Remote Sensing of Environment 115 (2011) 143–157

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

# Validating satellite phenology through intensive ground observation and landscape scaling in a mixed seasonal forest

Liang Liang [a,*], Mark D. Schwartz [b], Songlin Fei [a]

[a] Department of Forestry, University of Kentucky, 216 T. P. Cooper Building, Lexington, KY 40546-0073, United States
[b] Department of Geography, University of Wisconsin-Milwaukee, Bolton 410, P.O. Box 413, Milwaukee, WI 53201-0413, United States

ARTICLE INFO

*Article history:*
Received 15 May 2010
Received in revised form 15 August 2010
Accepted 19 August 2010

*Keywords:*
Phenology
Land surface phenology
Landscape phenology
MODIS
Validation
EOS Land Validation Core Sites

ABSTRACT

Phenology is a key component of monitoring terrestrial ecosystem variations in response to global climate change. Satellite-measured land surface phenology (LSP) has been widely used to assess large scale phenological patterns and processes. However, the accuracy of LSP is rarely validated with spatially compatible field data due to the significant spatiotemporal scale mismatch. In this study, we employ intensive field observations specifically designed to address this deficiency. High density/frequency spring phenological observations were collected in a mixed seasonal forest during 2008 and 2009. A landscape up-scaling approach was used to derive landscape phenology (LP) indices from plot-level observations in order to validate Moderate-resolution Imaging Spectroradiometer (MODIS) based LSP. Results show that the MODIS Enhanced Vegetation Index (EVI) derived start of spring season (SOS) measure was able to predict LP full bud burst date with absolute errors less than two days. In addition, LSP derived SOS captured inter-annual variations and spatial differences that agreed with ground observations. Comparison of complete time series of LP and LSP revealed that fundamental differences exist between the two observation means, e.g., LP development had increased influence on LSP during the course of spring onset. Therefore, inferring continuous LP processes directly from LSP measures could be problematic. However, using LSP derived from techniques such as logistic curve modeling for extracting seasonal markers appears more promising. This study contributes to a more explicit understanding of the linkages between remotely sensed phenology and traditionally observed (ground-based) phenology.



## 1. Introduction

Phenology tracks climate-driven plant/animal life cycle events, such as bud burst, flower bloom and bird migration (Schwartz, 2003). Among the many environmental impacts of climate change, phenology and its broader ecological impacts deserve special attention (IPCC, 2007; Morisette et al., 2009). Notably, a warming climate has led to phenological timing shifts over the past half century for many temperate land regions, as supported by both observations and retrospective models (Menzel et al., 2006; Schwartz et al., 2006; Van Buskirk et al., 2009). Phenological timing is sensitive to small fluctuations of climate and is easy to observe. Plant phenology, in particular, has the advantage of being relatively static in space compared to phenologies of birds or other animals, therefore providing better opportunities for recurrent observations of the same individuals at fixed locations across years. Hence, it is useful as a basis for establishing spatially explicit phenology models relative to climate change (Cleland et al., 2007; Khanduri et al., 2008).

Traditionally, plant phenology has been monitored by human observers (Lieth, 1974). Over the last three decades, aggregated effects of leaf phenology across broad landscapes – seasonal greening of vegetated land cover – has become detectable through satellite sensors and associated vegetation indices (Fischer, 1994; Goward et al., 1991; Huete et al., 2002; Reed et al., 2009). Sensors such as the Advanced Very High Resolution Radiometer (AVHRR) and Moderate-resolution Imaging Spectroradiometer (MODIS) onboard polar orbiting satellites have allowed calculation of vegetation indices (VI) time series from near-daily frequency measurements. Data from these two sensors (AVHRR and MODIS) are currently the major data sources in addressing regional to global vegetation phenology monitoring needs. As dates of phenological events (such as bud break or leaf senescence) are traditionally the primary measures of phenological timing, information derivation from satellite data likewise has been focused on estimating key dates of phenological change reflected on annual VI curves. In past studies, multiple techniques have been developed to extract phenological markers from VI time series (de Beurs & Henebry, 2010; Lloyd, 1990; Reed et al., 1994; White et al., 1997; Zhang et al., 2001). In an annual cycle, primary phenological markers include critical times such as the growing season start and end dates. The satellite-derived growing season start date is particularly

* Corresponding author. Tel.: +1 859 257 5666; fax: +1 859 323 1031.
*E-mail address:* liang.liang@uky.edu (L. Liang).

0034-4257/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.rse.2010.08.013

important in tracking inter-annual phenological variations and is typically referred to as start of spring season (SOS, a list of acronyms used in this paper is provided in Table 1) or start of season (Robin et al., 2007; Schwartz & Reed, 1999; White et al., 2009). Because observation of vegetation phenology through satellite sensors integrates signals from land surface processes in addition to reflectance from plants, satellite-based phenology is often called land surface phenology (de Beurs & Henebry, 2004; Friedl et al., 2006). With remote sensing's ability to detect land surface phenology (LSP) at regional, continental, and global scales, modeling spatially explicit phenological patterns related to climate variability has become possible (de Beurs & Henebry, 2008; Jolly et al., 2005; Kathuroju et al., 2007; Zhang et al., 2007, 2009).

However, to date, measures of LSP usually compare poorly with field-observed phenology (Schwartz & Hanes, 2009; White et al., 2009). A primary hurdle is the incompatibility of spatial scales at which two types of observations are commonly obtained, in addition to different characteristics of the data. In particular, conventional phenology observes growth phases diagnosed for a limited number of individual plants, which are often found at isolated locations without areal coverage; while LSP records electromagnetic reflectance over contiguous land patches. Therefore, reconciliation of ground and satellite-based phenological observations has been extremely challenging. White et al. (2009) performed a comprehensive intercomparison of ten AVHRR-based LSP SOS methods over broad regions of North America, relating them to ground phenology data, bioclimatic model predictions, and cryospheric/hydrologic seasonality metrics. The tested LSP SOS measures deviate from each other by up to two months. And when compared to ground data, point-vs.-pixel comparison errors were considered a major source of uncertainty. Schwartz and Hanes (2009) also found few significant correlations between ten MODIS-based LSP SOS measures and phenologies observed at three sites in the eastern U.S. These studies strongly suggest that field data collected with systematic areal sampling to address satellite LSP validation are essential to enable fair comparisons of the two phenological data types.

