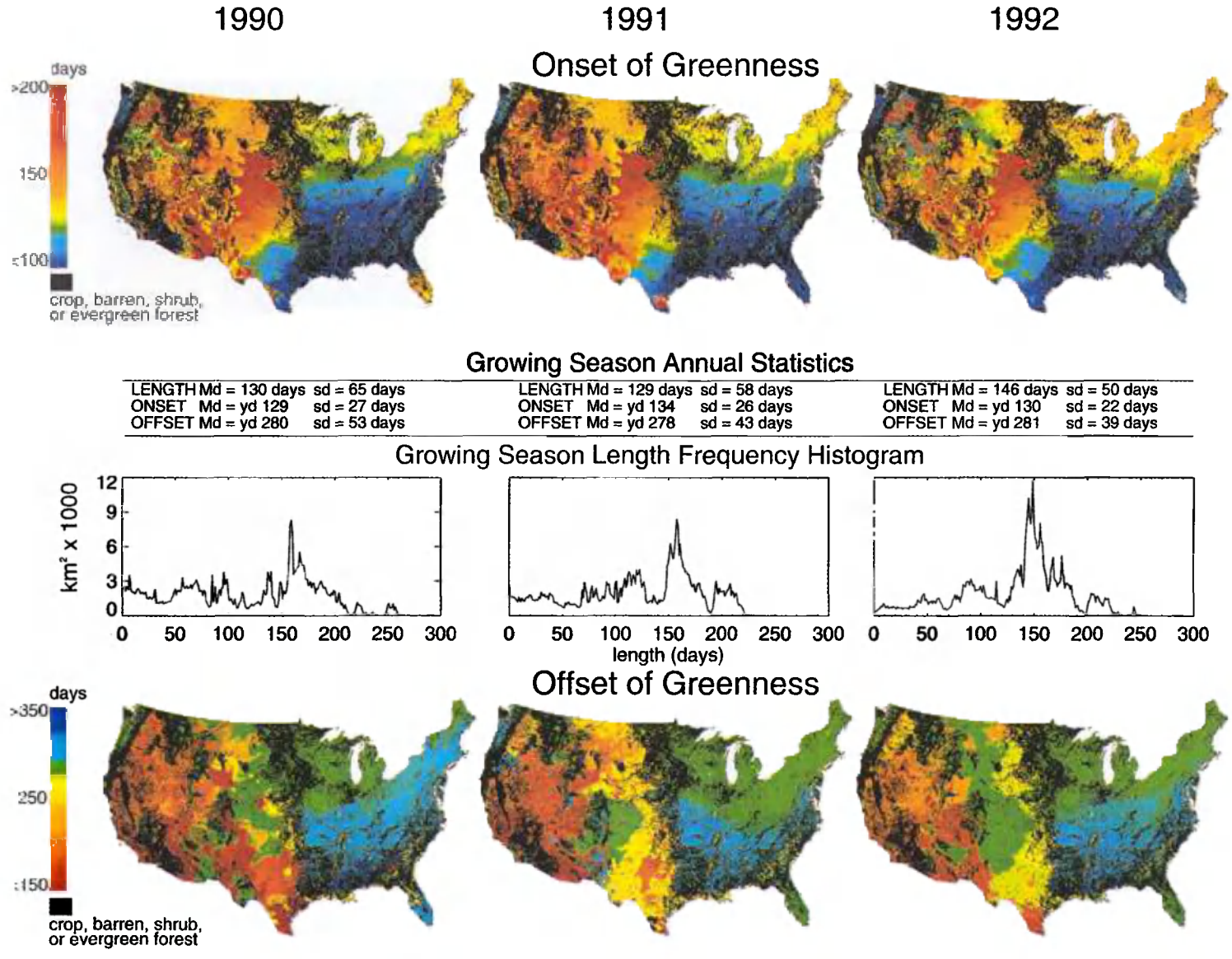

Plate 2. Ecosystem phenology applied to continental scales. Models developed from the cross-validation are applied to continental scales using a landcover map and Daymet distributed meteorology. Upper panels illustrate onset; lower panels illustrate offset. Middle panels are frequency distributions of growing season length. Colored areas are either grass or deciduous broadleaf forest (from landcover by *Nemani and Running* [1995]). Black areas are crop, barren, or evergreen (forest or desert shrubland) and were not modeled.

shrub, (3) grass (including cereal crops), (4) broadleaf crops, (5) evergreen needleleaf forest, and (6) deciduous broadleaf forest (DBF).

Although considerable seasonality can exist in evergreens, we elected to remove evergreen tree sites from the analysis for two reasons. First, because of red reflectance saturation at greenness minima (described by *Huete* [1988], *Spanner et al.* [1990a], and *Baret and Guyot* [1991]), satellite observation of southern pine sites are unlikely to detect the type of large seasonal variation in leaf area as reported by *Curran et al.* [1992]. Second, in western pines, while satellite detection of seasonality is common, it is usually related to snowmelt and deciduous understory dynamics, not to seasonality of the dominant landcover [*Spanner et al.*, 1990a, b; *White and Running*, 1994].

Since our goal was to develop models for pure classes in the continental United States, sites classified by *Hood* [1993] as Canadian (outside of meteorology network), urban, wetland, or mixed landcover were removed from the analysis. Barren areas (class 1) were not modeled due to a lack of vegetation dynamics. Modeling of crops, which require detailed ancillary information, and deserts, where, without extensive corrections, vegetation signals are often indistinguishable from soil reflectance, was beyond the scope of this study. The remaining sites were grouped from classifications by *Hood* [1993] into one of *Nemani and Running's* [1995] classes, ultimately producing 3 years of data for eight DBF and five grassland sites (Table 3).

For both biomes, we investigated patterns of onset and offset with a suite of climatological and meteorological variables. We followed a basic two-step methodology. First, the satellite-observed dates of onset and offset are used as pointers to dates in the meteorological files. For example, if onset is observed to occur at yearday 125, the soil temperature summation at yearday 125 is extracted from the meteorology file. This procedure is applied to all DBF sites until 24 summation values are extracted. Second, the general meteorological or climatological conditions (such as summations) which best predict the satellite-observed dates are identified. Within the deciduous broadleaf forest class (class 5), we focused primarily on soil temperature (SOILT) summations with various thresholds to predict onset (soil temperature calculated as an 11-day running average after *Zheng et al.* [1993]) and photoperiod to predict offset. Assuming that grasslands were dominated by precipitation and temperature controls, we applied a variety of precipitation stress indicators, absolute temperature limitations, and temperature summations. Initial model development included all sites and all years for a given landcover.

Table 3. Summary of Satellite Observation Sites

Site	Latitude	Longitude	Landcover[a]	Landcover[b]
	46.440	-090.520	NF	DBF
	43.870	-075.250	NH	DBF
	39.290	-086.350	MH	DBF
	38.650	-080.590	MH	DBF
	35.640	-083.850	MH	DBF
	35.080	-086.150	MH	DBF
7	44.250	-071.333	NH	DBF
8	42.533	-072.333	DF	DBF
9	47.520	-111.130	G	G
10	40.960	-104.840	G	G
11	40.650	-099.610	G	G
12	46.830	-103.720	G	G
13	48.940	-113.325	G	G

NF, northern forest; NH, northern hardwood; MH, mixed hardwood; DF, deciduous forest; G, grassland; DBF, deciduous broadleaf forest

[a] Original classifications from *Hood* [1993]

[b] Reclassifications using *Nemani and Running* [1995]

Table 4. Daymet Cross-Validation Error Statistics

	Daily Predictions From Daily Observation		Annual Predictions From Daily Observations	
	MAE	BIAS	MAE	BIAS
TMAX, °C	1.91	0.53×10^{-5}	0.77	0.0019
TMIN, °C	1.78	-1.8×10^{-5}	0.99	-0.0067
PPT, %	NA	NA	19.4	2.7
PPT, cm	NA	NA	13.2	-2.5

Averaged values for 1990-1992. Daymet, daily meteorology; MAE, mean absolute error; TMAX, maximum temperature; TMIN, minimum temperature; PPT, precipitation; NA, not applicable

In an iterative process, we compared the predicted dates of onset and offset with the satellite observations and refined the models to reduce the Mean Absolute Error (MAE) associated with the predictions. After the general form of the models was established, we used a cross-validation procedure to select final model equations and parameters. In the cross-validation, one observation is withheld, the models are developed from the remaining observations, and a prediction is generated for the withheld observation. Since each observation is eventually used in model construction, but is independently predicted, the cross-validation method is ideally suited for small data sets (for cross-validation method, see *Stone* [1974]). Thus model and parameter selection is objective and repeatable. Error statistics are calculated from the cross-validation runs, and a final model is developed from average parameters generated by the cross-validation. Following model development, we completed continental runs in which the landcover image was integrated with the model equations to produce images of onset and offset for the DBF and grass biomes in the continental United States.

