INTERNATIONAL JOURNAL OF CLIMATOLOGY
Int. J. Climatol. **30**: 1614–1626 (2010)
Published online 27 August 2009 in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/joc.2008

Intercomparing multiple measures of the onset of spring in eastern North America†

Mark D. Schwartz* and Jonathan M. Hanes
Department of Geography, University of Wisconsin-Milwaukee, P.O. Box 413, Milwaukee, WI 53201-0413, USA

ABSTRACT: Measuring the onset of deciduous tree leaf flush and subsequent development during the spring season in temperate climates can be accomplished using multiple ground and satellite-based techniques. Although all these measurements are valid (i.e. record a real characteristic related to plant development), they typically are poorly inter-related due to incompatible levels of spatial representation and differing methodologies. Given recent and likely future impacts of climate change on spring leaf development, the need to reconstruct past patterns, and the lack of standardised vegetation change measurements around the world, more work is needed to determine the relationships among the various measures, and the degree to which they may serve as substitutes for each other. In this article, we use observations and measurements at two phenology 'super-sites' in eastern North America and four other supporting sites to evaluate the relationships among multiple spring leaf development measures, and explore strategies to standardise their intercomparison. The results show infrequent significant correlations among 10 satellite-derived 'start of season' (SOS) measures (which suggests they are often not detecting the same phenomena), along with more common significant correlations among six ground phenology measures. However, when ground phenology and satellite-derived SOS are compared, there are few significant correlations, even at sites with extensive native species phenology available. Modelled phenology, based on daily temperature data (Spring Indices First Bloom date) does as well as any of the direct native species measures, and is well suited to facilitate intercomparisons. In order to effectively compare ground-based and satellite-derived SOS measures, approaches that use limited numbers of individual plants face considerable challenges. Given that satellite-derived measures are areal and at a scale of 250 m and larger, we suggest collecting ground phenology data at the same areal scale in order to make effective comparisons. Copyright © 2009 Royal Meteorological Society

KEY WORDS onset of spring; phenology; indices; remote sensing; native species

Received 23 March 2009; Revised 15 July 2009; Accepted 25 July 2009

1. Introduction

Many methods are used to measure the onset of deciduous tree leaf flush and subsequent development during the spring season in temperate climates (Badeck *et al*., 2004; Cleland *et al*., 2007). These can be grouped into at least three major categories: (1) satellite-derived measurements of vegetative development (derived from reflectance values); (2) instrumental measurements of latent-sensible heat energy balance and carbon dioxide flux (related to plant photosynthesis and transpiration) and (3) conventional phenological observations of numerous native and indicator plants (both cloned and normal types) recorded by human observers (Schwartz and Crawford, 2001; Menzel *et al*., 2006; Schwartz *et al*., 2006). All of these measurements are valid (i.e. record a real characteristic related to plant development), but the three categories look at different features, typically have incompatible levels of spatial representation, and thus are often poorly inter-related (Schwartz *et al*., 2002). Furthermore, differences between measures within the categories can be equally great. Satellite-derived 'start of season' (SOS) methodologies target different processes and produce widely varying results (White *et al*., (in press)). With energy and carbon flux values, meaningful 'change points' (comparable to SOS timing) in the continuous data sequences may be difficult to identify. Even among conventional phenological observations, the differential influences of genetics and weather/climate (typically temperatures) can make direct comparison between sites difficult.

Given recent and likely future impacts of climate change on spring leaf development, the need to reconstruct past patterns, and the lack of standardised vegetation change measurements around the world, more intercomparison work is needed to determine the relationships among the various measures, and the degree to which they may serve as substitutes for each other. In this article, we use measurements at two phenology 'super-sites' in eastern North America (where large numbers of plants species are monitored together with detailed recordings of atmospheric and remotely sensed data), and four other supporting sites (with a lesser mix of

* Correspondence to: Mark D. Schwartz, Department of Geography, University of Wisconsin–Milwaukee, P.O. Box 413, Milwaukee, WI 53201-0413, USA. E-mail: mds@uwm.edu
† This article was published online on 27 August 2009. An error was subsequently identified. This notice is included in the online and print versions to indicate that both have been corrected 25 September 2009.

Figure 1. Study sites in eastern North America.

measures) to evaluate the relationships among multiple spring vegetative growth measures, and explore strategies to improve their intercomparison.

2. Materials and methods

2.1. Remote sensing

We obtained Moderate Resolution Imaging Spectroradiometer (MODIS, 1 km resolution 16-day composite) Normalised Difference Vegetation Index (NDVI) and Enhanced Vegetation Index (EVI) data (MOD13A2, from the sensor aboard the NASA TERRA satellite), for 7 km × 7 km (49-pixel) areas centred on three sites from January 2000 to December 2006 (Oak Ridge National Laboratory ORNL Distributed Active Archive Center DAAC, 2007). These locations (Figure 1) were Harvard Forest, Massachusetts (42.532 °N, 72.188 °W), Park Falls/WLEF, Wisconsin (Park Falls, 45.946 °N, 90.272 °W), and UW-Milwaukee Field Station, Wisconsin (UWM Field Station, 43.387 °N, 88.023 °W). The MODIS land cover product (also available from the ORNL DAAC web site) classified these sites as follows: Harvard Forest [76% mixed forest and 24% deciduous broadleaf forest (DBF)]; Park Falls (98% mixed forest and 2% DBF) and UWM Field Station (47% croplands, 43% cropland/natural vegetation mosaic, 6% mixed forest and 4% DBF).

Among the five satellite-derived SOS techniques represented in this study (described below), there are two different 'options' that in various combinations produce four distinct methodological approaches (Table I, see also Reed *et al.* (2003) for more comparative information on SOS methods). These two options are: (1) relative time on the NDVI or EVI generically referred to as Vegetation Indices (VI) curve used to indicate SOS, with alternatives being either the 'first upturn' (the first sustained upward movement, which would be earlier), or 'midpoint' (the middle point between the lowest and highest VI values, which would be later); and (2) approach to representing the VI curve, with alternatives being either 'actual data' (actual data values with attempts to remove all erroneous data), or 'curve fitting' (the actual data are used to produce a best fit statistical distribution, such a Gaussian or Logistic).

The VI data were provided for processing to collaborators using the Delayed Moving Average (courtesy of B. Reed) and Seasonal Midpoint (courtesy of M. White) methods and then returned to us as SOS dates for both EVI and NDVI each spring. The Delayed Moving Average SOS method (summarised in Table I) identifies when the VI exhibits a rapid, sustained increase that signals the onset of significant photosynthetic activity by employing a 'backward-looking' or delayed moving average. VI data values are compared with the average of the previous (user-defined) *n* VI observations to identify departures from an established trend. The delayed moving average value serves as a predicted value with which the real VI values are compared. A trend change is detected where the VI value departs from (becomes greater than) the value of the moving average, such as when low VI values

Table I. Satellite-derived start of season (SOS) measures compared in this study.