Lacking compatible field data, multiple efforts have been made to compare satellite phenology to ground-based phenology models (Schwartz & Reed, 1999; Schwartz et al., 2002). Bioclimatic models such as the Spring Indices can support ground–satellite integration at the continental scale with geographically consistent variations (Schwartz, 1997). As another alternative to intensive field data, medium resolution (e.g., 30 m) imagery was used as an intermediary to bridge coarse resolution (e.g., 500 m) LSP and plot-level phenology (Fisher & Mustard, 2007; Fisher et al., 2006). Fisher et al. (2006) made comparisons between LSP (derived from 30 m multi-year integrated Landsat data) and plot-level phenology in New England deciduous forests. Results suggested consistent spatial variability between Landsat LSP and ground data within masked deciduous vegetation covers. The spatial variability was investigated in relation to local climate gradients as well as regional landscape controls. Fisher and Mustard (2007) further investigated inter-annual variability of Landsat LSP and field data by incorporating coarse resolution (500 m) MODIS-based LSP, and found temporal variability generally agreed across scales in two experimental forests in New England. However, large uncertainties were suggested due to the inability to account for within-pixel phenological variability with coarse-scale LSP. Moreover, similar research conducted in Europe suggested an overall agreement between ground and satellite measurements in more homogeneous and internally consistent vegetation covers (such as pure deciduous stands), while also inferring landscape heterogeneity to be a significant source of errors and uncertainties (Doktor et al., 2009; Soudani et al., 2008). Although ground-based/satellite phenology comparisons have been previously attempted, a major hurdle is always the lack of ground-based in situ observations sufficient to account for spatial heterogeneity at larger (landscape-level) scales. Thus, past studies suggest that in situ data collection designed specifically for assessing LSP with appropriate scaling methods holds the best promise to further clarify the relationship between LSP and field-observed phenology.

Efforts of the USA-National Phenology Network (USA-NPN) and collaborators are expanding the spatial extent of ground phenology data collected across the country (USA-NPN: http://www.usa-npn.org). Continued data accumulation will provide unprecedented opportunity for better understanding of phenological phenomenon over broad regions and facilitate the integration with remotely sensed data, as well as continental-scale phenology models. However, most ground data being collected are still essentially isolated point features, lacking spatial coverage that can match satellite pixels. Also, most ground data available are for discrete phenological events, lacking temporal continuity compared to the time series of remote sensing imagery. This is likely the major reason that to date most comparisons have been made between phenological markers (e.g., SOS) rather than to complete time series. Recently, near surface remote sensing (using automated digital camera systems) has emerged as a useful tool for capturing repeat canopy phenology (Ahrends et al., 2008; Richardson et al., 2007). Such techniques hold promise for more objective and direct comparison with satellite measures. The tradeoff is that it mainly records vegetation reflectance change and loses the biological inferences carried by conventional phenological observation, being more akin to satellite remote sensing. Again, in order to improve the synergy among different phenological observation means, studies with intensively observed plot-level phenology and associated scaling methods are crucial.

High spatial/temporal density data can be utilized to generate integrated phenology representations by incorporating landscape patterns and processes (Liang & Schwartz, 2009). Landscape phenology (LP) derived from in situ observations can serve as a direct counterpart of satellite-based LSP, making compatible comparisons possible. Our study is a full-scale implementation and major improvement of LP methodology specifically geared towards validating LSP measures. To achieve this goal, two procedures are critical. First, we systematically sample dominant tree species in target forest patches, as well as collect springtime canopy phenology in a conventional but spatially and temporally concentrated manner. Second, a hierarchical scaling approach is employed to bring plot observation to the satellite pixel level, thus enabling comprehensive comparisons of field-based and satellite-derived measures of spring phenology.

**Table 1**
Acronyms and corresponding terms.

| Acronyms | Terms |
|---|---|
| AVHRR | Advanced Very High Resolution Radiometer |
| AMSL | Above mean sea level |
| BBCH | Biologische Bundesanstalt, Bundessortenamt and Chemical industry |
| DOY | Day of year |
| EOS | Earth Observing System |
| EVI | Enhanced Vegetation Index |
| FBB | Full bud burst |
| LP | Landscape phenology |
| LSP | Land surface phenology |
| MODIS | Moderate-resolution Imaging Spectroradiometer |
| NASA | National Aeronautics and Space Administration |
| NDVI | Normalized Difference Vegetation Index |
| ORNL DAAC | Oak Ridge National Laboratory Distributed Active Archive Center |
| RE | Relative errors |
| RMSE | Root mean square errors |
| SOS | Start of spring season/start of season |
| USA-NPN | USA-National Phenology Network |
| VI | Vegetation indices |

## 2. Study area and data

The study area is located in the Park Falls Range District of the Chequamegon National Forest in Wisconsin, near an AmeriFlux eddy

covariance tower site (Park Falls/WLEF, 45.946°N, 90.272°W), which is also a National Aeronautics and Space Administration (NASA) Earth Observing System (EOS) Land Validation Core Site (Morisette et al., 2002). Pleistocene glaciers eroded away most pre-existing topography of the area and covered the bedrock with up to 30 m thick glacial till and about 10 m thick outwash deposits (Martin, 1965). Hence, the entire landscape is composed of forested ground moraines and outwash plains, spotted with lakes and swamps. There is approximately 30 m of local relief ranging from 470 to 500 m above mean sea level (AMSL) near the site. Soils in the study area are mainly loamy under forests and wetland soils which all developed from glacial–fluvial parent materials. The region features a humid continental climate with long winters and relatively mild summers. Annual temperatures fluctuate between −18 °C and 25.8 °C, with an average of 4.8 °C (Wisconsin Online: http://www.wisconline.com/counties/price/climate.html). Total annual precipitation is around 810 mm including 126.2 cm of snowfall. Median growing season length (0 °C criteria) is 131 days. Full leaf emergence usually occurs in mid-late May and leaf fall in late September/early October (AmeriFlux General Site Information: http://public.ornl.gov/ameriflux/Site_Info/siteInfo.cfm?KEYID=us.park_falls.01).

Vegetation near the WLEF tower is characterized as a secondary mixed temperate forest with both deciduous (70%) and coniferous (30%) tree species. Four major subtypes of forest were differentiated by earlier studies in connection to soil nutrient availability, hydrologic condition, and landforms (Burrows et al., 2002; Ewers et al., 2002; Kassnacht & Gower, 1997). These forest types are: 1) upland conifers: dominated by red pine (*Pinus resinosa*) and jack pine (*Pinus banksiana*) in areas of excessively drained, sandy soils derived from glacial outwash; 2) northern hardwoods: dominated by sugar maple (*Acer saccharum*), and other deciduous species such as red maple (*Acer rubrum*), white ash (*Fraxinus americana*), yellow birch (*Betula alleghaniensis*), and basswood (*Tilia americana*) that occur on the finer-textured soils derived from moraines and drumlins; 3) aspen/fir forests: dominated by quaking aspen (*Populus tremuloides*) and balsam fir (*Abies balsamea*), with soils of intermediate characteristics supporting a wide variety of other tree species, such as paper birch (*Betula papyrifera*), bigtooth aspen (*Populus grandidentata*), red maple, and red and white pine (*P. resinosa* and *P. strobus*); and 4) forested wetlands: dominated by white cedar (*Thuja occidentalis*), balsam fir (*A. balsamea*), and speckled alder (*Alnus regosa*) that occurs on the poorly drained organic soils, with other mesic species like white spruce (*Picea glauca*), black spruce (*Picea mariana*), and tamarack (*Larix laricina*). These four cover types comprise more than 80% of the area around the WLEF tower.