Interannual Variability

The 3 years of data available from the satellite data set is an inadequate test of the model's ability to represent interannual variability in vegetation dynamics. To assess this temporal ability, we test the DBF onset model against a lilac clone (*Syringa vulgaris*) phenology data set from 1962-1994. The data set consists of first leaf, 95% leaf, first bloom, full bloom, and end bloom data for 178 stations in the eastern United States (see *Schwartz* [1994, 1997] for data set description and site locations). This data set, collected

through a United States Department of Agriculture program, is highly unique. It is geographically and temporally extensive, collected with precisely controlled methodologies, and is by far the best available long-term phenology data set in the United States. Here we consider data for first leaf. Additionally, meteorological data (TMAX, TMIN, snow on ground) is included from the nearest weather station. In many cases, phenology data is missing, especially during the initial and final years of the study. Meteorology data is also incomplete. We used a one-dimensional version of Daymet to calculate RAD for each site. We then used the DBF model to generate phenological predictions at each site in which the meteorology data passed the Daymet screening algorithms. Predictions and observations were then weighted by 1°C climatic zones so that geographic areas with heavy representation would not bias results. The average date of observed first leaf was calculated, and interannual variability of predictions and observations were investigated relative to this date.

Results

Daymet

Final cross-validation results for Daymet are presented in Table 4. Daily temperature MAE is 1.78°C for TMIN and 1.91°C for TMAX. At an annual scale, the pattern is reversed, where TMIN MAE exceeds TMAX MAE. Bias is consistently low for TMAX and TMIN. Errors for annual PPT predictions from daily observations are 19.4% or, in absolute terms, 13.2 cm. PPT bias is 2.7% or -2.5 cm, indicating that precipitation tends to be slightly underpredicted. In general, error statistics are similar to errors found by *Thornton et al.* [1997].

Plate 1 illustrates spatial and temporal variability in temperature and precipitation. The upper large image shows that in the East, temperature increases from north to south, while in the West, the influence of the Rocky Mountains, large riparian zones, basin and range topography, and the presence of large deserts create a more varied pattern. The difference from mean panels illustrate that 1990 was generally hot and that 1992, the year after the Mount Pinatubo eruption, was cool, especially in the East. Mean PPT is sharply divided between wet and dry areas by a north-south border running approximately along the short-tall grass prairie division. To the east of the border, PPT is highest in the south; to the west of the border, PPT, strongly influenced by topography, is more variable. Difference from mean images underscore this pattern, showing that in each of the 3 study years, high PPT areas were located in different western regions. In the east, the location of the Texas to New England storm track shifted in each year: north in 1990, south in 1991, and even farther south in 1992.

BISE and Detection of Onset and Offset

Initial BISE runs with the default 30 day sliding period revealed that in some cases, the off-nadir screening removed a quantity of values sufficient to cause numerous failures in the curve fitting algorithm. The problem occurs if there are no observations within the length of the sliding period, in which case 30 consecutive days are set to 0. Increasing the sliding period to 40 days removed this problem without detracting from the algorithm's ability to detect rapid decreases in NDVI (see *Viovy and Arino* [1992] for discussion). The default truncation threshold of 0.1 NDVI removed artificially large jumps in NDVI while still allowing for periods of sharply increasing growth.

Table 5. Mean 95% Confidence Intervals

	1990	1991	1992
Tree onset	±1.6	±1.6	±0.6
Tree offset	±0.7	±0.8	±1.7
Grass onset	±1.5	±1.3	±1.8
Grass offset	±2.1	±2.3	±1.0

While a consistent detection methodology based on the BISE algorithm is useful, it is equally important to know the confidence intervals for the obtained observation dates. Table 5 presents the 95% confidence interval (CI) for the observed dates of onset and offset for the DBF and grass biome (1990-1992). In 1990, for example, there is a 95% likelihood that on average, the true date when $NDVI_{ratio}$ exceeded 0.5 in the DBF sites occurred within ±1.6 days of the reported date. During the 3 study years, the onset CI is roughly the same for DBF and grasses, but the offset CI is considerably less variable for DBF than for grasses. Across years and biomes, the CI ranges from a minimum of 0.6 days to a maximum of 2.3 days.

Figure 3 suggests that satellite-based estimates of onset and offset are comparable to the same general states of phenological development at the Harvard Forest and Hubbard Brook LTER sites. At the Harvard Forest, the satellite observations of onset correspond to an average of 30% leaf expansion in the spring (Figure 3a) while observations of offset correspond to an average of 15% leaf drop in the fall (Figure 3b). At Hubbard Brook, observed onset occurred at an average stage of 1.8, which represents leaves in the initial stages of expansion, or about 1 cm long (Figure 3c). Offset occurred at an average value of 1.6, or roughly 30% leaf fall (Figure 3d). Although considerable variability exists, especially at Hubbard Brook, and offset data was unavailable at the Harvard Forest for 1990, results suggest that an $NDVI_{ratio}$ of 0.5 corresponds to a ground state of initial leaf expansion (1 cm or 30%) in the spring and to a 15-25% leaf drop in the fall.

Model Building: Cross-Validation

DBF. Initially, we believed that one degree day summation or precipitation level could be found to represent onset. Within the eight DBF landcover sites (1-8), for example, it seemed plausible that all sites might require the same degree of warming to initiate growth. This did not prove to be the case. Calculation of the thermal summation as a function of mean annual temperature, such that warmer sites require larger summations, resulted in vastly improved predictions. Further, incorporating radiation improved the model. For each step of the cross-validation, a combined

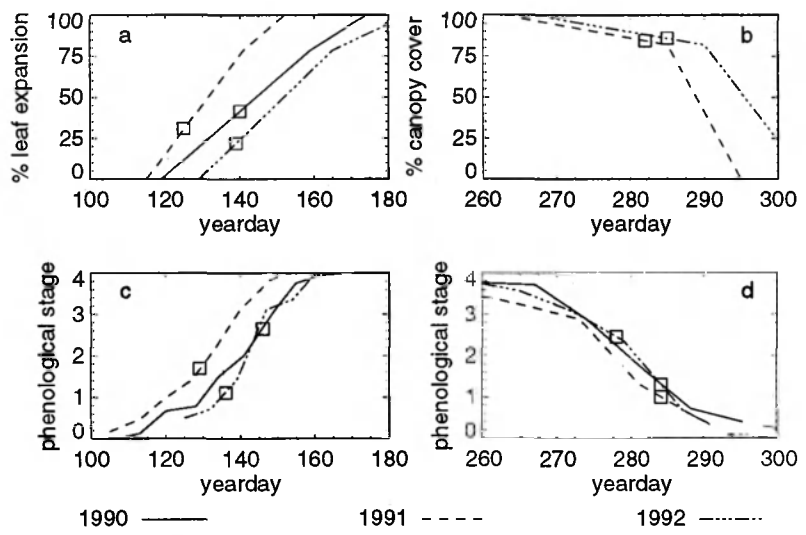

Figure 3. Comparison of field observations and satellite observations. The curves represent trends of field-measured leaf phenology for the Harvard Forest (a) spring and (b) fall and the Hubbard Brook Experimental Forest (c) spring and (d) fall. At Hubbard Brook, a numerical scheme is used where spring conditions range from 0 (unexpanded buds) to 4 (full leaf expansion and summer canopy) and fall conditions range from 4 (full summer canopy) to 0 (winter conditions, all leaves fallen). Squares are superimposed on the curve at the day of satellite-observed onset or offset. For example, in Figure 3a, onset was observed on yearday 140 in 1990, yearday 125 in 1991, and yearday 139 in 1992.

function of temperature summation * radiation summation was developed with a multiple linear regression. The threshold value, with averaged parameters from the 24 cross-validation runs, is calculated as

$$SUM_{com} = -16873361 + 760757 * T_{avg} + 81322 * RAD_{avg} \quad (2)$$

Where SUM_{com} is the combined soil temperature * radiation summation; T_{avg} is the 1990-1992 mean air temperature; and RAD_{avg} is the 1990-1992 mean daylight short wave radiation in watts per square meter. Figure 4 illustrates that warm, high-radiation sites require a larger SUM_{com} to induce onset than do cool, low-radiation sites. Considered alone, radiation summations at onset are relatively constant with changes in site radiation regime but yield unacceptable results. Addition of precipitation parameters did not improve model predictions. Onset is predicted to occur when

$$\sum_{0}^{i}(STSUM * RADSUM) \geq SUM_{com} \quad (3)$$

where STSUM is a soil temperature summation calculated as

$$\sum_{0}^{i} SOILT \geq 0.0°C \quad (4)$$

and RADSUM is a radiation summation calculated as

$$\sum_{0}^{i} RAD \quad (5)$$

where SOILT and RAD are daily values from Daymet. Varying the summation threshold changed the parameters in (2), but not the fundamental relationships. We chose the 0°C threshold based on the rationale that cytokinins, produced by fine roots and required for leaf expansion, cannot be produced until the soil is thawed. We calculated summations and critical values using air temperatures and soil temperatures. Results were similar in all cases, but using mean air temperature for the critical value and soil temperature for the summations produced marginally superior results.