Name	Source data*	SOS relative timing	VI curve approach	Resolution	Primary reference
Boston University	NBAR	First upturn	Curve fitting	1 km	Zhang *et al.* (2003)
Delayed Moving Average	EVI	First upturn	Actual data	1 km	Reed *et al.* (1994)
Delayed Moving Average	NDVI	First upturn	Actual data	1 km	Reed *et al.* (1994)
Logistic Curve	NBAR	Midpoint	Curve fitting	500 m	Fisher *et al.* (2007)
Modified TIMESAT	EVI	First upturn	Curve fitting	250 m	Tan *et al.* (2008)
Modified TIMESAT	NDVI	First upturn	Curve fitting	250 m	Tan *et al.* (2008)
Modified TIMESAT	EVI	First upturn	Curve fitting	500 m	Tan *et al.* (2008)
Modified TIMESAT	NDVI	First upturn	Curve fitting	500 m	Tan *et al.* (2008)
Seasonal Midpoint	EVI	Midpoint	Actual data	1 km	White *et al.* (2002)
Seasonal Midpoint	NDVI	Midpoint	Actual data	1 km	White *et al.* (2002)

* The MODIS bi-directional reflectance distribution function (BRDF)/Albedo algorithm generates one nadir BRDF-adjusted reflectance (NBAR) for each MODIS land band. The other data sources are Normalised Difference Vegetation Index (NDVI) and Enhanced Vegetation Index (EVI).

are predicted by the moving average, but actual VI values are higher. The date of this departure is labeled as the SOS [see Reed *et al.* (1994) and Schwartz *et al.* (2002) for additional information about this technique].

The Seasonal Midpoint method (summarised in Table I) first selects the annual cloud-screened minimum and maximum VI for every pixel, and then calculates the midpoint between them. This is repeated for every year in the study. The average of these values is the Seasonal Midpoint VI. The following steps are then completed for every pixel: (1) the complete annual time series is obtained and screened to remove cloud-contaminated pixels; (2) a spline curve is fit to the data to interpolate VI to daily values and (3) the Seasonal Midpoint VI is used as a threshold to identify the Seasonal Midpoint SOS date [see White *et al.* (1997, 2002) and Schwartz *et al.* (2002) for additional information about this technique].

Three additional and comparable SOS time series for these same three sites were also obtained directly from researchers who developed and applied their techniques to MODIS VI data: the 'Boston University' (courtesy of M. Friedl), 'Logistic Curve' (courtesy of J. Fisher), and 'modified TIMESAT' (courtesy of B. Tan and J. Morisette) methods (summarised in Table I). The Boston University method fit logistic functions to annual time series of MODIS EVI measurements for 49 pixels (1 km resolution) and identified SOS dates based on the times at which the rate of change in the logistic functions exhibited local maxima during spring (Zhang *et al.*, 2003). The Logistic Curve method fit sigmoid logistic functions to annual time series of MODIS NDVI measurements for 196 pixels (500 m resolution) and defined SOS dates based on when the fitted NDVI values reached half of the annual maximum values during spring (Fisher and Mustard, 2007). The Modified TIMESAT method initially used TIMESAT software (Jonsson and Eklundh, 2004) to fit asymmetric Gaussian curves to annual time series of gap-filled MODIS NDVI and EVI measurements for 784 pixels (250 m resolution) and 196 pixels (500 m resolution). SOS dates were determined based on when the third derivative of the asymmetric Gaussian curves exhibited local maxima during the spring (Gao *et al.*, 2008; Tan *et al.*, 2008). After obtaining all these SOS data, we calculated medians from among the pixel values at each site every year for each of the 10 SOS techniques (Table I, some were not available for every year due to processing limitations), and these yearly median values were used for all further comparison analyses.

2.2. Phenological model output and surface phenology

Extensive conventional (surface) phenological observations are available at UWM Field Station and Harvard Forest. These include bud burst dates (defined as when 50% of buds have recognisable leaves; reported dates are the average among observed plants of the same species) for native tree and shrub species (taken since 1990 at Harvard Forest, and at both sites using the same protocol since 2000, see Table II).

The timing of all measures in this study were standardised (rendered as positive or negative departures in days) at each site/year relative to corresponding Spring Indices (SI) First Bloom dates (previously shown to be best correlated among SI output with native tree species bud burst), which were calculated using 1.5 m level daily maximum–minimum air temperatures, available from co-located or nearby (in the case of Morgan-Monroe) COOP weather stations (Schwartz *et al.*, 2006, Table III). The extensively validated SI phenological models (which can be calculated for any site from daily maximum and minimum temperatures) are composite indicators of spring phenological variability for clonal lilac (*Syringa chinensis*) and honeysuckle (*Lonicera tatarica, L. korolkowii*) and represent the response of temperature-sensitive and non-water limited plant species to seasonally integrated changes in temperature (Schwartz, 1997; Schwartz and Reiter, 2000; Schwartz *et al.*, 2006). Thus, SI simulations are most applicable to forest trees and shrubs, crops planted in temperate regions with adequate rainfall, or in irrigated temperate dry regions; outputs are valid in all areas where lilac and honeysuckle plants would receive sufficient chilling in the cold season and sufficient heat in the warm season to survive (survival is theoretical because the models do not address potential summer water stress or winter cold mortality). Although SI does

Table II. Native tree species observations and statistical cluster membership by site*.

Latin name	Common name	UWM Field Station species/Cluster #	Harvard Forest species/Cluster #
Acer saccharum	**Sugar maple**	Yes/1	Yes/2
Betula alleghaniensis	Yellow birch	Yes/1	Yes/1
***Crataegus* spp.**	**Hawthorn**	Yes/1	Yes/2
Fagus grandifolia	Beech	Yes/1	Yes/1
Fraxinus americana	**White ash**	Yes/1	Yes/1
Hamamelis virginiana	**Witch hazel**	Yes/1	Yes/2
Betula papyrifera	Paper birch	Yes/2	Yes/2
Cornus alternifolia	Pagoda dogwood	Yes/2	Yes/3
Populus tremuloides	Quaking aspen	Yes/2	Yes/2
Quercus rubra	Red oak	Yes/2	Yes/1
Acer pensylvanicum	Striped maple	No	Yes/1
Acer rubrum	Red maple	No	Yes/1
Amelanchier species	Shadbush	No	Yes/2
Betula lenta	Black birch	No	Yes/1
Quercus alba	White oak	No	Yes/1
Quercus velutina	Black oak	No	Yes/1
Carya cordiformis	Yellowbud hickory	Yes/1	No
Carya ovata	Shagbark hickory	Yes/1	No
Fraxinus nigra	Black ash	Yes/1	No
Tilia americana	Basswood	Yes/1	No
Ulmus americana	American elm	Yes/1	No
Carpinus caroliniana	Musclewood	Yes/2	No
Cornus racemosa	Grey dogwood	Yes/2	No
Cornus sericea	Red osier dogwood	Yes/2	No
Larix laricina	Tamarack	Yes/2	No
Ostrya virginiana	Ironwood	Yes/2	No
Prunus serotina	Black cherry	Yes/2	No
Dirca palustris	Leatherwood	Yes/3	No
Viburnum lentago	Nannyberry	Yes/3	No

* Four highest species correlations with SI First Bloom date at both sites shown in bold.

not reproduce all the details of multi-species phenology data at any site, or the specific phenology of some types of plants, the models provide a baseline assessment of general phenological response over a standard period, supplying a needed context for evaluating and comparing other phenological measures. Furthermore, SI is expected to be robust under future conditions where climate departs significantly from the historical mean, as the models are optimised for continental-scale applications and include training data bounded by the USA Northeast, North Carolina, Oklahoma, and North Dakota (approximately 35 to 49°N and −104 to −68 °W, Hayhoe *et al*., 2007).