Systematic sampling and detailed phenological observation of sampled trees were carried out in a temporally and spatially concentrated manner. In summer 2007, transects were completed for two 625×625 m study areas (Fig. 1). The north study area featured upland conifers and northern hardwoods; and the south study area was more heterogeneous, dominated by aspen/fir forests, forested wetlands, as well as patches of pine plantation and hardwood stands. In particular, our tree sampling scheme consistently followed a two dimensional 3/7 cyclic sampling design, which often allows efficient detection of spatial autocorrelations (Burrows et al., 2002; Clayton & Hudelson, 1995). Along each transect, three plots were sampled from seven potential plot candidates which were evenly spaced with a 25 m sampling interval (see Fig. 1). Selected plots were separated by 25 m, 50 m, and 100 m intervals respectively in a repeated manner. This 3/7 cycle repeated four times on both the latitudinal and longitudinal dimensions within each study area, yielding 12 plots on each transect and 288 plots in total for both study areas. The three largest trees (with preference for dominant species) within a 10 m radius from each plot center were sampled/tagged, accumulating to 864 trees (24 basswood were also added in the north study area for comparison to concurrent phenological observations of this species on the UW-Milwaukee campus in future studies). This study used observations collected in spring 2008 and spring 2009 from both study areas. Observations were conducted by observers during approximately a five-week period (mid April to late May) each spring at a bi-daily frequency for all tagged trees. The bi-daily observation frequency was the best we were able to accomplish in the field due to limited resources.

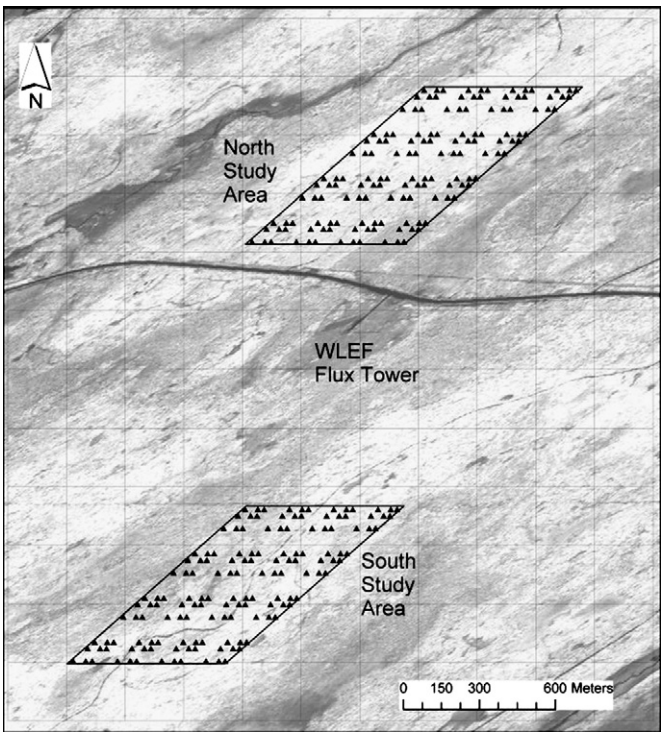


**Fig. 1.** Study area map (in sinusoidal projection) showing: 1) phenological observation transects/plots (according to a cyclic sampling design, see Burrows et al., 2002); 2) MODIS pixel grids (250×250 m squares); and 3) 2.4 m resolution NDVI image derived from a May 18, 2007 QuickBird image underlying the plots and MODIS pixel grids.

A detailed field protocol was used to quantify canopy leaf/needle development stages following a general structure prescribed in the German Biologische Bundesanstalt, Bundessortenamt and Chemical industry (BBCH) scheme (Meier, 2001), and adapted from a native species protocol used for phenological surveys at the Indiana Morgan–Monroe forest AmeriFlux site (J.C. Randolph, personal communication). A zero starting point was assigned to “no buds visible” stage, followed by a coding system for five major phases: buds visible (100), buds swollen (200), buds open (leaves or candles visible, 300), leaf clusters/candles out (not fully unfolded, 400), and leaves/needles fully unfolded (500). A leaf expansion level (600) was added for broad-leafed species. Then, a percentage estimate for the canopy was used to further delineate sub-phases within each phenological phase. To facilitate efficient field identification, we empirically divided the percentage range into four consecutive levels: 0–10% (0), 10–50% (10), 50–90% (50), and 90–100% (90). Following this protocol, for example, a tree canopy with estimated 50–90% buds open would be assigned a value of 350.

Satellite data used for deriving LSP were from MODIS onboard the EOS Terra satellite. Collection 5 VI product (MOD13Q1, available at 250 m spatial resolution and a 16-day composite window) was acquired from the Oak Ridge National Laboratory Distributed Active Archive Center (ORNL DAAC: http://www.daac.ornl.gov/MODIS/modis.html). Both Normalized Difference Vegetation Index (NDVI) and Enhanced Vegetation Index (EVI) were included. The time span of VI products downloaded was from January to August in 2008 and 2009. In addition to MODIS data, high resolution imagery was also utilized in our study. A summer (July 5, 2002) IKONOS image set with

1 m (panchromatic) and 4 m (multispectral) spatial resolutions was available from the EOS Land Validation Core Sites webpage (Park Falls/WLEF/CHEAS: http://landval.gsfc.nasa.gov/coresite.php?SiteID=30). Three 2.4 m resolution QuickBird multispectral images captured on April 17, 2006, May 3, 2007, and May 18, 2007 for the study area were provided directly by NASA. The acquisition dates of these images covered both deciduous canopy leaf-on (May 18, 2007) and leaf-off (April 17, 2006; May 3, 2007) conditions. These images were used to facilitate landscape characterization and scaling.