Photoperiod alone could be used to reasonably predict offset. The photoperiod associated with satellite observed offset was nearly constant across all DBF sites, with a mean value of 655 min. Since much research has suggested that photoperiod triggers are active only within a certain temperature range, we added upper and lower limits. If the temperature is warm, growth is permitted to continue, while extremely low temperatures will induce offset regardless of photoperiod. Through the cross-validation procedure, we identified upper and lower temperature limits which minimized MAE for the general landcover predictions. Figure 5 shows a contour plot of the MAE for a range of upper and lower temperature thresholds. As long as 2°C is chosen as the lower limit, a wide range of upper limits will nearly minimize MAE; a small local minimum exists at 11.15°C. DBF offset is predicted to occur when

$$\left[DAYL \leq 655\text{min and } SOILT \leq 11.15°C\right] \text{ or } SOILT \leq 2.0°C \quad (6)$$

where DAYL is daylength in minutes and SOILT is daily soil temperature.

Grass. Onset of greenness in the grassland sites was strongly controlled by temperature. A relationship of average

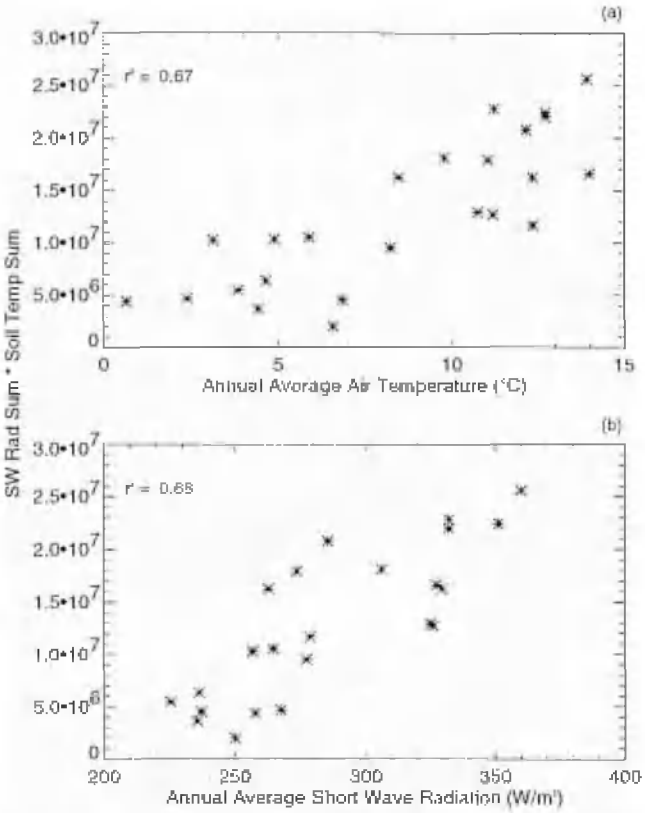

Figure 4. Formulation of the DBF onset model. The (a) annual average air temperature and (b) radiation are plotted on the x axes for each DBF site for each year. The thermal summation multiplied by the radiation summation (extracted from the meteorological files at the date of satellite-observed onset) is plotted on the y axis. A multiple linear regression is used to fit an equation to the points. The multiple linear correlation coefficient is 0.74.

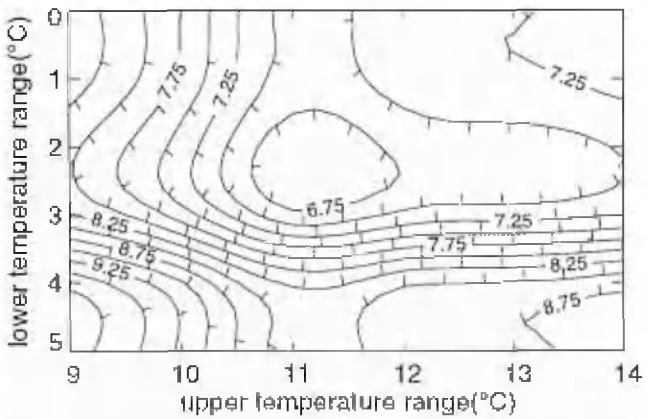

Figure 5. Contour plot of mean absolute errors from the DBF offset cross-validation. The upper (11.15°C) and lower (2.0°C) temperature modifications to photoperiod induced offset are selected from the contour plot at the point where errors are minimized. Temperatures greater than the upper limit can prolong the growing season, while temperatures less than the lower limit can prematurely end the growing season.

annual air temperature and predicted summation indicated that two primary groups of grasslands exist: cool grassland requiring low summations and warm grasslands requiring high summations. To capture this variability within a landcover while maintaining a single equation, we implemented a normalized step function of the form:

$$\mathrm{STSUM_{crit}} = \frac{\mathrm{STSUM_{max}} - \mathrm{STSUM_{min}}}{2}\left(\frac{e^{\left(a\left(T_{avg}-T_{mid}\right)\right)}-1}{e^{\left(a\left(T_{avg}-T_{mid}\right)\right)}+1}\right) + \mathrm{STSUM_{mid}} \quad (7)$$

where $\mathrm{STSUM_{crit}}$ is the threshold soil temperature summation (summation calculated with (4)); $\mathrm{STSUM_{max}}$ is the maximum summation of warm grasslands at onset (extracted from meteorology files); $\mathrm{STSUM_{min}}$ is the mean summation of the cool grasslands at onset (extracted from meteorology files); T_{mid} is the temperature halfway between the upper and lower annual temperature range of North American grasslands (9°C, defined by *Sims* [1988]); $\mathrm{STSUM_{mid}}$ is the summation midway between $\mathrm{STSUM_{max}}$ and $\mathrm{STSUM_{min}}$; and T_{avg} is annual average temperature (a Daymet parameter). The equation then contains four constants; one input parameter (T_{avg}); and one variable, a (controls shape of the curve) which is calculated using the CURVEFIT function in the Interactive Data Language (version 4.0.1, 1995, Research Systems, Inc.). Figure 6 displays the extracted summations as a function of temperature. A clear and sharp division is evident between the warm and cool sites.

For model building in the study sites, addition of precipitation parameters did not improve results. Nonetheless, in dry grasslands areas (not available in satellite database), sufficient moisture is a precursor to growth. To account for

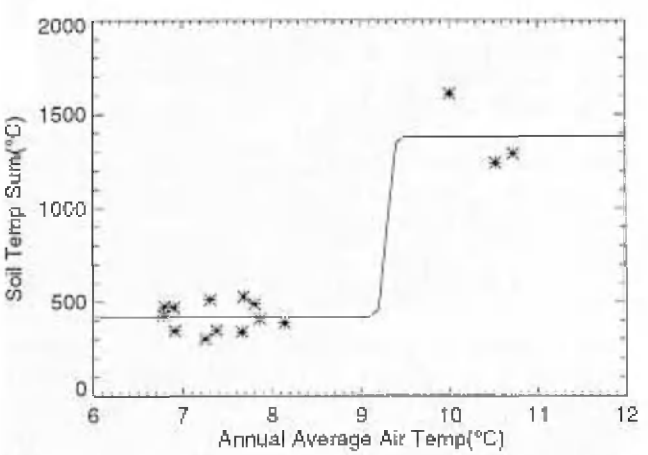

Figure 6. Formulation of the grass onset model. Points represent annual average air temperature plotted against the thermal summation extracted from the meteorological files at the date of satellite-observed onset. A normalized step function requiring only temperature as an input parameter is fit to the data points and provides a single-equation method of representing different thermal requirements of cool and warm season grasses.

this necessity in the model, we included a precipitation summation requirement of the form:

$$PPTSUM_{crit} = k\ PPT_{avg} \quad (8)$$

where $PPTSUM_{crit}$ is a precipitation summation and k is a proportion of the average annual precipitation for the site. We increased k to 0.15 at which point cross-validation MAE began to increase. Thus grasslands must attain sufficient moisture (8) and temperature summations (7) before onset is predicted to occur. Equation (8) is included solely to account for dry grassland moisture requirements in broader model applications. The selected 0.15 level is therefore somewhat arbitrary and is simply the level beyond which cross-validation results were negatively affected. At levels below 0.15, (8) had no effect on error results, suggesting that the study sites were not moisture limited during 1990-1992. Future satellite phenology research should closely examine this threshold level.