We used three alternative pre-processing techniques to prepare the native species phenological data for comparison with the other measures. First, four species (*Acer saccharum, Crataegus* sp., *Fraxinus americana* and *Hamamelis virginiana*) were identified that had highest correlated bud bursts with SI First Bloom date among the species available at both Harvard Forest and UWM Field Station (Table II). We calculated the average bud burst date of these four species. Second, hierarchical clustering was used in SPSS with a 3 and 4 cluster solution to group the phenological responses (bud burst dates) of the species at each site. Finally, we used tree relative frequency information at each site to weight the bud burst information when averaging all species responses (Harvard Forest, personal communication from John O'Keefe; UWM Field Station, Dunnum, 1972). These percentages are much generalised (and in the case of UWM Field Station only address the roughly half of the landscape that is forested), but we considered them sufficient to test whether weighting ground phenology according to species distribution would improve correlations with other measures. The general species percentage distributions at each site are: (1) Harvard Forest (30% coniferous, 28% *Quercus rubra* and *Q. velutina*, 18% *Acer rubrum*, 8% *Betula alleghaniensis* and *B. lenta* and 16% other deciduous species); and (2) UWM Field Station (42% *A. saccharum*, 13% *Fagus americana*, 12% *Fraxinus americana*, 11% *Tilia americana*, 11% *Ostrya virginiana* and 11% other species).

2.3. Latent/sensible heat and carbon flux

Latent/sensible heat (energy flux), and carbon flux data at 30-min intervals were obtained from four eddy covariance flux tower sites (Figure 1 and Table III): Harvard Forest and Park Falls (locations previously described, Wofsy *et al*., 1993; Davis *et al*., 2003; Ricciuto *et al*., 2008), plus Morgan-Monroe State Forest, IN (39.323 °N,

Table III. Spring Indices First Bloom dates (as days after December 31) used to standardise (render as positive or negative departures in days) all other phenology measures, and flux data availability by year at each study site (blank cells indicate no data).

Year	Harvard Forest	Lamont	Morgan-Monroe	Oak Ridge	Park Falls	UWM Field Station
1992	*					
1993	132*	*				
1994	142*	89*				
1995	143*	83*		*		
1996	146*	97*		102*		
1997	149*	87*		76*	*	
1998	134*	93*	*	86*	132*	
1999	139*		115*	98*	135*	
2000	137*		111*	81*	135	127
2001	135				134	129
2002	135				151	140*
2003	140				141	139*
2004	137				145	139*
2005	142				142	138*
2006	136				136	131*

* Flux data available in this year.

86.413 °W, Schmid *et al.*, 2000) and Oak Ridge, Tennessee (35.931 °N, 84.332 °W, Wilson and Baldocchi, 2001). We calculate latent and sensible heat values at UWM Field Station (location previously described) using the Bowen ratio technique. At the Lamont, OK ARM Southern Great Plains central facility site (Lamont, 36.605 °N, 97.488 °W, Figure 1), latent and sensible heat were also obtained (Bowen ratio technique, Schwartz and Crawford, 2001). Carbon flux (Net Ecosystem Exchange, NEE) values were available for comparison at Harvard Forest, Park Falls, Morgan-Monroe, and Oak Ridge (Table III). All energy and carbon flux data were converted to average 'daytime' values for each day at every site, by excluding data for times when net radiation had a value of zero or less.

An approach was needed to objectively determine dates in the daily sequence of flux dates that would be comparable to SOS and conventional phenological event dates, for comparison with these other measures. In order to facilitate this analysis, latent and sensible heat values were combined by simply subtracting the average daily sensible heat value from the average daily latent heat value (L-S difference). Although more sophisticated combination techniques are certainly possible, we concluded that this simple difference was sufficient to detect changing relationships among sites. We next used an objective technique to identify slope break-points within 30 days of average SI First Bloom date at each site for sequences of: (1) greatest positive slope of the L-S difference (based on 6 days following each date), signifying the first sustained upward growth of latent heat values in spring and (2) greatest negative slope of NEE (based on the 30 days following each date), signifying the first downward draw of carbon from the atmosphere in spring. After these analyses were conducted, the resulting break-point dates from each site were correlated with the average SI First Bloom dates for each site to look for relationships.

After these initial steps, the remaining analyses were to simply calculate Pearson r correlations among all the measures, and prepare two sets of graphical displays: (1) showing comparative interannual variations in all the measures at Harvard Forest, Park Falls, and UWM Field Station and (2) use box plots for each measure, standardised to SI First Bloom date at Harvard Forest and UWM Field Station in order to visualise their relative timing.

3. Results

The timing standardisation to SI First Bloom dates allows the relative sequence of spring development as represented by ground and satellite-derived measurements of phenology to be visually compared at each site and between the Harvard Forest and UWM Field Station sites (Figures 2 and 3). Within this context, the timing of different measures varies between sites, with three of the four native species events at UWM Field Station proceeding later in phenological time than at Harvard Forest, and earliness/lateness reversed for SOS measures (with the exception of Logistic Curve, which is only slightly different).

Comparisons of the phenological measures at the Harvard Forest, UWM Field Station and Park Falls sites are shown in Tables IV–VI and Figures 4–6. Considering first the satellite-derived SOS measures, none of the 10 are consistently inter-related, with only Delayed Moving Average using EVI data, Delayed Moving Average using NDVI data, and Modified TIMESAT at 500 m resolution using EVI data showing significant correlations with