## 3. Methodology

### 3.1. Scaling rationale

In order to compare satellite pixel level VI signals with in situ observed tree canopy phenology, a scaling approach was required to bring the two types of observation to a comparable scale and extent. Tree canopy phenology was spatially represented by point features within a MODIS pixel-sized area. Given that our field phenology data were collected in a concentrated manner over study areas that were larger than the spatial resolution (250 m) of the MODIS product used, we employed an up-scaling approach towards generating a ground-based areal representation of landscape phenology (LP) to enable a comprehensive comparison with MODIS-based land surface phenology (LSP). We endeavored to achieve the highest precision possible by accounting for landscape heterogeneity using high resolution IKONOS and QuickBird imagery (as detailed later). Most importantly, akin to the Hierarchical Patch Dynamics paradigm (Wu, 1999; Wu et al., 2006), we employed a nested hierarchical scaling ladder following the predefined ecosystem structure. A step-wise scaling was performed to transfer individual tree phenology to population phenology, then to community phenology, and finally to landscape phenology. These integrated phenologies can be understood as corporate phenological patterns associated with different ecosystem organizational levels. At each step of scale transition, information specific to the next level were added. For instance, population phenology required accounting for intraspecific variations (variations within a species); and community phenology required incorporating spatial heterogeneity and interspecific variations (variations among species), in addition to the population phenology estimates. Such a scaling ladder allowed for ecologically coherent aggregation of phenological information through respective levels. A flowchart showing the rationale of the scaling approach, as well as primary procedures (discussed in following sections) is provided in Fig. 2.

### 3.2. Landscape phenology up-scaling

#### 3.2.1. Population phenology estimation

Assuming spatial autocorrelations exist among individual phenological observations, spatial interpolation based on predictable geographic variations would provide the most accurate population phenology estimates. However, according to Moran's I statistics and Semivariogram diagnosis, no consistent spatial autocorrelations were found in our phenological data. Instead, phenologies of individual plants of the same species expressed rather discrete characteristics. Thus, population phenology estimation fell back to a conventional statistical approach using the simple arithmetic average. An empirical minimum sample size (20) was recommended by past studies to represent population phenology when using averaging (Liang & Schwartz, 2009; West & Wein, 1971). Similarly, a smaller minimum sample size (15) was suggested for acquiring population phenology in tropical forests (Morellato et al., 2010). Therefore, in each of the study areas samples greater than 20 were averaged to generate population phenology for each species. In particular, sugar maple (sample size = 173), red maple (83), basswood (24), and balsam fir (21) populations were representative in the north study area. Quaking aspen (119), balsam fir (99), red maple (82), speckled alder (34), and white birch (23) were dominant in the south study area. Red pine had sufficient sample sizes in both study areas, but did not show observable spring phenology, hence was omitted in our study. As a remedy, the role of potential red pine growth was estimated with adjusted balsam fir phenology which had similar minimal influence on satellite data (as detailed later). In addition, to account for a pure aspen stand in the north study area, quaking aspen with a sample size of 17 was accepted to generate population phenology of this species.

#### 3.2.2. Community phenology derivation

Scaling population phenology to community phenology first involved accurately delineating community distribution and their boundaries. Our field survey provided information concerning species present at sampling plots, yet because the sampling was not exhaustive, exact community boundary locations could not be inferred directly from the field data. Hence high resolution satellite imagery was utilized to facilitate community boundary delineation. A 4 m resolution 2002 summer IKONOS image was first resolution-merged with a 1 m panchromatic image using a principle component approach to generate a 1 m resolution multispectral image (ERDAS Inc., 2008). Then we used a segmentation approach to divide the study area into contiguous spectrally similar image segments. Finally, these image segments were manually combined and attributed to different communities according to collected tree information and field notes. Given the classification was heavily supervised for the entire study areas, we did not do an accuracy assessment. The resultant maps of forest communities for the north and south study areas are respectively shown in Fig. 3. In the north study area, hardwood forests with sugar maple and red maple as the primary species occupied almost half of the study area. Pine plantations were also dominant in the north study area. On the east edge of north study area, grass/shrub land and tamarack-dominated wetland occurred with a highly mixed transitional zone between them. A small pure quaking aspen stand was found on the northwest corner of north study area. The south study area was more fragmented, with large areas of aspen/fir wood on the upland and forested wetland on the west. Transitional areas were composed of communities dominated by speckled alder bushes, grass/shrub patches, highly mixed areas and sparse aspen/fir wood. A sphagnum bog was located on the east edge with old growth residues distributed along its peripherals. In the southeast quarter of the south study area, pine plantations and maple-dominated hard woods were present, similar to those found in the north study area.

A further step to upscale population phenology to community phenology required an estimation of species abundance in each community. Dominant species members of each community were known from the field survey, but only rough estimates can be made from the field data in regards to species abundance in each community. Hence, in order to increase the scaling precision, multi-temporal QuickBird images were used to facilitate further characterization of community compositions. Specifically, in the mixed forest under study, deciduous and coniferous covers can be differentiated using images acquired on contrasting seasonal occasions when deciduous species were leaf-off and leaf-on. Utilizing this phenological asynchrony, we were able to estimate the proportions of deciduous, coniferous, and non-vegetated background within communities.

Three QuickBird images captured from two deciduous leaf-off dates (May 3, 2007 and April 17, 2006) and one deciduous leaf-on date (May 18, 2007) were available for a multi-temporal land cover analysis. The QuickBird Standard Imagery product has a positioning accuracy of 23 m (circular error with the 90% confidence level) excluding off-nadir viewing angle and terrain distortions (Cheng et al., 2003). Locational mismatch across images was geometrically corrected by coregistering the QuickBird images to the 2002 IKONOS image. Using 25 ground control points each, all QuickBird images were transformed with 2nd order polynomial models and the nearest