No one measure of environmental conditions proved adequate to predict offset, and we ultimately developed two separate scenarios which could induce offset and which reasonably predicted the observed dates: (1) hot, dry conditions or (2) cold weather. Scenario 1 is triggered when

$$PPTSUM_{prev} < CRIT_{prev} \text{ and } PPTSUM_{next} \quad CRIT_{next}$$

$$\text{and } TMAX > CRIT_{TMAX} \quad (9)$$

where $PPTSUM_{prev}$ is the PPT total of the previous 31 days; $PPTSUM_{next}$ is the PPT total of the next 7 days; $CRIT_{prev}$ and $CRIT_{next}$ are critical values in centimeters; and $CRIT_{TMAX}$ is a critical temperature defined as a percent value of $TMAX_{ann}$, the single annual maximum temperature average for 1990-1992. Scenario 1 requires low PPT for the previous month, low PPT for the next week, and temperatures approaching annual maximums. The requirement of low PPT for the next week is included to allow continuation of the main growing season when temporarily harsh conditions are followed by a return of favorable conditions and rapid regrowth, a common occurrence in many grasslands. Cross-validation results showed that MAE is low when $CRIT_{TMAX}$ is between 96 and 89% of $TMAX_{ann}$ but that a minimum exists at 92%. For a $CRIT_{max}$ of 92%, $CRIT_{prev}$ is minimized at 1.14 cm, and $CRIT_{next}$ is minimized at 0.97 cm (Figure 7). Scenario 2 is triggered when

$$TMIN_{smooth} \leq TMIN_{avg} \text{ and Yearday} \geq 243 \quad (10)$$

where $TMIN_{smooth}$ is the boxcar-smoothed daily minimum temperature and $TMIN_{avg}$ is the annual average minimum temperature. Smoothing length, from the cross-validation, is optimized at 3 days. A yearday minimum is included to avoid predicting offset in late spring, an event which was not observed at any site.

Error Statistics

At a 1% significance level, t tests for onset and offset indicate that within both the DBF and grass biome, predictions are not significantly different from observations.

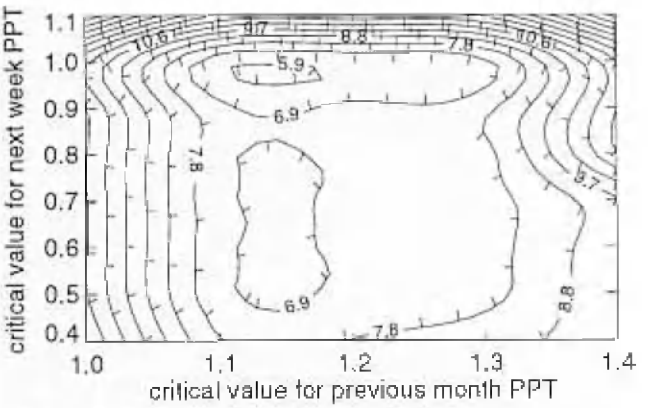

Figure 7. Contour plot of mean absolute errors from the grass offset cross-validation. The previous month's (1.14 cm) and the next week's (0.96 cm) critical precipitation summations are selected at the point where mean absolute errors are minimized. Summations are used to indicate moisture stresses.

However, for most users of predictive ecological models, this is at best an unrevealing statistic. By assessing model precision with the standard deviation, the t test can lead to the impression that highly inaccurate models give very precise predictions. Far more useful information includes (1) the maximum expected error and (2) the expected bias. *Reynolds* [1984] developed a methodology designed to assess these statistics. The maximum expected error, e^* is calculated as

$$e^* = \sqrt{\frac{\sum_{i=1}^{n} D_i^2\, \chi^2_{1-\alpha}(1)}{\chi^2_{1-\alpha'}(n)}} \quad (11)$$

where $\sum D_i^2$ is the sum of the squared differences between predictions and observations, $\chi^2_{1-\alpha}$ is the quantile of the chi-squared value distribution with (1) or (n) degrees of freedom at the $\alpha = 0.05$ probability level. Equation (11) relies on a normal distribution of errors. Wilks-Shapiro and Kolmogorov-Smirnoff tests of normality found no significant differences from the null hypothesis of normality.

Additionally, confidence intervals may be used to assess the potential difference between the expected error distribution $E(D)$ and the mean sample error $\overline{D}$:

$$CI = \overline{D} \pm \frac{S_x\, t_{1-\alpha/2}(n-1)}{\sqrt{n}}$$

where S_x is the standard deviation of the errors, $t_{1-\alpha/2}\ (n-1)$ is the $1-\alpha/2$ quantile of the t distribution with n-1 degrees of freedom. The CI gives an indication of discrepancy (bias) between $E(D)$ and $\overline{D}$ which are likely to occur in future applications.

Error statistics are summarized in Table 6. Mean Absolute Errors (MAE) suggest that on average, the models predicted the observed dates within about 1 week. The maximum critical error at a 95% probability level, e^* , ranges from 14 to

Table 6. Cross-Validation Error Statistics for Differences Between the Predicted and Observed Dates of Onset and Offset of Greenness for DBF and Grassland Sites, 1990-1992

	Tree Sites		Grass Sites	
	Onset of Greenness	Offset of Greenness	Onset of Greenness	Offset of Greenness
Number of events	24	24	15	15
Mean of differences, days	0.2	-1.8	0.0	2.7
Standard deviation	9.1	9.9	7.3	7.6
Critical error e^*, days	14.2	10.8	10.6	11.9
Confidence interval, days	-2.5 to +2.8	-5.0 to +1.3	-2.3 to +2.3	-0.2 to +5.6
MAE, days	7.2	5.3	6.1	6.3

11 days. CI for satellite measurements range from 0.6 to 2.3 days. For some applications, this range of uncertainty in prediction accuracy may be unacceptable. If the user requires maximum errors less than 14 days, or a smaller CI, the model is not appropriate. However, if a potential user requires that predictions of onset for a DBF site have a 95% probability of having errors no more than 14 days, then the model is appropriate. CI distribution is around zero, indicating that none of the models are severely biased. The grass offset model, though, appears to have a slight bias toward overprediction. Figure 8 presents predicted versus observed results in a scatter plot format. Results for grassland onset and offset (Figure 8c and 8d) both exhibit high r^2 values and uniform scatter about the 1:1 line. DBF onset (Figure 8a) has a lower r^2 of 0.57. In both grassland and DBF onset, the model tends to overpredict 1990 and underpredict 1991. DBF offset (Figure 8b) has a relatively low r^2 of 0.39, indicating that offset tends to be underpredicted at low values and overpredicted at high values. However, DBF offset also has by far the smallest range of yeardays, both for predicted and observed values (Figure 8b), the lowest MAE, and nearly the lowest e^* (Table 6).

Continental and Regional Application

Figure 9 shows that the DBF model tends to capture both the major interannual trends in the lilac first leaf observations and the overall trend toward slightly earlier dates of first leaf. In some years, such as 1966 and 1967, the short-term trends are missed. Absolute values are considerably higher (y intercept = 26 days) in the predictions. The crucial result, though, is the remarkably similar overall shape of the two curves.

Plate 2 presents the application of the final phenology models to areas in the continental United States classified by *Nemani and Running* [1995] as DBF or grass. Crop, barren, and evergreen (forest and shrub) areas were not modeled and

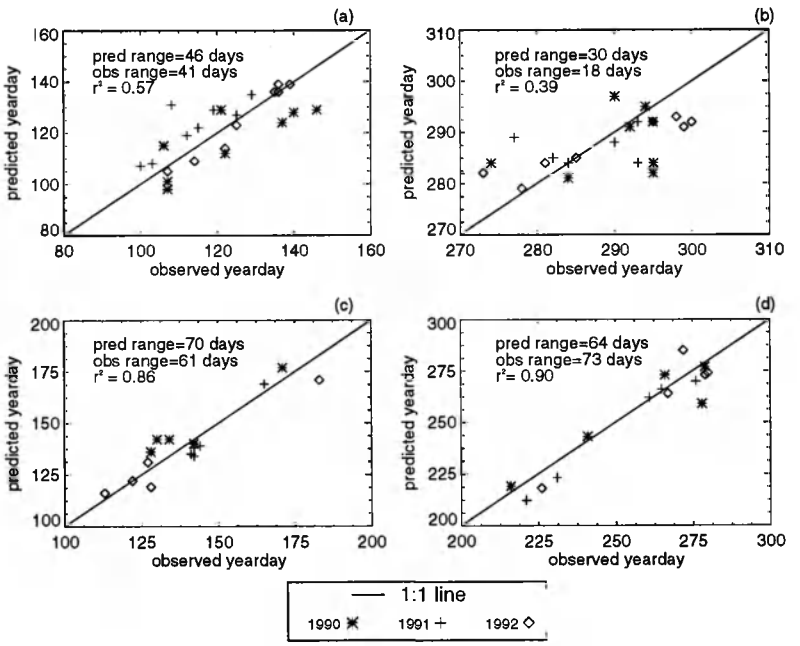

Figure 8. Scatter plots of predicted versus observed dates of onset (a) DBF and (c) grass and offset (b) DBF and (d) grass. Correlation coefficients are fairly high in all cases except DBF offset (Figure 8b) which has the shortest range of occurrences.