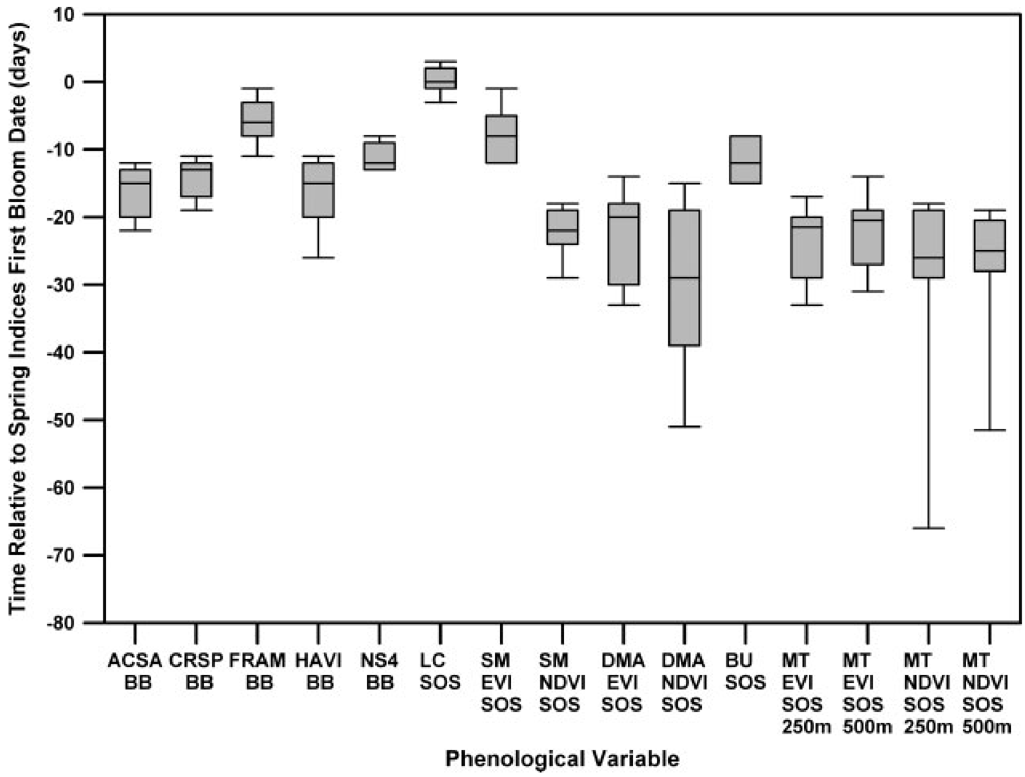

Figure 2. Harvard Forest box plot array of phenological variables, relative to Spring Indices First Bloom date (SIFB). ACSA BB (*Acer saccharum* [Sugar maple] Bud Burst); CRSP BB (*Crataegus* ssp. [Hawthorn] Bud Burst); FRAM BB (*Fraxinus americana* [White ash] Bud Burst); HAVI BB (*Hamamelis virginiana* [Witch hazel] Bud Burst); NS4 BB (Native Species Average Bud Burst date among four species with highest correlation to SIFB); LC SOS (Logistic Curve); SM EVI SOS (Seasonal Midpoint using EVI data); SM NDVI SOS (Seasonal Midpoint using NDVI data); DMA EVI SOS (Delayed Moving Average using EVI data); DMA NDVI SOS (Delayed Moving Average using NDVI data); BU SOS (Boston University); MT EVI SOS 250 m (Modified TIMESAT at 250 m resolution using EVI data); MT EVI 500 m (Modified TIMESAT at 500 m resolution using EVI data); MT NDVI 250 m (Modified TIMESAT at 250 m resolution using NDVI data) and MT NDVI 500 m (Modified TIMESAT at 500 m resolution using NDVI data).

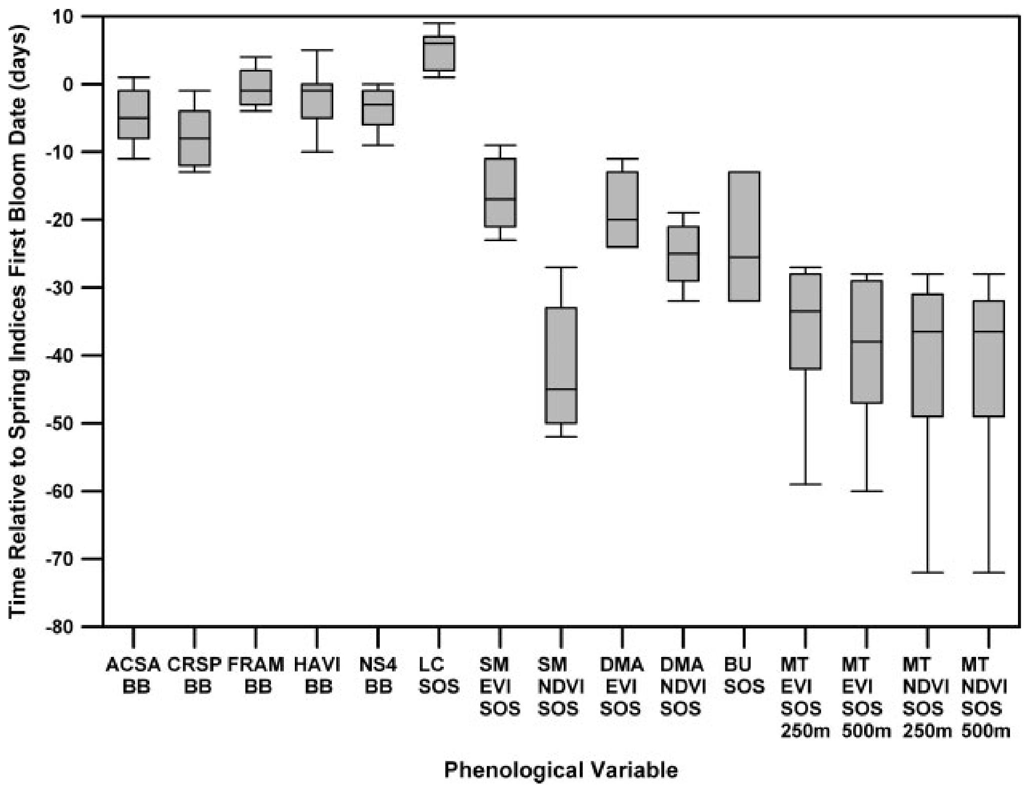

Figure 3. UWM Field Station box plot array of phenological variables, relative to Spring Indices First Bloom date (SIFB, see Figure 2 for label descriptions).

Table IV. Harvard Forest site Pearson r correlations ×100* (sig. @ 0.05 [α] level shown in bold).

	BU	DMAE	DMAN	LC	MT2E	MT2N	MT5E	MT5N	SME	SMN	SIFB	NS4	NSC1	NSC2	NSC3	NSW
BU	X															
DMAE	−17	X														
DMAN	−60	+68	X													
LC	+94	+50	+26	X												
MT2E	−54	**+88**	**+93**	+19	X											
MT2N	−86	+73	**+86**	−12	+74	X										
MT5E	−55	**+84**	**+90**	+17	**+99**	+68	X									
MT5N	−84	+76	**+88**	−07	+74	**+99**	+68	X								
SME	+72	+47	+35	**+81**	+29	−02	+30	+01	X							
SMN	−94	+62	**+86**	+44	+75	+66	+74	+67	+31	X						
SIFB	+82	+11	+19	+72	+01	−15	+01	−12	+37	+60	X					
NS4	+85	−05	−16	+65	−15	−55	−09	−54	+38	+26	**+82**	X				
NSC1	+94	+14	−07	**+84**	−15	−42	−11	−40	+62	+31	**+86**	**+92**	X			
NSC2	+44	−29	−31	+10	−15	−56	−05	−60	−13	+07	+46	**+79**	+50	X		
NSC3	+22	−35	−49	+01	−63	−30	−73	−26	−40	−35	+15	−05	+02	−15	X	
NSW	+92	+06	+03	+55	+08	−40	+19	−42	+56	+31	+62	**+88**	**+80**	+73	−50	X

* BU (Boston University); DMAE (Delayed Moving Average using EVI data); DMAN (Delayed Moving Average using NDVI data); LC (Logistic Curve); MT2E (Modified TIMESAT at 250 m resolution using EVI data); MT2N (Modified TIMESAT at 250 m resolution using NDVI data); MT5E (Modified TIMESAT at 500 m resolution using EVI data); MT5N (Modified TIMESAT at 500 m resolution using NDVI data); SME (Seasonal Midpoint using EVI data); SMN (Seasonal Midpoint using NDVI data); SIFB (Spring Indices First Bloom date); NS4 (Native Species Average Bud Burst date among four species with highest correlation to SIFB); NSC1 (Native Species Bud Burst Cluster 1); NSC2 (Native Species Bud Burst Cluster 2); NSC3 (Native Species Bud Burst Cluster 3) and NSW (Native Species Average Bud Burst date weighted by site species frequency estimate).

Table V. UW–Milwaukee Field Station site Pearson r correlations ×100 (sig. @ 0.05 [α] level shown in bold, see Table IV for label descriptions).