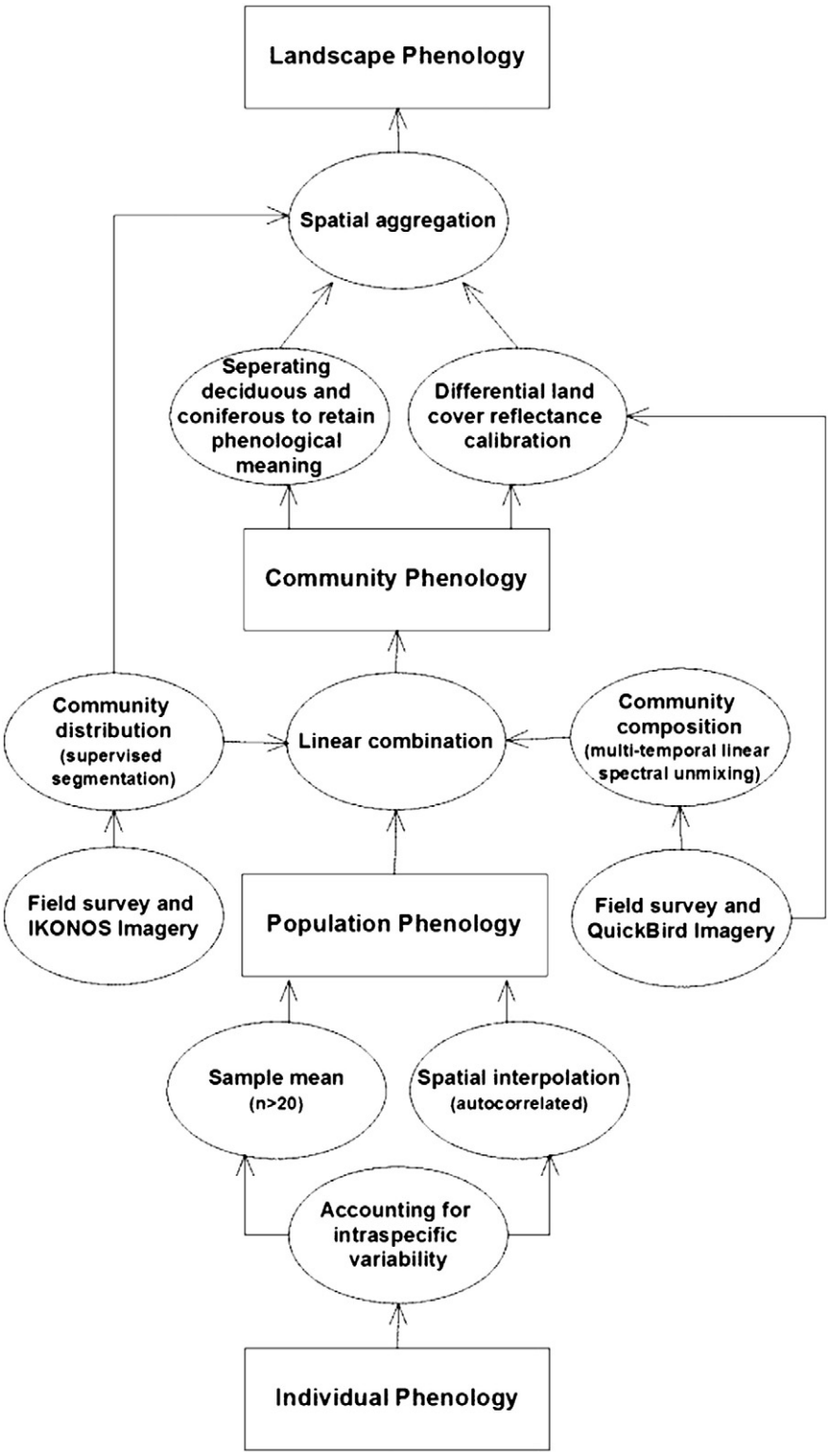


**Fig. 2.** Conceptual diagram and flow chart of hierarchical phenological organization and up-scaling towards landscape phenology (LP); LP as of an ecosystem patch sits on the same level as ecosystems; the procedure of up-scaling is composed of three primary stages connecting the four levels of phenological representation.

neighbor resampling approach. Root mean square errors (RMSE) achieved were less than 1 m for all three images.

Using the multi-temporal QuickBird images, we adopted a linear spectral unmixing approach to estimate respective portions of coniferous, deciduous, and non-vegetated covers at the sub-pixel level. Multi-temporal linear spectral unmixing models classify vegetation cover types with improved accuracy over traditional methods (DeFries et al., 2000; Lu et al., 2003; Roberts et al., 1998). This approach can be particularly useful to separate coniferous, deciduous, and non-vegetated covers based on their different phenological characteristics. Due to cloud contamination (10.1%) which affected a small portion of the study area in the May 3, 2007 image, this image was not used for the sub-pixel analysis. Consequently, the April 17, 2006 image and May 18, 2007 image were stacked to form a multi-temporal image series, upon which

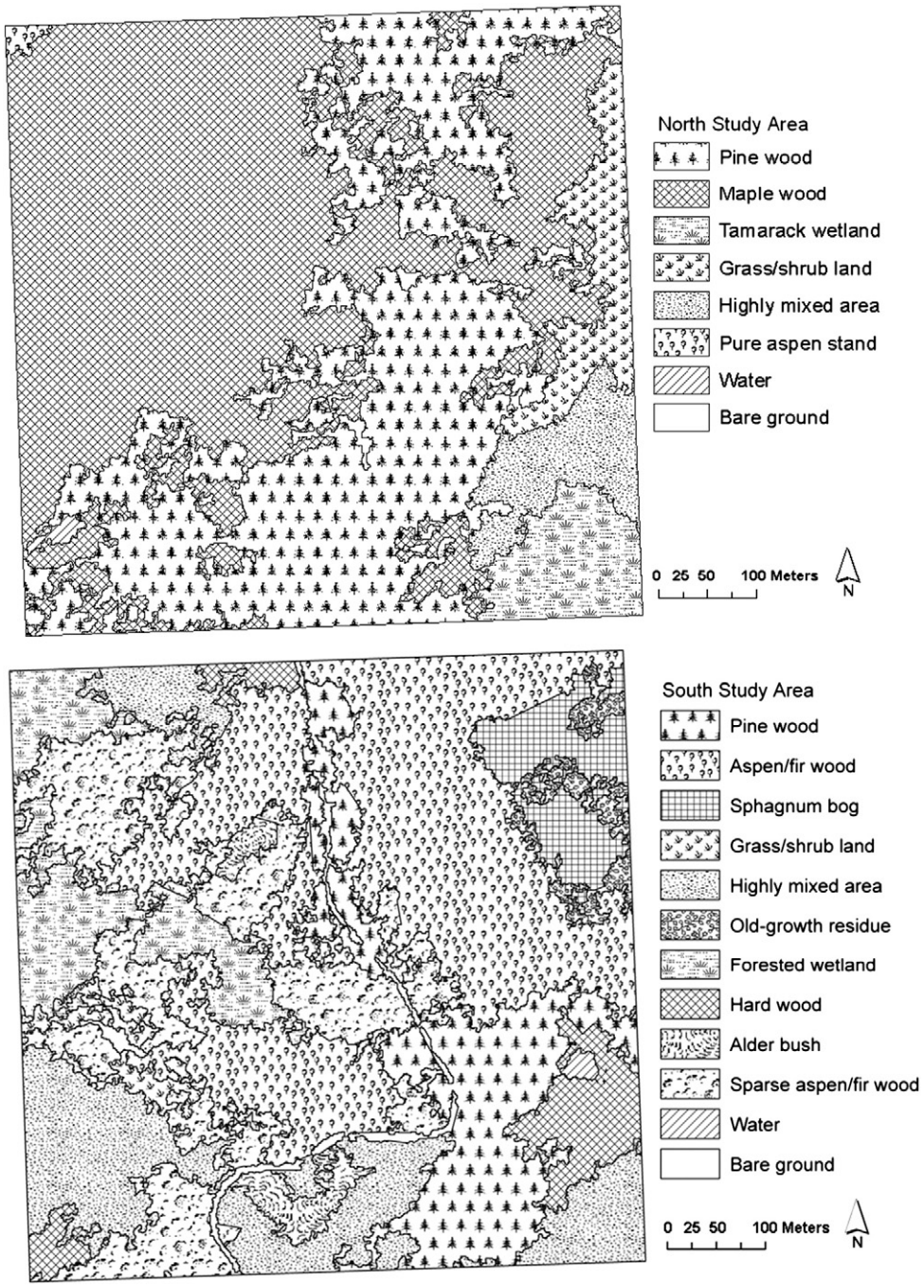


**Fig. 3.** Forest communities in the study areas, delineated using segmentation classification of a summer IKONOS image (UTM projection) and manual post-processing based on field survey; the study areas appear tilted due to the projection.

the linear spectral unmixing was performed. Manually digitized pure coniferous and deciduous cover patches were used to generate end-member signatures. The non-vegetated cover signature was extracted from reflectance of homogeneous bare grounds such as stream point bars, exposed lakebeds, and roads. We reviewed these signatures afterwards with feature space representations of the first three components from principal component analysis and minimum noise fraction transformation (Wu & Murray, 2003). Resultant end-member images are provided in Fig. 4, showing respective portions (in 0–1 ratios) of deciduous, coniferous, and non-vegetated covers at the sub-pixel level across the region surrounding study areas.