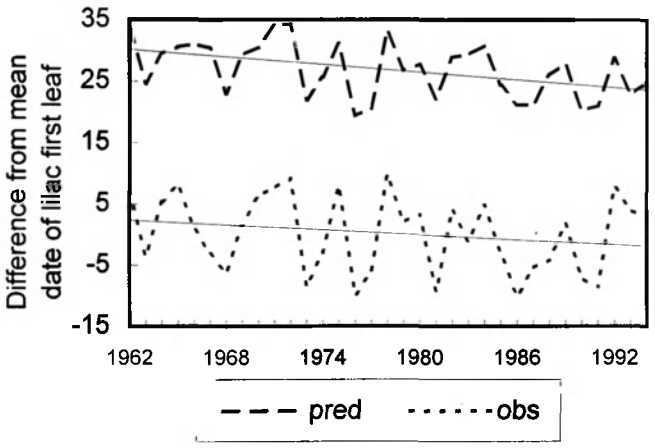

Figure 9. Interannual variability of observed lilac first leaf and DBF model predictions. The lower line shows annual observed deviations from the mean date of first leaf for 193 stations of lilac clones throughout the eastern United States. The upper line shows the deviations of DBF model predictions from the same mean date. Relative interannual variability is well represented. Large absolute deviations are due to the differences in metrics: the lower curve is first leaf of an early growing understory species, while the upper curve is a prediction of 50% maximum greenness for the entire ecosystem.

are colored black in Plate 2. Extensive interannual variability in vegetation dynamics is evident in both biomes, although due to strong influences from moisture stress and great spatio-temporal variability in precipitation, grassland phenology is more variable than DBF phenology. Onset of greenness, relatively spatially consistent in the East, was earliest in 1990 and latest in 1992. In the West, onset was spatially and temporally variable. Onset in the Big Bend area of Texas, for example, ranged from July in 1990 to May in 1992. In general, onset in the West occurred earlier in 1992 than in 1990 or 1991. One consistent feature, from Montana to Texas, is the inverted green wave, in which onset occurs earliest in the northern great plains, latest around Nebraska, and earlier again to the south. This result was also found by B.C. Reed (personal communication, 1995) in a phenological analysis based solely on remote sensing. Offset (Plate 2) was similarly variable. North-South stratification of offset in the East is a function of photoperiod-induced offset. In both 1990 and 1992, there was extensive interaction between temperature and photoperiod, resulting in an extension of the growing season, especially in the South in 1990 and 1992. Greatest variation again occurred in the West where dynamic precipitation levels caused significant interannual variability in offset.

Frequency distributions further reveal the extensive interannual variation in growing season length and distribution (Plate 2). Drought stresses in 1990 led to predictions that thousands of square kilometers had little or no growing season. The reverse occurred in 1992, when very few pixels had a growing season less than 50 days. The median length ranged from 129 days in 1991 to 146 days in 1992, a difference of 17 days. In contrast, the maximum range of median dates was 5 days for onset and only 3 days for offset. In general, because of variability in Western precipitation, standard deviations are much higher for offset than onset. The standard deviation of growing season length followed a trend toward decreasing with increases in growing season length and decreases in histogram skewness: 65 days in 1990, 58 days in 1991, and 50 days in 1992.

Discussion

The purpose of our research was to develop simple ecosystem phenology models requiring only basic knowledge of landcover, climate, and meteorology. Throughout the research, simplicity and widest possible applicability were predominant concerns. In some cases, this approach may have limited the absolute accuracy of the models. For example, use of the six-class landcover image eliminated the possibility of developing equations for multiple grassland types (such as C_3 and C_4). Inclusion of detailed landcover information, such as that provided by *Loveland et al.* [1991], might have allowed for more accuracy in model development but would have presented other difficulties. Detailed landcover maps of this type require large amounts of ancillary information that is difficult to acquire for large areas, making regular updates at continental or global scales impractical. More importantly, in many areas of the world, detailed landcover information is unavailable; basic classification as grassland or forest may be all that is available. The classification presented by *Nemani and Running* [1995] is purely based on remote sensing. As such, it is globally applicable and annually repeatable.

Figure 8 and Table 6 illustrate that for the major North American deciduous landcovers, traditional meteorological models can be used to predict satellite observed dates of onset and offset of greenness with MAE's of about 1 week, maximum critical errors of 10-14 days, and no significant bias. Correlation coefficients are high, except for DBF offset where the range of dates is relatively short. The exception to traditional modeling is the combined radiation and temperature summation, which considerably improved model precision. Temperature, though, is the dominant control. Reasonably good results can be obtained using only a temperature summation, but use of only a radiation summation leads to poor results. That combined summations increase with both annual average temperature and radiation across DBF and grassland sites may initially appear somewhat surprising, but actually this finding is consistent with known phenological controls. First, it is possible that warmer sites may not fulfill their chilling requirements until later in the year and thus are observed to have larger summations at onset. Second, partial fulfillment of chilling requirements, as may occur in warmer areas, requires a longer period of warming to initiate growth [Hänninen, 1987].

DBF onset models, while based on established techniques from traditional phenology research, are empirically determined and do not necessarily represent the physiological processes which control onset and offset. Commenting on temperature summations, *Hunter and Lechowicz* [1992, page 597] state "...good predictions can be obtained using the wrong biological model, and in some cases the appropriate causal model may not give good predictions." On the basis of literature findings, we had expected to find a latitudinal

variation in the photoperiod required to induce offset. In the range of latitudes in this study, though, there was no significant variation, suggesting that below the boreal study region of much photoperiod research, critical phototperiods are relatively constant. Photoperiodic controls of offset and interactions with temperature are as effective as summation models but even more poorly understood.

As suggested in Figure 6, it appears that the grasslands could have been subdivided based on climatic zonation. Instead, we used the normalized step function to provide a transition between these zones. Figure 6 shows that a sharp decrease in predicted summations occurs at an annual average air temperature of 9.0-9.5°C. *Sims* [1988] states that 10°C is the approximate division between grasslands dominated by cool-season (C_3) and warm-season (C_4) grasses. Acknowledging that the sample size is small, it appears that different grassland lifeforms may require dramatically different thermal summations to induce onset. *Monson and Williams* [1982] and *Dickinson and Dodd* [1976] found similar results in phenological research of warm and cool season species which suggested that warm season species are physiologically adapted to higher temperatures and tend to initiate growth several weeks later than cool season species. This is consistent with known C_3 versus C_4 dichotomies. C_3 plants, due to the costs of photorespiration, tend to have a lower optimum temperature for photosynthesis than do C_4 plants. Therefore C_3 grasslands should evolve phenological triggers (low summations) which initiate growth in the cool season, while C_4 grasslands should initiate growth in the hot summer months (high summations). We speculate that the inverted green wave is a function of the switch from C_3-dominated grasslands to C_4-dominated grasslands at roughly the latitude where the inversion occurs. Cool season grasses in the North require low summations and begin growth early. As C_4 species become common but temperatures remain fairly cool, the high summations are not fulfilled until quite late (red colors). As the climate warms with decreases in latitude, onset occurs earlier.

Grassland equations are based on general literature findings with modifications for satellite phenology detection. Grasslands are often composed of numerous species which occupy different portions of the growing season. Onset obviously occurs when the first species initiate growth and is relatively simply modeled with a temperature and precipitation summation. Growth continues until a hot dry spell occurs, causing wilting and brownoff of grasses. A further PPT event, though, can cause regrowth of the same or other species. If regrowth occurs, it is likely that the BISE algorithm will treat the brief drop in NDVI as an aberration and the $NDVI_{ratio}$ will not drop below 0.5. For this reason, we included the $CRIT_{next}$ requirement that the next week, as well as the previous month, must be dry in order to induce offset. In this manner, regrowth is allowed to occur, continuing the main growing season. It is likely that, as found by *Pitt and Wikeem* [1990], soil moisture is a strong control of growth in many grasslands. For this research, we chose to employ simple, commonly available meteorological data to drive the models. Further research focusing on grasslands should investigate incorporating soil moisture stresses in (8) and (9). In some cases, no drought is encountered, and the growing season is curtailed by cold weather.

While Figure 8 indicates errors in representation of interannual variability, the time period of these data is too short to draw substantive conclusions. When Figure 9 is considered, though, it is apparent that the DBF onset model accurately captures most of the observed relative variation in large-scale field phenology observations. It is critical to note that we are concerned with the relative, not the absolute accuracy of the model. Absolute temporal differences are due to several factors: understory species initiate growth earlier than overstory species; lilac is an early growing understory species; and most importantly, the model predicts 0.5 $NDVI_{ratio}$ for the entire ecosystem, not first leaf for a particular species. This is the most extensive test of the model's interannual accuracy we can envisage within the scope of this research. Further tests should seek out more extensive data sets covering grasses and DBF offset which may exist in Europe.