	BU	DMAE	DMAN	LC	MT2E	MT2N	MT5E	MT5N	SME	SMN	SIFB	NS4	NSC1	NSC2	NSC3	NSW
BU	X															
DMAE	−16	X														
DMAN	−04	**+97**	X													
LC	−25	**+79**	**+88**	X												
MT2E	−26	**+92**	**+88**	+80	X											
MT2N	−41	**+80**	+71	+65	**+94**	X										
MT5E	−16	**+91**	**+86**	+72	**+98**	**+95**	X									
MT5N	−37	**+83**	+74	+66	**+95**	**+99**	**+96**	X								
SME	+04	**+91**	**+89**	+63	+68	+48	+69	+54	X							
SMN	+78	−31	−09	+13	−40	−62	−49	−62	−11	X						
SIFB	−48	+72	+75	**+87**	+63	+57	+55	+56	+60	+07	X					
NS4	+04	**+91**	**+93**	**+82**	+79	+62	+76	+65	**+88**	+03	**+84**	X				
NSC1	−03	**+80**	+75	+63	+66	+70	+74	+71	+68	−35	+74	**+78**	X			
NSC2	+23	+55	+47	+08	+25	+02	+27	+08	**+78**	−12	+15	+56	+27	X		
NSC3	−04	+15	+08	−22	−17	−35	−21	−31	+44	−02	−21	+10	−26	**+82**	X	
NSW	−07	**+96**	**+97**	**+84**	**+81**	+65	+79	+68	**+94**	−06	**+82**	**+98**	**+79**	+57	+16	X

five or more (50%) of the other measures within at least one site. The correlations among satellite-derived measures are also weaker at Park Falls than the other two sites, with none of the 10 SOS from that site showing a significant correlation to more than three others.

We chose the three cluster solution for the native species bud burst grouping, as the four cluster solution did not add additional useful distinctions between groups (see cluster membership by species and site in Table II). Among the SI and five native species measures, four (all except Native Species Bud Burst Cluster 2 and Native Species Bud Burst Cluster 3) show significant correlations to at least two other ground phenology measures within at least one site. Harvard Forest shows noticeably less overall relationship between satellite-derived and ground phenology than the other two sites, with Logistic Curve/Native Species Bud Burst Cluster 1 being the only significant correlation (however, some pairings such as Logistic Curve/SI First Bloom date and Boston University/Native Species Bud Burst Cluster1 are just below the 0.05 [α] significance threshold). At the other two sites (UWM Field Station and Park Falls), Delayed Moving Average using EVI data, Logistic Curve

Table VI. Park Falls site Pearson r correlations ×100 (sig. @ 0.05 [α] level shown in bold, see Table IV for label descriptions).

	BU	DMAE	DMAN	LC	MT2E	MT2N	MT5E	MT5N	SME	SMN	SIFB
BU	X										
DMAE	**+97**	X									
DMAN	−02	−11	X								
LC	**+95**	**+76**	−05	X							
MT2E	+75	−07	−58	+23	X						
MT2N	+60	−22	−62	+11	**+98**	X					
MT5E	X	X	X	X	X	X	X				
MT5N	+62	−20	−61	+12	**+98**	**+99**	X	X			
SME	+95	**+95**	−27	+71	+09	−04	X	−02	X		
SMN	+36	−49	**+87**	+26	−28	−43	X	−41	+24	X	
SIFB	**+97**	**+81**	−06	**+79**	+28	+11	X	+13	**+82**	+39	X

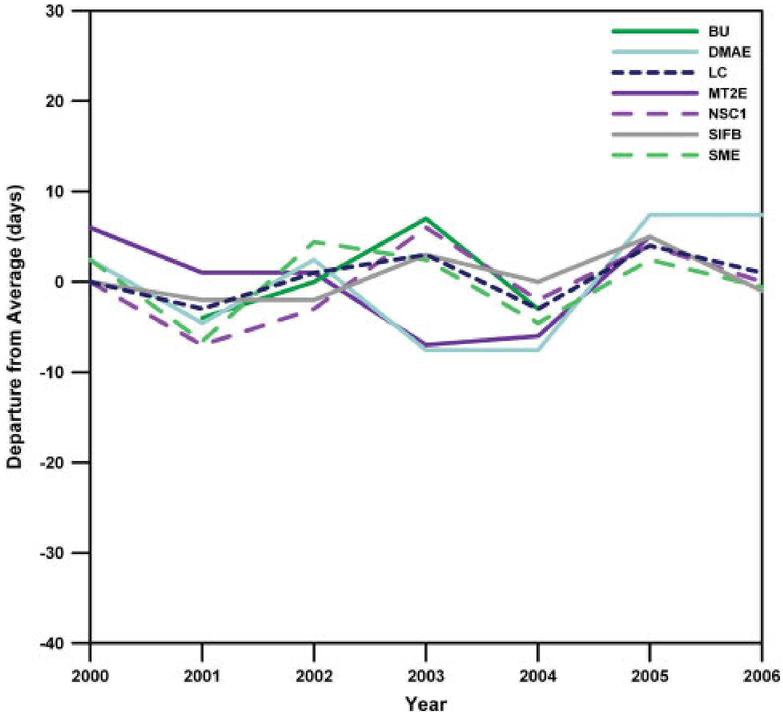

Figure 4. Harvard Forest selected variable comparisons, 2000–2006. BU (Boston University); DMAE (Delayed Moving Average using EVI data); LC (Logistic Curve); MT2E (Modified TIMESAT at 250 m resolution using EVI data); NSC1 (Native Species Budburst Cluster 1); SIFB (Spring Indices First Bloom date) and SME (Seasonal Midpoint using EVI data).

and Seasonal Midpoint using EVI data all show consistent correlations with ground phenology, but Logistic Curve (by virtue of its one correlation at Harvard Forest, mentioned above) is the only SOS technique to have at least one significant correlation to a ground phenology measure at all three sites. Conversely, SI First Bloom date is significantly correlated to Logistic Curve at UWM Field Station and Park Falls, and nearly so at Harvard Forest (as mentioned above).

The relationships among energy flux, native species clusters and SI First Bloom date at UWM Field Station are shown in Figure 7. Native Species Bud Burst Cluster 1, SI First Bloom date and the flux crossover point (where downward trending sensible heat and upward trending latent heat overlap) all occur at approximately the same time. This same relationship appears to extend to other energy flux sites, such that average SI First Bloom date can predict the average flux crossover date with modest error (Figure 8). An even stronger relationship seems to exist between average SI First Bloom date and NEE downturn date (Figure 9).

4. Discussion

The overall relationships among ground phenology and satellite-derived SOS at the Harvard Forest, UWM Field Station, and Park Falls sites present interpretive challenges. First, significant correlations among the various satellite-derived SOS dates at the three sites are infrequent, which suggests that these SOS techniques are often not detecting the same phenomena, in agreement with recent continental-scale results (White *et al.*,

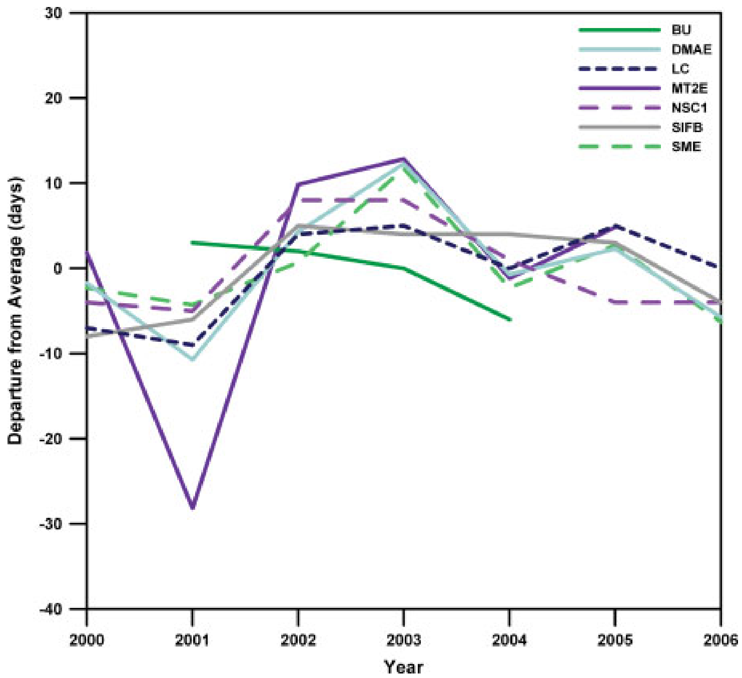

Figure 5. UWM Field Station selected variable comparisons, 2000–2006 (see Figure 4 for label descriptions).