We estimated community level phenology from dominant population phenologies in each particular community and corresponding composition information derived from field survey and high resolution image analysis (as described earlier in detail). A linear function was used to mathematically compose community phenology from presence and abundance estimates: community phenology = deciduous proportion × deciduous population phenology + coniferous proportion × coniferous population phenology. When more than one dominant deciduous species were present in a community, a weighted average of population phenologies was used. The weights were empirically assigned to relevant populations according to their respective population size rankings.

### 3.2.3. Landscape phenology composition

In compositing landscape-level phenology representation, we used two approaches. The first was aimed at generating a LP index

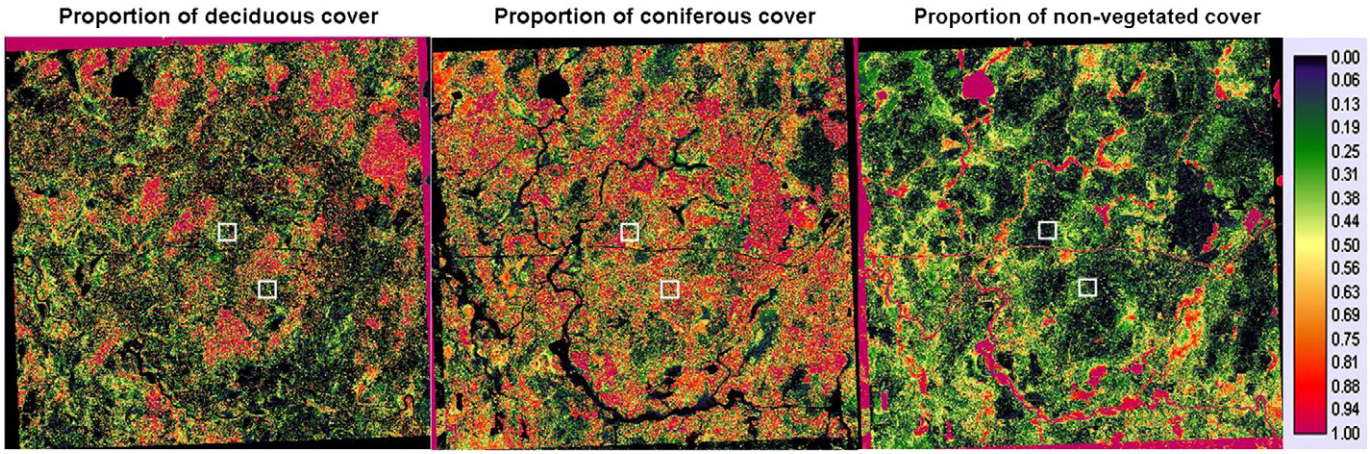


**Fig. 4.** Sub-pixel proportion maps of deciduous, coniferous, and non-vegetated covers generated from multi-temporal linear spectral unmixing; represented in grayscale showing proportions from 0 (black) to 1 (white) for print and color ramp for online; the maps show surrounding areas of the study site with the WLEF flux tower site visible (a circular feature) in the center of area; the study areas are located near the WLEF tower, as marked with squares.

compatible with satellite-measured land surface reflectance, thus sacrificing its biological significance. In the study area, coniferous canopy cover, deciduous canopy cover, and mixtures of both were the primary vegetation types. In addition, small areas of grass/shrub land were present. Given that phenologies of coniferous, deciduous, and grass cover contribute to LSP differently, a simple spatial aggregation of community phenologies would not produce a satisfactory index. Therefore, we used an alternative method to address the varying degrees of reflectance contribution from these vegetation covers. Using two 2007 QuickBird images, NDVI values were computed for homogeneous deciduous, coniferous, and grass covers. Then linear regression models were fitted between the QuickBird-based NDVI and averaged population phenologies for deciduous, coniferous, and grass covers respectively. Due to a lack of grass phenology measurements in 2008 and 2009, speckled alder phenology was used as a surrogate of grass phenology given its strong correlation (0.935, $p$-value$<$0.01) among sampled species with averaged grass phenology photographed in 2007 (Liang & Schwartz, 2009). The linear regression model between QuickBird NDVI of grass covers and speckled alder phenology (also based on 2007 data) was applied to simulate grass phenology in 2008 and 2009. Speckled alder was also present as a major understory species and dominant shrub in relatively open areas that would be recognized as grass/shrub land from satellite imagery. As such, population phenologies of deciduous and coniferous species were adjusted by applying respective fitted linear models to generate calibrated community phenologies. Simulated grass phenology was combined into deciduous phenology for communities with observed grass/shrub patches and/or highly gapped canopies. Finally we aggregated all community level phenological estimates to represent phenology at the landscape level, and then spatially averaged to generate a land cover calibrated LP index.

The second approach was intended to retain the biological significance of phenology protocols used in the field throughout the scaling process. Due to different phenological characteristics of deciduous and coniferous phenology, an index fitting both was not practical. Therefore subcomponent LP indices for deciduous and coniferous covers were derived separately. We achieved this by forming simplified landscape scenarios with either pure deciduous or pure coniferous cover. The deciduous portions or the coniferous portions of community phenologies were aggregated over the landscape with an area-weighted averaging approach. Because each cover type only occupied a fraction of the study area, that fractional value was used to adjust the derived index into the value range of the phenology protocol.