All methodologies for detecting phenological development from satellites are somewhat arbitrary. We selected a threshold method which defines onset and offset as the period of greatest increase and decrease in NDVI because of its relevance for ecological models, many of which initiate growth at full photosynthetic activity, a condition inconsistent with budburst. Additionally, we chose to include the field observations from the Harvard Forest and Hubbard Brook (Figure 3). These data, while limited to northeastern DBF and thus certainly not a complete validation, are extremely valuable and suggest that (1) observed onset occurs after initial leaf expansion and that (2) offset is observed between 15 and 50% leaf fall. In essence, an $NDVI_{ratio}$ of 0.5 corresponds to a considerably smaller canopy at onset than at offset. Photosynthetic activity, though, is likely to be similar at onset and offset. *Sellers* [1985] observed that in well watered canopies, vegetation indices are correlated with photosynthetic activity. Further, in deciduous canopies, where leaf life span is less than 1 year, photosynthetic capacity is highest in newly expanded, high-chlorophyll leaves and much reduced in older, low-chlorophyll leaves approaching abscission [*Reich et al.*, 1995]. Thus greatly different percent canopy cover may be associated with the same photosynthetic potential. Consequently, the models are probably best considered as predictors of physiological, not structural, conditions. Unfortunately, no similar data were available for grassland sites, and similar canopy conditions may or may not exist at 0.5 $NDVI_{ratio}$.

Model development for desert and crop biomes was beyond the scope of this study. In desert areas, large amounts of bare soil of varying brightness and accompanying remote sensing difficulties [*Huete et al.*, 1992] interact with low NDVI signals to produce temporal curves which are difficult to accurately interpret. Desert vegetation dynamics are strongly controlled by precipitation, which presents two problems for the methodology. First, the 10 km pixel resolution of the study precludes accurate representation of highly localized and irregular desert precipitation dynamics [*Barry and Chorley*, 1992]. Second, lowest prediction accuracy occurs in low station density desert areas. Low station density, combined with low precipitation, inevitably

lead to difficulties in predicting PPT. Errors in precipitation interpolation combined with remote sensing problems lead to problems in model development. Higher spatial resolution meteorology data and rigorous atmospheric and soil reflectance corrections to the NDVI (see soil adjusted vegetation index, *Huete et al.* [1992]) are critical precursors to development of accurate models for desert areas.

Crops present problems of a different sort. An underlying assumption of the methodology is that the vegetation being modeled is responding to environmental cues in a deterministic fashion. Crops do not behave this way. Rather, crops are genetically engineered species whose date of onset is controlled by planting, fertilization, and irrigation. Offset is not controlled by photoperiod, temperature, or precipitation; harvest dates are the primary factors. Onset and offset, then, are divergent from natural vegetation and cannot realistically be modeled with this methodology. Detailed information about crop type, irrigation, and farming methods are required to accurately model crop phenology.

Fortunately, for many of the envisioned users of this methodology, these liabilities may not be a major concern. Large-scale ecosystem simulation models often use potential vegetation maps as their landcover layer and ignore crops entirely. A recent global simulation estimated that only 11% of global net primary production (NPP) was produced in crop and desert biomes [*Prince et al.*, 1995]. In their study, *Prince et al.* [1995] found that forest and grass landcover types accounted for ~ 85% of estimated NPP, underscoring the usefulness of developing phenology equations for these biomes. For large-scale application, we suggest that use of the grassland phenology models in the place of crops should not lead to major errors in NPP estimates. Local applications requiring more accurate estimates of seasonality for deserts or crops will require detailed crop data including species, planting dates, fertilization, etc., and finer scale precipitation data.

Phenological response to climate change is a crucial research topic. Will the equations we developed change with global warming? This is a many-faceted, speculative question (see *Cannell and Smith* [1986], *Murray et al.* [1989], *Hänninen et al.* [1993], *Hänninen* [1994], and *Farnsworth et al.* [1995] for discussions of many issues). The general trend toward earlier growth shown in Figure 9 is consistent with the warming trend beginning in the late 1960s found by *Karl et al.* [1996], indicating a phenotypically plastic response in vegetation phenology. It is likely that as long as climatic conditions remain within normal limits of interannual variability, current models will adequately represent vegetation phenology. However, if climate changes beyond normal bounds, as is forecast by many GCMs, will vegetation reach a genetically controlled limit to phenological responses? Further, as climate changes at very long timescales, community composition is also likely to change, possibly introducing species with entirely different phenological controls. Our dynamic model formulation may be adequate for the current range of climatic conditions: as temperatures increase, so will the required summations. However, as outlined by *Lechowicz* [1984], phenological response is likely to be a combination of phylogenetic, evolutionary, and competitive interactions. Because of these complex controls, models incorporating biogeochemistry with community changes [*Friend et al.*, 1993] may be best suited to incorporate phenological dynamics with climate change.

This research was undertaken as an initial test of the methodology. It appears that satisfactory results can be obtained in the continental United States, but caution should be exercised before applying these equations beyond the midlatitude temperate zone in which they were developed. In contrast to midlatitude forests, for example, photoperiodic controls of offset in the northern boreal forest vary with latitude [*Heide*, 1974; *Hänninen et al.*, 1990; *Oleksyn et al.*, 1992]. In dry tropical areas, phenology is predominantly controlled by highly seasonal precipitation which can result in multiple growing seasons [*Reich*, 1994]. Seasonality in wet tropical forests may be due to poorly understood autogenic physiological controls [*Reich*, 1994] or optimization of radiation absorption [*Wright and van Schaik*, 1992]. Including these types of dynamics will enable global application of meteorological phenology models.

Conclusions and Applications

This research provides tested phenology models which may be used to predict the onset and offset of greenness for the deciduous broadleaf forest and grassland biomes in the temperate midlatitudes. Mean prediction errors are ~1 week, and maximum expected errors are 10-14 days. Tested with a long-term field data set, the deciduous broadleaf forest onset model accurately represents relative trends in interannual phenological variability. Major assumptions of the research are (1) point and species-level phenology models are scaleable to a regional level, (2) average NDVI curves for 20 km x 20 km observation sites represent regional greenness, (3) a threshold method can be used to assess onset and offset of greenness, (4) biome models developed from a limited number of sites are applicable to continental scales, and (5) a normalized step function adequately predicts thermal requirements of different grasslands.

There are two main potential uses for such models. First, large scale biogeochemical models require tested algorithms to regulate the timing and length of the growing season. Models developed here are applicable for two major deciduous landcovers, but much of the globe is covered by evergreen forests, evergreen shrublands, crops, and woodlands, modeling of which was beyond the scope of this research. More detailed satellite analysis, ancillary crop information, and field data will be required before tested algorithms are available for all major global landcovers. Second, research testing impacts of climate change scenarios on vegetation dynamics [*VEMAP*, 1995], accurate phenology models may be used as sensitive monitoring tools to detect vegetation response to interannual climatic variability. For example, research using detailed and widely distributed meteorological records, sometimes dating back a century, could provide an understanding of vegetation dynamics free of the limitations imposed by sparse phenology records and the short history of satellite observations. In short, few other terrestrial biophysical processes present such an intimate, immediate, and detectable connection to climatic variability.

Acknowledgments. The authors wish to thank Nan Rosenbloom at the National Center for Atmospheric Research for providing meteorological data and Joy Hood at the Earth Resource Observation Systems for providing the daily AVHRR data set. Thanks to John O'Keefe for providing the Harvard Forest phenology records and to Hubbard Brook for their excellent web site (some data used in this publication was obtained by scientists of the Hubbard Brook Ecosystem Study; this publication has not been reviewed by those scientists. The Hubbard Brook Experimental Forest is operated and maintained by the Northeastern Forest Experimental Station, United States Department of Agriculture, Radnor, Pennsylvania). Mark Schwartz generously provided the lilac phenology data set. We also thank Joseph White for numerous insightful methodological suggestions, Hubert Hasenauer for statistical assistance, and Stephanie Best for editing the manuscript. Three anonymous reviewers provided helpful critiques. This research was funded in part by NASA grants NAS5-31368 and NAGW-4511.

References

Badhwar, G.D., Automatic corn-soybean classification using Landsat MSS data, I, Near-harvest crop proportion estimation, *Remote Sens. Environ.*, *14*, 15-29, 1984.

Baret, F., and G. Guyot, Potentials and limits of vegetation indices for LAI and APAR assessment, *Remote Sens. Environ.*, *35*, 161-173, 1991.

Barry, R.G., and R.J. Chorley, *Atmosphere, Weather, and Climate*, pp. 219-220, Routledge, New York, 1992.

Bristow, K.L., and G.S. Campbell, On the relationship between incoming solar radiation and daily maximum and minimum temperature, *Agric. For. Meteorol*, *31*, 159-166, 1984.

Brown, D.M., Corn yield responses to irrigation, plant population and nitrogen in a cool, humid climate, *Can. J. Plant Sci.*, *66*, 453-464, 1986.

Burgan, R.E., and R.A. Hartford, Monitoring vegetation greenness with satellite data, *Gen. Tech. Rep. INT-297*, Intermt. Res. Stn., For. Serv., Ogden, Utah, 1993.

Burk, J.H., Phenology, germination, and survival of desert ephemerals in deep canyon, Riverside county, California, *Madrono*, *29*, 154-163, 1982.

Cannel, M.J.R., and R.I. Smith., Climatic warming, spring budburst and frost damage on trees, *J. Appl. Ecol.*, *23*, 177-191, 1986.