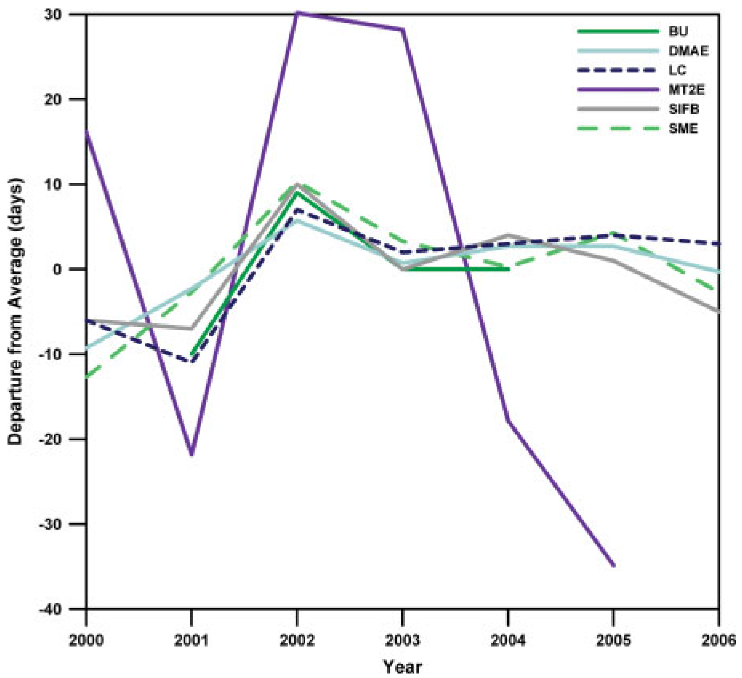

Figure 6. Park Falls selected variable comparisons, 2000–2006 (see Figure 4 for label descriptions).

(in press))[†] Indeed, contrasting approaches of the individual methodologies are at least one reason why this outcome might be expected (Table I).

Second, significant correlations among ground phenology measures are more common, suggesting that there may be multiple ways to represent the overall ground response accurately (at least within the limitations of point-based measures). However, when ground phenology and satellite-derived SOS are compared among sites, there are few significant correlations. Much of the difficulty in comparing these two categories of phenological measures comes from their contrasting properties. Currently, satellite-derived SOS dates are typically produced from continuous VI data acquired at a very coarse temporal (no more than once every 14–16 days) and spatial (most at 1 km) resolution. Ground phenology data are generally point values (discrete phenological events at specific locations).

[†] Correction made here after initial publication.

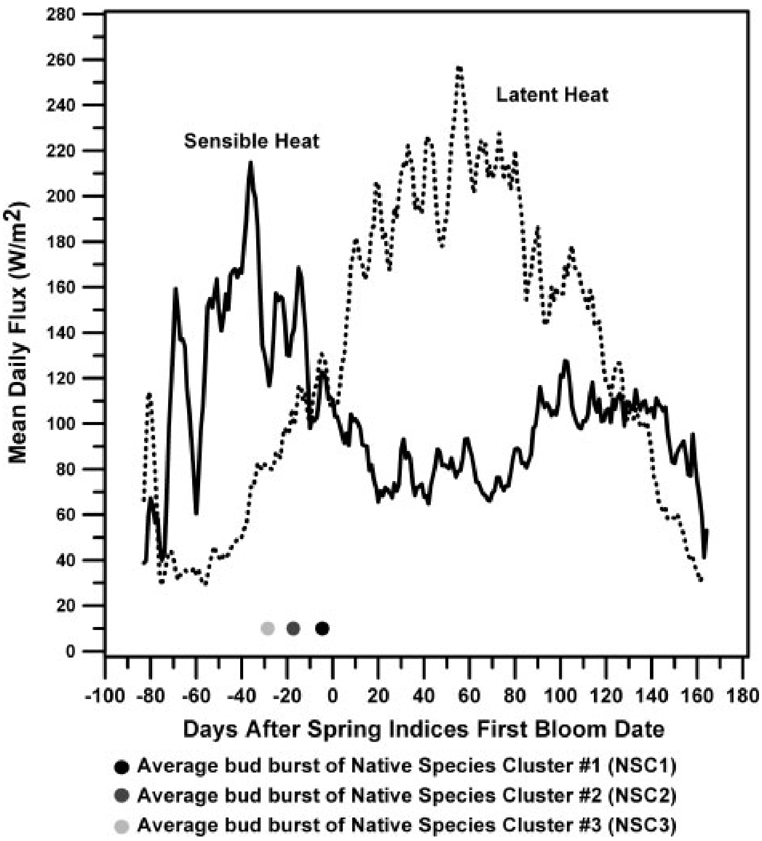

Figure 7. Energy flux comparisons and native species phenology at UWM Field Station, 2002–2006.

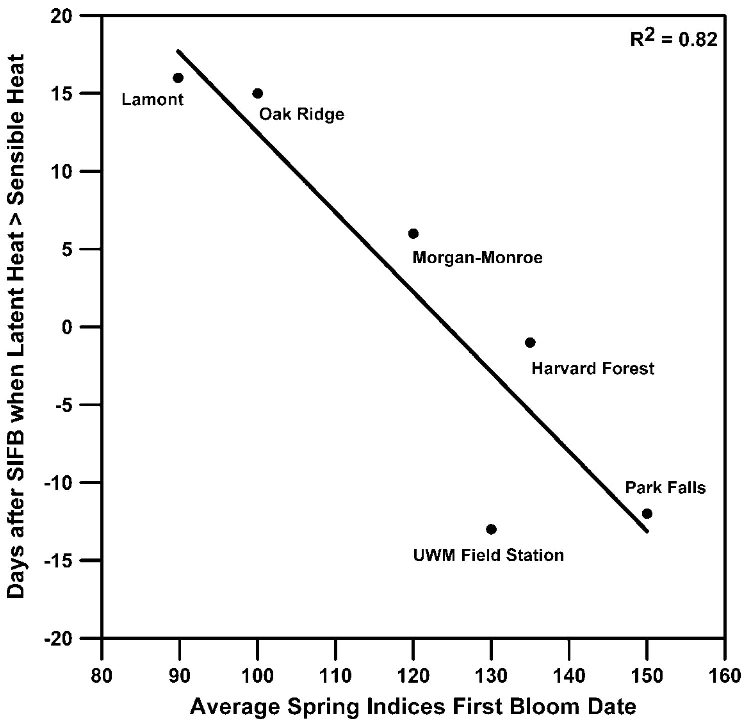

Figure 8. Relationship between Spring Indices First Bloom date (SIFB) and latent-sensible heat crossover date (sig. @ 0.05 [α] level).