### 3.3. Deriving land surface phenology

Satellite-based LSP was derived from MODIS VI time series. Two study area frames (625×625 m each, in WGS84 UTM15N projection) were transformed to the sinusoidal projection that is native to global MODIS products, since a UTM-sinusoidal reprojection yields higher accuracy compared to its reverse (Seong et al., 2002). The reprojected study area frames were then overlaid on a 250 m MODIS VI grid, showing twelve pixels each overlapping the north or the south study areas respectively (see Fig. 1 and Table 2). The geolocation accuracy for MODIS Terra is 45 m (Qu et al., 2006), which is relatively small compared to the study area size, allowing for fairly precise spatial alignment and comparison of satellite and ground data. Therefore, instead of using the mean value of the pixels overlapping with the study areas, we utilized a pixel-by-pixel method to extract VI values more precisely. In particular, the respective portions of VI values affected by the study areas were estimated using a weighted extraction approach. This was facilitated with high resolution QuickBird images, taking into account brightness variations within each MODIS pixel. By overlaying the study area frames and MODIS pixel grid over a QuickBird-based NDVI change map, we calculated mean NDVI change within the intersected polygons (intersected portions of MODIS pixels and the study area, Table 2). The ratios of mean NDVI change values within intersected portions to the mean NDVI change values within the entire corresponding MODIS pixels were used as weights. Such weights quantified the degree to which each MODIS pixel acquired information from specific fractions of the study areas. All the derived MODIS VI fractions were then averaged according to their areas over the north study area and south study area respectively.

We constructed MODIS VI time series for spring 2008 and spring 2009. In the MODIS VI product, each pixel value was accompanied with a reliability rank, as well as more detailed quality assessment ratings. Only data with all "good" pixels (pixel reliability = 0) were used. The majority of VI data before mid-April were either contaminated with clouds (pixel reliability = 3) or affected by snow/ice coverage (pixel reliability = 2), occasionally mixed with a few marginal pixels (pixel reliability = 1). So we used only the time period from mid-April to mid-July to form spring VI time series, which matched the field data collection time span and the growing season onset time in the study area. Actual image acquisition dates (instead of the 16-day composite period start dates) were used for improved accuracy, given the relatively small study areas (Table 2). For a study area, when there were inconsistent dates associated with the

**Table 2**
Summary of MODIS pixels intersecting the study areas; with QuickBird NDVI derived weights, area fractions within the study area, and data acquisition dates used in day of year (DOY).

| | Pixel# | Weight | Fraction | Dates used in 2008 (DOY) | Dates used in 2009 (DOY) |
|---|---|---|---|---|---|
| North study area | 1 | 1.144 | 0.306 | 126,144,149,165,185,194 | 105,118,142,155,171,187,194 |
| | 2 | 1.021 | 0.821 | 121,144,149,165,185,194 | 105,118,142,155,171,187,194 |
| | 3 | 1.001 | 0.817 | 121,144,149,165,185,194 | 105,118,142,155,171,187,194 |
| | 4 | 1.102 | 0.280 | 121,144,149,165,185,194 | 105,118,142,155,171,187,194 |
| | 5 | 0.996 | 0.411 | 121,144,149,165,185,194 | 102,118,142,155,171,187,194 |
| | 6 | 0.999 | 0.987 | 126,144,149,165,185,194 | 100,118,142,155,171,187,194 |
| | 7 | 1.005 | 0.970 | 121,144,149,165,185,194 | 101,118,142,155,171,187,194 |
| | 8 | 1.134 | 0.332 | 121,144,149,165,185,194 | 105,118,142,155,171,187,194 |
| | 9 | 0.997 | 0.397 | 126,144,149,165,185,194 | 105,118,142,155,171,187,194 |
| | 10 | 1.005 | 0.867 | 121,144,149,165,185,194 | 105,118,142,155,171,187,194 |
| | 11 | 1.017 | 0.841 | 121,144,149,165,185,194 | 105,118,142,155,171,187,194 |
| | 12 | 1.054 | 0.239 | 121,144,149,165,190,194 | 101,118,135,155,171,187,194 |
| South study area | 13 | 1.074 | 0.214 | 121,144,149,165,183,194 | 105,118,142,155,171,187,194 |
| | 14 | 1.020 | 0.667 | 121,144,149,165,183,194 | 105,118,142,155,171,187,194 |
| | 15 | 0.997 | 0.670 | 121,144,149,165,183,194 | 98,118,135,155,171,187,194 |
| | 16 | 1.021 | 0.251 | 126,144,149,165,183,194 | 105,118,142,155,171,187,194 |
| | 17 | 1.003 | 0.303 | 121,144,149,165,183,194 | 105,118,142,155,171,187,194 |
| | 18 | 1.000 | 0.960 | 121,144,149,165,183,194 | 105,118,142,155,171,187,194 |
| | 19 | 1.000 | 0.991 | 121,144,149,165,183,194 | 105,118,142,155,171,187,194 |
| | 20 | 1.005 | 0.438 | 121,144,149,165,183,194 | 98,118,142,155,171,187,194 |
| | 21 | 1.054 | 0.409 | 121,144,149,165,183,194 | 101,118,142,155,171,187,194 |
| | 22 | 1.001 | 0.988 | 121,144,149,165,183,194 | 105,118,142,155,171,187,194 |
| | 23 | 0.998 | 0.969 | 121,144,149,165,183,194 | 105,118,142,155,171,187,194 |
| | 24 | 1.046 | 0.331 | 121,144,149,165,183,194 | 105,118,142,155,171,187,194 |

intersected MODIS pixels, the median date was used. We then fitted the selected VI values with a logistic function for generating continuous LSP curves according to Zhang et al. (2003) and also provided as follows:

$$y(t) = \frac{c}{1 + e^{a+bt}} + d$$

where $t$ is time in day of year, $y(t)$ is the VI value at time $t$, $a$ and $b$ are fitting parameters, $c+d$ is the maximum VI value, and $d$ is the initial background VI value. Since only good quality data were used, no smoothing technique was applied to any of these data series. The first maximum curvature change date was extracted to represent the time of SOS according to algorithms detailed in Zhang et al. (2003). This approach was adopted to generate MODIS Global Land Cover Dynamics product (MOD12Q2) (Ganguly et al., 2010; user guide: http://www-modis.bu.edu/duckwater1/mod12q2/doc/MOD12Q2_V4_user_guide.doc.pdf) and used in national canopy phenology monitoring systems (Hargrove et al., 2009). Given its broad application, we chose this method as the primary validation target in our study. As the MODIS LSP data (from MOD12Q2) for 2008 and 2009 are not yet available through the ORNL DAAC over the study area to date, and because Global Land Cover Dynamics product will have a lower (500 m for collection 5) resolution (Ganguly et al., 2010), we generated SOS dates manually for our study areas from the 250 m VI data following the prescribed methods. The fitted logistic curves are provided in Fig. 5.

### 3.4. Validating land surface phenology

We made a comprehensive comparison of the LP index and LSP over the same spatial extent and temporal span: first for the MODIS pixel intersected fractions; then for LP indices and LSP time series over both study areas. For phenological comparison at the fraction level, only VI values acquired during the field observation time period were used. All the VI values acquired on dates overlapped the field data collection time periods were selected for both study areas and for both years. They were compared with LP subset values of study area fractions for each corresponding date. When there was no ground observation available on a date with satellite data, we used the average of two adjacent LP index values. VI and LP index values were standardized using the formula: (observation − minimum)/(maximum − minimum). Both reflectance-weighted VI values extracted for overlapping fractions and VI values of entire intersected MODIS pixels were compared to LP index values, with RMSE calculated accordingly.