Caprio, J.M., Flowering dates, potential evapotranspiration and water use efficiency of *Syringa vulgaris* L. at different elevations in the western United States of America, *Agric. For. Meteorol.*, *63*, 55-71, 1993.

Curran, P.J., J.L. Dungan, and H.L. Gholz, Seasonal LAI in Slash Pine estimated with Landsat TM, *Remote Sens. Environ.*, *39*, 3-13, 1992.

Dickinson, C.E., and J.L. Dodd, Phenological patterns in the shortgrass prairie, *Am. Midland Nat.*, *96*, 367-378, 1976.

Doraiswamy, P.C., and D.R. Thompson, A crop moisture stress index for large areas and its application in the prediction of spring wheat phenology, *Agric. Meteorol*, *27*, 1-15, 1982.

Duncan, J., D. Stow, J. Franklin, and A. Hope, Assessing the relationship between spectral vegetation indices and shrub cover in the Jornada Basin, New Mexico, *Int. J. Remote Sens.*, *14*, 3395-3416, 1993.

Farnsworth, E.J., J. Núñez-Farfán, S.A. Careaga, and F.A. Bazzaz, Phenology and growth of three temperate forest life forms in response to artificial soil warming, *J. Ecol*, *83*, 967-977, 1995.

Fischer, A., A model for the seasonal variations of vegetation indices in coarse resolution data and its inversion to extract crop parameters, *Rem. Sens. Environ.*, *48*, 220-230, 1994.

French, N., and R.H. Sauer, Phenological studies and modeling in grasslands, in *Phenology and Seasonality Modeling*, edited by H. Leith, pp. 227-236, Springer-Verlag, New York, 1974.

Friend, A.D., H.H. Shugart, and S.W. Running, A physiology-based gap model of forest dynamics, *Ecology*, *74*, 792-797, 1993.

Fuller, D.O., and S.D. Prince, Rainfall and foliar dynamics in tropical southern Africa: Potential impacts of global climatic change on savanna vegetation, *Clim. Change*, *33*, 69-96, 1996.

Glassy, J.M., and S.W. Running, Validating diurnal climatology logic of the MT-CLIM model across a climatic gradient in Oregon, *Ecol. Appl.*, *4*, 248-257, 1994.

Goetz, S.J., and S.D. Prince, Remote sensing of net primary production in boreal forest stands, *Agric. For. Meteorol.*, *78*, 149-179, 1996.

Goward, S.N., C.J. Tucker, D.G. Dye, North American vegetation patterns observed with the NOAA advanced very high resolution radiometer, *Vegetation*, *64*, 3-14, 1985.

Goward, S.N., B. Markham, D.G. Dye, W. Dulaney, and J. Yang, Normalized difference vegetation index measurements from the advanced very high resolution radiometer, *Rem. Sens. Environ*, *35*, 257-277, 1991.

Hänninen, H., Effects of temperature on dormancy release in woody plants: Implications of prevailing models, *Silva Fenn.*, *21*, 279-299, 1987.

Hänninen, H., Effects of climate change on trees from cool and temperate regions: An ecophysiological approach to modelling of bud burst phenology, *Can. J. Bot.*, *73*,183-199, 1994.

Hänninen, H., R. Häkkinen, P. Hari, and V. Koski, Timing of growth cessation in relation to climatic adaptation of northern woody plants, *Tree Physiol.*, *6*, 29-39, 1990.

Hänninen, H., S. Kellomäki, K. Laitinen, B. Pajari, and T. Repo, Effect of increased winter temperature on the onset of height growth of Scots pine: A field test of a phenological model, *Silva Fenn.*, *27*, 251-257, 1993.

Hari, P., and R. Häkkinen, The utilization of old phenological time series of budburst to compare models describing annual cycles of plants, *Tree Physiol.*, *8*, 281-287, 1991.

Heide, O.M., Growth and dormancy in Norway Spruce ecotypes (*Picea abies*), I, Interaction of photoperiod and temperature, *Physiol. Plant.*, *30*, 1-12, 1974.

Henderson-Sellers, A., Z.-L. Yang, and R.E. Dickinson, The project for intercomparison of land-surface parameterization schemes, *Bull. Am. Meteorol. Soc.*, *74*, 1335-1349, 1993.

Hicken, R.P., and M.T. Vittum, The importance of soil and air temperature in spring phenoclimatic modeling, *Int. J. Biometeorol*, *20*, 200-206, 1976.

Hodges, T., and V. French, Soyphen: Soybean growth stages modeled from temperature, day length, and water availability, *Agron. J.*, *77*, 500-505, 1985.

Holben, B.N., Characteristics of maximum-value composite images from temporal AVHRR data, *Int. J. Remote Sens.*, *7*, 1417-1434, 1986.

Hood, J.J., Advanced very high resolution radiometer validation data set, paper presented at the 25th International Symposium, Remote Sensing and Global Environmental Change, Graz, Austria, April 4-8, 1993.

Huete, A.R., A soil-adjusted vegetation index (SAVI), *Remote Sens. Environ.*, *25*, 295-309, 1988.

Huete, A.R., G. Hua, J. Qi, A. Chehbouni, and W.J.D.G. Van Leeuwem, Normalization of multidirectional red and NIR reflectances with the SAVI, *Remote Sens. Environ.*, *40*, 1-20, 1992.

Hungerford, R.D., R.R. Nemani, S.W. Running, and J.C. Coughlan, MTCLIM: A mountain microclimate simulation model, *USDA For. Serv. Res. Pap. INT-414*, Intermt. For. and Range Exp. Stn., U. S. Dep. Agric., For. Serv., Ogden, Utah, 1989.

Hunter, A.H., and M.J. Lechowicz, Predicting the timing of budburst in temperate trees, *J. Appl. Ecol.*, *29*, 597-604, 1992.

Janecek, A., G. Benderoth, M.K.B. Lüdeke, J. Kindermann, and G.H. Kohlmaier, Model of the seasonal and perennial carbon dynamics in deciduous-type forests controlled by climatic variables, *Ecol. Modell.*, *49*, 101-124, 1989.

Justice, C.O., J.R.G. Townshend, B.N. Holben, and C.J. Tucker, Analysis of the phenology of global vegetation using meteorological satellite data, *Int. J. Remote Sens*, *6*, 1271-1318, 1985.

Kaduk, J., and M. Heimann, A prognostic phenology model for global terrestrial carbon cycle models, *Clim. Research*, *6*, 1-19, 1996.

Karl, T.R., R.W. Knight, D.R. Easterling, R.G. Quayle, Indices of climate change for the United States, *Bull. Am. Meteorol. Soc.*, *77*, 279-292, 1996.

Kemp, P.R., Phenological patterns of Chihuahuan desert plants in relationship to the timing of water availability, *J. Ecol.*, *71*, 427-436, 1983.

Kikuzawa, K., Leaf phenology as an optimal strategy for carbon gain in plants, *Can. J. Bot.*, *73*, 158-163, 1995.

Kogan, F.N., Droughts of the late 1980's in the United States as derived from NOAA polar-orbiting satellite data, *Bull. Am. Meteorol. Soc.*, *76*, 655-668, 1995.

Koski, V., and J. Selkäinaho, Experiments on the joint effect of heat sum and photoperiod on seedlings of *Betula pendula*, *Commun. Inst. For. Fenniae*, *105*, 1-34, 1982.

Kramer, K., Selecting a model to predict the onset of growth of *Fagus sylvatica*, *J. Appl. Ecol.*, *31*, 172-181, 1994.

Lavender, D.P., Environment and shoot growth of woody plants, *Res. Pap. 45*, For. Res. Lab, Sch. of For., Oreg. State Univ., Corvallis, 1981.

Lechowicz, M.J., Why do temperate deciduous trees leaf out at different times? Adaptation and ecology of forest communities, *Am. Nat., 124*, 821-842, 1984.

Lindsay, A.A., and J.E. Newman, Uses of official weather data in spring time-temperature analysis of an Indiana phenological record, *Ecology, 37*, 812-823, 1956.

Lloyd, D., A phenological classification of terrestrial vegetation cover using shortwave vegetation index imagery, *Int. J. Remote Sens., 11*, 2269-2279, 1990.

Lomas, J., and H. Herrera, Weather and rice yield relationships in tropical Costa Rica, *Agric. For. Meteorol., 35*, 133-151, 1985.

Loveland, T.R., J.W. Merchant, D.O. Ohlen, and J.F. Brown, Development of a land-cover characteristics database for the conterminous United States, *Photogramm. Eng. Remote Sens., 57*, 1453-1463, 1991.

Markon, C.J., M.D. Fleming, and E.F. Binnian, Characteristics of vegetation phenology over the Alaskan landscape using AVHRR time-series data, *Polar Rec., 31*, 179-190, 1995.

Monson, R.K., and G.J. Williams III, A correlation between photosynthetic temperature adaptation and seasonal phenology patterns in the shortgrass prairie, *Oecologia, 54*, 58-62, 1982.