So anything that can bridge these differing properties will likely improve their correlation, such as recording ground phenological status continuously (every 2–4 days) rather than just recording the dates of discrete events (Liang and Schwartz, 2009). Also, given that MODIS VI data are collected daily, efforts should be made to reduce the temporal coarseness of generally available VI data to the limits allowable by cloud cover

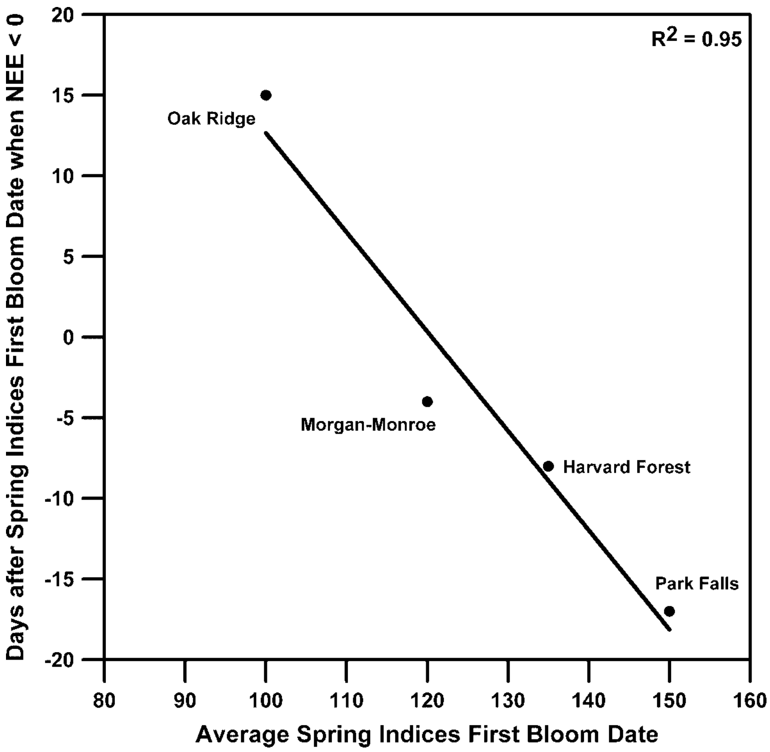

Figure 9. Relationship between Spring Indices First Bloom date and Net Ecosystem Exchange (NEE) downturn date (sig. @ 0.05 [α] level).

interference (down to at least 7–10 days) and improve typical spatial resolution to 500 m or less. Finally, the positive comparative results of this study for a logistic curve fitting SOS technique using the midpoint of the VI curve as a point of reference (Logistic Curve), should encourage further development and testing of this relatively new approach.

Another interesting result is that SI First Bloom date does as well as any of the direct native species measures, which is a bit curious, but perhaps the most unexpected result is that surface–satellite relationships are stronger overall at UWM Field Station, rather than Harvard Forest. This would initially appear to be counter intuitive, given that Harvard Forest is 100% forested and UWM Field Station is less than half so.

One possible explanation is that despite the equally large number of native species monitored at both sites, those at UWM Field Station are more representative precisely because they occupy a smaller area of the total landscape. Even though we know the general distribution of trees at both sites, and have phenology data from the appropriate species, perhaps the sample size at Harvard Forest is inadequate, given the heterogeneity of the Harvard Forest site, while at UWM Field Station (with almost 50% one species, *A. saccharum*), these same factors are not an issue. These results strongly suggest that ground phenology should preferably be monitored in a fashion that is not only representative of the diversity of species at a site, but also their abundance and distribution. Simply put, ground phenology must be monitored in an explicitly areal fashion if there is any hope to effectively mesh it with satellite-derived SOS measures, which are inherently areal (Liang and Schwartz, 2009).

Using SI First Bloom date as a proxy for ground phenological responses and late winter–early spring at sites in the eastern USA DBF biome shows an ordering of differences between cool northern sites and warmer southern sites. This relationship allows average SI First Bloom date to also serve as a proxy for average spring NEE drawdown date, and latent-sensible heat crossover date (Figures 8 and 9). Schwartz and Crawford (2001) showed similar results at a smaller number of sites, and (using a completely different methodology) Baldocchi *et al.* (2005) likewise demonstrated that there is a comparable significant relationship between inferred leaf-out date and the time in spring when NEE reaches zero. These relationships exist because northern species apparently grow more 'efficiently' (reach comparable phenological development stages with less energy) than their southern counterparts. This phenological effect has been demonstrated in numerous experiments where members of the same tree species have been brought from multiple sites to a single location (Schlarbaum and Bagley, 1981). Schwartz (1992) showed that this effect could be detected on a continental-scale using meteorological measurements, and Kathuroju *et al.* (2007) recently demonstrated that the same effect (with similar magnitude) could be identified using satellite-derived SOS measures and SI First Bloom dates.

In order to provide a context for understanding future changes, we need to develop a strategic approach for intercomparison of phenological events between sites.

For those sites where daily maximum–minimum air temperatures are available among applicable plant types (forest trees, shrubs and crops planted in temperate regions with adequate rainfall, or crops in irrigated temperate dry regions), the SI models are one effective method to simulate ground phenological responses for effective comparison to satellite-derived SOS [as validated by Schwartz *et al*. (2002) and most recently by White *et al*. (in press)]† and may facilitate studies aiming to reconstruct past patterns or determine comparability of various satellite-derived and ground phenology measures.

Regarding the critical need to effectively compare ground-based and satellite-derived SOS measures, it is clear that attempts using limited numbers of individual plants (even if large numbers of species and ancillary species abundance data are available) will face considerable challenges. Given that satellite-derived measures are areal and at a scale of 250 m and larger, collecting ground phenology data at the same areal scale should be considered to make effective comparisons. Although some of this type of information can be gathered by satellite-derived or ground-based measurements of leaf area index (LAI) or various other ground-based sensors, these approaches have less potential to reveal the underlying diverse response and variation of species phenology. By sampling a large number of trees in appropriately large areas, the variation of species phenology in relation to environmental and genetic factors can be discerned, and provide the means to build ground-based phenology 'pixels' that are truly a fair and valid comparison with satellite-derived phenology (Liang and Schwartz, 2009). A study exploring these issues is underway in northern Wisconsin in the footprint of the Park Falls/WLEF tall tower, funded by National Science Foundation grant number BCS-0649380.

Acknowledgements

This article is partially based upon work supported by the National Science Foundation under grant number ATM-0085224. Some data were also obtained from the Atmospheric Radiation Measurement (ARM) Program sponsored by the US Department of Energy, Office of Science, Office of Biological and Environmental Research, Environmental Sciences Division. We are grateful to J. Fisher, Boston University (via M. Friedl), NASA Goddard Space Flight Center (via J. Morisette), the U.S. Geological Survey (via B. Reed), B. Tan, and M. White who either provided or helped us produce the satellite-derived SOS values used in this study. Our thanks to the following Ameriflux principal investigators who allowed us to use carbon and energy flux data from their sites: D. Baldocchi/T. Meyers (Oak Ridge, TN), K. Davis (Park Falls/WLEF, WI), D. Dragoni (Morgan-Monroe, IN), and J. Munger/S. Wofsy (Harvard Forest, MA). We also appreciate the use of daily weather data from the Harvard Forest site, and Harvard Forest phenology data, which have been collected by J. O'Keefe. UW–Milwaukee Field Station flux and phenology data were collected in cooperation with G. Meyer and J. Reinartz.

† Correction made here after initial publication.

References

Badeck FW, Bondeau A, Böttcher K, Doktor D, Lucht W, Schaber J, Sitch S. 2004. Responses of spring phenology to climate change. *New Phytologist* **162**: 295–309.

Baldocchi DD, Black TA, Curtis PS, Falge E, Fuentes JD, Granier A, Gu L, Knohl A, Pilegaard K, Schmid HP, Valentini R, Wilson K, Wofsy S, Xu L, Yamamoto S. 2005. Predicting the onset of net carbon uptake by deciduous forests with soil temperature and climate data: a synthesis of FLUXNET data. *International Journal of Biometeorology* **49**: 377–387.

Cleland EE, Chuine I, Menzel A, Mooney HA, Schwartz MD. 2007. Shifting plant phenology in response to global change. *Trends in Ecology and Evolution* **22**: 357–365.

Davis KJ, Bakwin PS, Yi C, Berger BW, Conglong Z, Teclaw RM, Isebrands JG. 2003. The annual cycles of CO_2 and H_2O exchange over a northern mixed forest as observed from a very tall tower. *Global Change Biology* **9**: 1278–1293.

Dunnum J. 1972. The upland hardwood forest of the cedar-sauk field station. *Field Station Bulletin* **5**: 17–23.

Fisher JI, Mustard JF. 2007. Cross-scalar satellite phenology from ground, Landsat, and MODIS data. *Remote Sensing of Environment* **109**: 261–273.

Gao F, Morisette JT, Wolfe RE, Ederer G, Pedelty J, Masuoka E, Myneni R, Tan B, Nightingale J. 2008. An Algorithm to Produce Temporally and Spatially Continuous MODIS-LAI Time Series. *IEEE Geoscience and Remote Sensing Letters* **5**: 60–64.

Hayhoe K, Wake C, Huntington TG, Luo L, Schwartz MD, Sheffield J, Wood E, Anderson B, Bradbury J, DeGaetano A, Troy TJ, Wolfe D. 2007. Past and Future Changes in Climate and Hydrological Indicators in the U.S. Northeast. *Climate Dynamics* **28**: 381–407.

Jonsson P, Eklundh L. 2004. TIMESAT-a program for analyzing time-series of satellite sensor data. *Computers and Geosciences* **30**: 833–845.

Kathuroju N, White MA, Symanzik J, Schwartz MD, Powell JA, Nemani RR. 2007. On the use of the Advanced Very High Resolution Radiometer for development of prognostic land surface phenology models. *Ecological Modelling* **201**: 144–156.

Liang L, Schwartz MD. 2009. Landscape phenology: an integrative approach to seasonal vegetation dynamics. *Landscape Ecology* **24**: 465–472.

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

Oak Ridge National Laboratory Distributed Active Archive Center. 2007. *MODIS subsetted land products, collection 4*, http://www.daac.ornl.gov/MODIS/modis.html, from ORNL DAAC, Oak Ridge, Tennessee (accessed October 8, 2007).

Reed BC, Brown JF, VanderZee D, Loveland TR, Merchant JW, Ohlen DO. 1994. Variability of land cover phenology in the United States. *Journal of Vegetation Science* **5**: 703–714.

Reed BC, White MA, Brown JF. 2003. Remote sensing phenology. In *Phenology: An Integrative Environmental Science*, Schwartz MD (ed). Kluwer: Netherlands; 365–381.

Ricciuto DM, Butler MP, Davis KJ, Cook BD, Bakwin PS, Andrews A, Teclaw RM. 2008. Causes of interannual variability in ecosystem–atmosphere CO_2 exchange in a northern Wisconsin forest using a Bayesian model calibration. *Agricultural and Forest Meteorology* **148**: 309–327.

Schlarbaum SE, Bagley WT. 1981. Intraspecific genetic variation of *Quercus rubra L.*, Northern Red Oak. *Silvae Genetica* **30**: 50–56.

Schmid HP, Grimmond CSB, Cropley F, Offerle B, Su HB. 2000. Measurements of CO_2 and energy fluxes over a mixed hardwood forest in the mid-western United States. *Agricultural and Forest Meteorology* **103**: 357–374.

Schwartz MD. 1992. Phenology and springtime surface layer change. *Monthly Weather Review* **120**: 2570–2578.

Schwartz MD. 1997. Spring index models: an approach to connecting satellite and surface phenology. In *Phenology of Seasonal Climates*, Lieth H, Schwartz MD (eds). Backhuys: Netherlands; 23–38.

Schwartz MD, Ahas R, Aasa A. 2006. Onset of spring starting earlier across the Northern Hemisphere. *Global Change Biology* **12**: 343–351.

Schwartz MD, Crawford TM. 2001. Detecting energy-balance modifications at the onset of spring. *Physical Geography* **21**: 394–409.

Schwartz MD, Reed BR, White MA. 2002. Assessing satellite-derived start-of-season measures in the conterminous USA. *International Journal of Climatology* **22**: 1793–1805.

Schwartz MD, Reiter BE. 2000. Changes in North American spring. *International Journal of Climatology* **20**: 929–932.

Tan B, Morisette JT, Wolfe RE, Gao F, Ederer GA, Nightingale J, Pedelty JA. 2008. Vegetation phenology metrics derived from temporally smoothed and gap-filled MODIS data. Proceedings of the 2008 *International Geoscience and Remote Sensing Symposium* (Boston, MA), 593–596.

White MA, de Beurs KM, Didan K, Inouye DW, Richardson AD, Jensen OP, Magnuson J, O'Keefe J, Zhang G, Nemani RR, van Leeuwen WJD, Brown JF, de Wit A, Schaepman M, Lin X, Dettinger M, Bailey A, Kimball J, Schwartz MD, Baldocchi DD, Lee JT, Lauenroth WK. Intercomparison, interpretation, and assessment of spring phenology in North America estimated from remote sensing for 1982 to 2006. *Global Change Biology* (in press), DOI: 10.1111/j.1365-2486.2009.01910.x.

White MA, Nemani RR, Thornton PE, Running SW. 2002. Satellite evidence of phenological differences between urbanized and rural areas of the eastern United States deciduous broadleaf forest. *Ecosystems* **5**: 260–273.

White MA, Thornton PE, Running SW. 1997. A continental phenology model for monitoring vegetation responses to interannual climatic variability. *Global Biogeochemical Cycles* **11**: 217–234.

Wilson KB, Baldocchi DD. 2001. Comparing independent estimates of carbon dioxide exchange over five years at a deciduous forest in the southern United States. *Journal of Geophysical Research* **106**: 34167–34178.

Wofsy SC, Goulden ML, Munger JW, Fan S-M, Bakwin PS, Daube BC, Bassow SL, Bazzaz FA. 1993. Net exchange of CO_2 in a mid-latitude forest. *Science* **260**: 1314–1317.

Zhang X, Friedl MA, Schaaf CB, Strahler AH, Hodges JCF, Gao F, Reed BC, Heute A. 2003. Monitoring vegetation phenology using MODIS. *Remote Sensing of Environment* **84**: 471–475.