Mueggler, W.F., Variation in production and seasonal development of mountain grasslands in western Montana, *USDA Forest Service Research Paper INT-316*, Intermt. For. and Range Exp. Stn., U.S. Dep. of Agric., For. Serv., Ogden, Utah, 1983.

Murray, M.B., M.J.R. Cannel, and R.I. Smith, Date of budburst of fifteen tree species in Britain following climatic warming, *J. Appl. Ecol., 26*, 693-700, 1989.

Myneni, R.B., and D.L. Williams, On the relationship of FAPAR and NDVI, *Remote Sens. Environ., 49*, 200-211, 1994.

Nemani, R.R. and S.W. Running, Satellite monitoring of global land cover changes and their impact on climate, *Clim. Change, 31*, 395-413, 1995.

Nooden, L.D., and J.A. Weber, Environmental and hormonal control of dormancy release in terminal buds of plants, in *Dormancy and Development Arrest,* edited by M. E. Clutter, pp. 222-268, Academic, San Diego, Calif., 1978.

O'Leary, G.J., D.J. Connor, and D.H. White, Effect of sowing time on growth, yield and water use of rain-fed wheat in the Wimmera, Vic, *Aust. J. Agric. Res., 36*, 187-196, 1985.

Oleksyn, J., M.G. Tjoelker, and P.B. Reich, Growth and biomass partitioning of populations of European *Pinus sylvestris* L. Under simulated 50E and 60EN day lengths: Evidence for photoperiodic ecotypes, *New Phytol., 120*, 561-574, 1992.

Pitt, M.D., and B.M. Wikeem, Phenological patterns and adaptations in an *Artemesia/Agropyron* plant community, *J. Range Manage., 43*, 350-367, 1990.

Prince, S.D., S.J. Goetz, and S.N. Goward, Monitoring primary productivity from Earth observing satellites, *Water Air Soil Pollut., 82*, 509-522, 1995.

Ram, J., S.P. Singh, and J.S. Singh, Community level phenology of grassland above treeline in central Himalaya, India, *Arct. Alp. Res.*, 20, 325-332, 1988.

Réaumur, R.A.F. de., Observations du thérmomètre, faites à Paris pendant l'année 1735, comparés avec celles qui ont été faites sous la ligne, à l'Isle de France, à Alger et en quelques-unes de nos isles de l'Amérique. *Mém. Acad. des Sci., Paris, 1735*, 545, 1735.

Reed, B.C., J.F. Brown, D. VanderZee, T.R. Loveland, J.W. Merchant, and D.O. Ohlen, Measured phenological variability from satellite imagery, *J. Veg. Sci., 5*, 703-714, 1994.

Reich, P.B., Phenology of tropical forests: Patterns, causes, and consequences, *Can. J. Bot., 73*, 164-174, 1994.

Reich, P.B., T. Koike, S.T. Gower, and A.W. Schoettle, Causes and consequences of variation in conifer leaf life-span, in *Ecophysiology of Coniferous Forests*, edited by W. K. Smith and T. M. Hinckley, pp. 225-254, Academic, San Diego, Calif., 1995.

Reynolds, M.R., Jr., Estimating the error in model predictions, *For. Sci., 30*, 454-469, 1984.

Running, S.W., and R.R. Nemani, Regional hydrologic and carbon balance responses of forests resulting from potential climate change, *Clim. Change, 19*, 349-368, 1991.

Running, S.W., Nemani, R.R. and Hungerford, R.D., Extrapolation of synoptic meteorological data in mountainous terrain and its use for simulating forest evapotranspiration and photosynthesis, *Can. J. For. Res., 17*, 472-483, 1987.

Schmidt, W.C., and J.E. Lotan, Phenology of common forest flora of the Northern Rockies--1928 to 1937, *USDA For. Serv. Res. Pap. INT-259*, Intermt. For. and Range Exp. Stn., United States Dep. of Agric., For. Serv., Ogden, Utah, 1980.

Schwartz, M.D., Phenology and springtime surface-layer change, *Mon. Weather Rev., 120*, 2570-2578, 1992.

Schwartz, M.D., Monitoring global change with phenology: The case of the spring green wave, *Int. J. Biometeorol, 38*, 18-22, 1994.

Schwartz, M.D., Spring index models: an approach to connecting satellite and surface phenology, *Phenol. Seas.*, in press, 1997.

Sellers, P.J., Canopy reflectance, photosynthesis and transpiration, *Int. J. Remote Sens., 6*, 1335-1372, 1985.

Sellers. P.J., et al., Comparison of radiative and physiological effects of doubled atmospheric CO_2 on climate, *Science, 271*, 1402-1406, 1996.

Sharifi, M.R., F.C. Meinzer, E.T. Nilsen, P.W. Rundel, R.A. Virginia, W.M. Jarrell, D.J. Herman, and P.C. Clark, Effect of manipulation of water and nitrogen supplies on the quantitative phenology of *Larrea tridentata* (creosote bush) in the Sonoran Desert of California, *Am. J. Bot., 75*, 1163-1174, 1988.

Sims, P.L., Grasslands, in *North American Terrestrial Vegetation,* edited by M. G. Barbour and W. D. Billings, pp. 265-286, Cambridge Univ. Press, New York, 1988.

Smit-Spinks, B., B.T. Swanson, and A.H. Markhart III, The effect of photperiod and thermoperiod on cold acclimation and growth of *Pinus sylvestris, Can. J. For. Res., 15*, 453-460, 1985.

Spanner, M.A., L.L. Pierce, D.L. Peterson, and S.W. Running, Remote sensing of temperate coniferous forests: The influence of canopy closure, understory vegetation and background reflectance, *Int. J. Remote Sens., 11*, 95-111, 1990a.

Spanner, M.A., L.L. Pierce, S.W. Running, and D.L. Peterson, The seasonality of AVHRR data of temperate coniferous forests: Relationships with leaf area index, *Remote Sens. Environ., 33*, 97-112, 1990b.

Sparks, T.H., and P.D. Carey, The response of species to climate over two centuries: An analysis of the Marsham phenological records, 1736-1947, *J. Ecol., 83*, 321-329, 1995.

Stone, M., Cross-validatory choice and assessment of statistical predictions, *J. R. Stat. Soc. Ser. B, 36*, 111-133, 1974.

Stone, T.A., S. Schlesinger, R.A. Houghton, and G.M. Woodwell, A map of the vegetation of South America based on satellite imagery, *Photogramm. Eng. Remote Sens., 60*, 541-551, 1994.

Thomson, A.J., and S.M. Moncrieff, Prediction of budburst in Douglas-fir by degree-day accumulation, *Can. J. For. Res., 12*, 448-452, 1981.

Thornton, P.E., S.W. Running, and M.A. White, Generating surfaces of daily meteorological variables over large regions of complex terrain, *J. Hydrol.*, in press, 1997.

Tucker, C.J., J.R.G. Townshend, and T.E. Goff, African land-cover classification using satellite data, *Science, 227*, 369-375, 1985.

Undersander, D.J. and S. Christiansen, Interactions of water variables and growing degree days on heading phase of winter wheat, *Agric. For. Meteorol., 38*, 169-180, 1986.

Valentine, H.T., Budbreak and leaf growth functions for modeling herbivory in some gypsy moth hosts, *For. Sci., 29*, 607-617, 1983.

Vegis, A., Dormancy in Higher Plants, *Annu. Rev. Plant Physiol, 5*, 185-222, 1964.

VEMAP Modeling Participants, Vegetation/ecosystem modeling and analysis project: Comparing biogeography and biogeochemistry models in a continental-scale study of terrestrial ecosystem responses to climate change and CO_2 doubling, *Global Biogeochem. Cycles, 9*, 407-437, 1995.

Viovy, N., and O. Arino, The best index slope extraction (BISE): A method for reducing noise in NDVI time-series, *Int. J. Remote Sens., 13*, 1585-1590, 1992.

White, J.D., and S.W. Running, Testing scale dependent assumptions in regional ecosystem simulation models, *J. Veg. Sci., 5*, 687-702, 1994.

Wright, S.J., and C.P. van Schaik, Light and the phenology of tropical trees, *Am. Nat., 143*, 192-199, 1992.

Yoder, B.J., and R.H. Waring, The normalized difference vegetation index of small Douglas-Fir canopies with varying chlorophyll concentrations, *Remote Sens. Environ., 49*, 81-91, 1994.

Zheng, D., E.R. Hunt, and S.W. Running, A daily soil temperature model based on air temperature and precipitation for continental applications, *Clim. Research, 2*, 183-191, 1993.

S.W. Running, P.E. Thornton, M.A. White, Numerical Terradynamic Simulation Group, School of Forestry, University of Montana, Missoula, MT 59812. (e-mail: swr@ntsg.umt.edu; peter@ntsg.umt.edu; mike@ntsg.umt.edu)

(Received October 24, 1996; revised January 31, 1997; accepted February 3, 1997)