For the entire north study area or south study area, interpolated LSP was compared to land cover calibrated LP index values on all dates with ground observations. In addition to RMSE statistics, relative errors (RE) were computed following the formula given in Heinsch et al. (2006):

$$RE = \sum_{i=1}^{n} \frac{(LSP - LP)}{LP} / n \times 100\%$$

where $LSP$ is land surface phenology prediction, $LP$ is landscape phenology index values, and $n$ is the number of pairs of LSP and LP measurements. Standardization using the previously given formula was applied to all LSP and LP values. Finally, we compared LSP SOS dates to full bud burst (FBB) dates of the LP indices (for deciduous and coniferous respectively) as well as dominant population phenologies. As only deciduous and coniferous based LP indices retained their original meaning in terms of the phenological protocols, FBB dates of these two subcomponent indices were used as phenological markers for comparison.

## 4. Results

### 4.1. VI-based land surface phenology and landscape phenology

We found a significant linear relationship (all $p$-values$<0.01$) between MODIS VI and the land cover calibrated landscape phenology (LP) index for the respective fractions of MODIS pixels contained in the study areas (Fig. 6). Most pairs of VI and LP index were distributed around the 1:1 line. These pairs demonstrated pixel-by-pixel correspondences between VI and LP index. As indicated by higher coefficients of determination and smaller RMSE terms, the linear dependency on LP was stronger for EVI (Fig. 6c and d) than for NDVI (Fig. 6a and b). Similarly, reflectance-weighted VI (Fig. 6b and d) had stronger dependency on LP than the raw VI values (Fig. 6a and c).

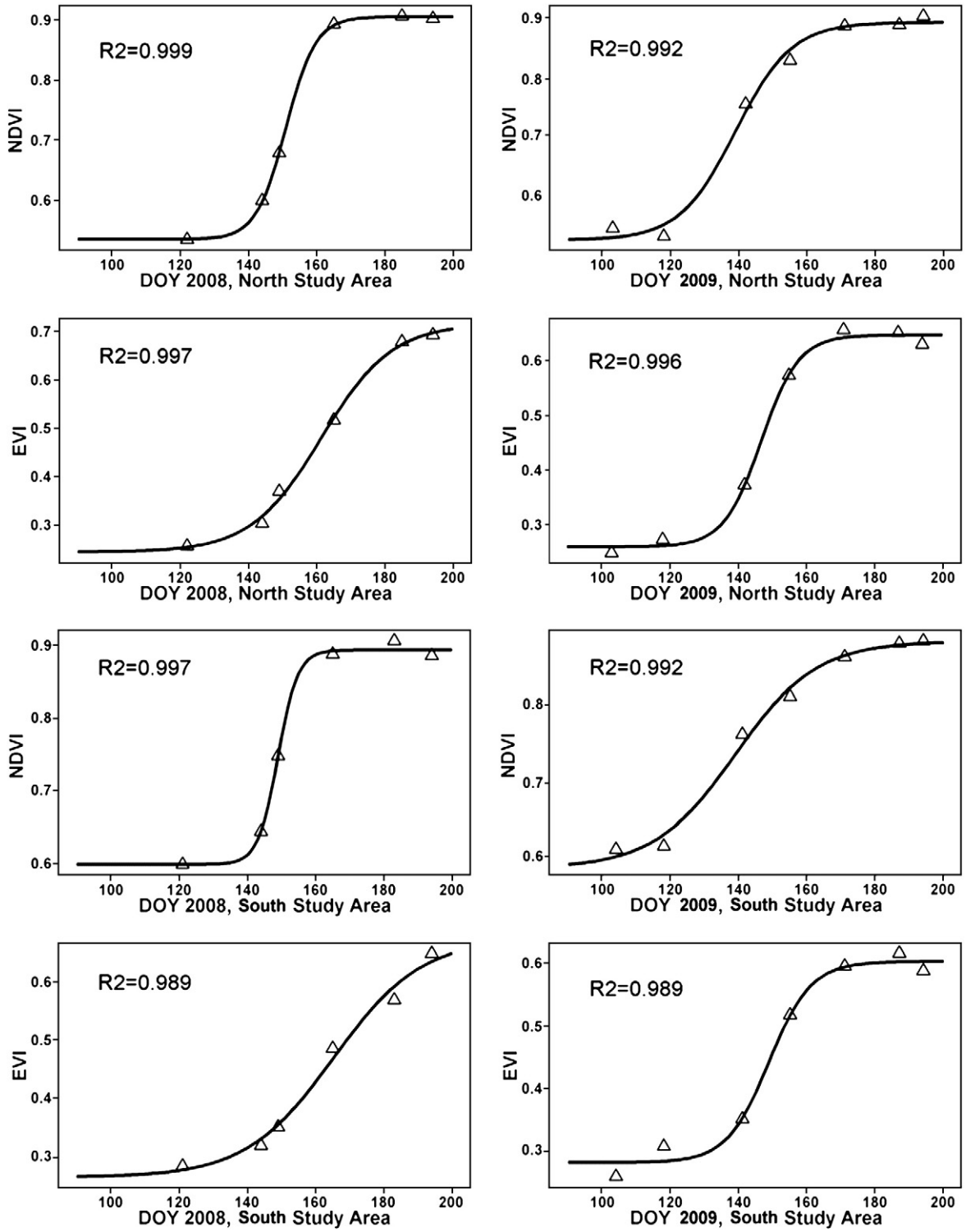


**Fig. 5.** Logistic growth curves fitted for vegetation indices (VI) data points for the two study areas during spring 2008 and spring 2009; triangles represent VI values corresponding to acquisition time in day of year (DOY); only VI values with good pixel quality were used, with no data smoothing; coefficients of determination (R2) are provided.

Comparison of logistic function fitted land surface phenology (LSP) and LP index time series suggested systematic differences (Fig. 7). In most cases, LSP departed from the LP index overall with large relative errors. The general relationship between LSP and LP appeared to be exponential, showing an increased influence of LP on LSP over time. Specifically in the early stages of LP development, there was relatively slow corresponding VI change compared to later stages. In other words, VI-based LSP change appeared to be less responsive to LP change at the early stages, but were affected heavily by LP change at later stages of phenological development, as also indicated by gradient variations. Such patterns were recognizable for most cases, except that for 2009 NDVI data (Fig. 7e and g) where the relationships showed a more consistent correspondence to LP with smaller errors.

### 4.2. Start of season validation

Start of season (SOS) dates derived from VI time series and full bud burst (FBB, equal to 400 in our protocol) dates of deciduous and

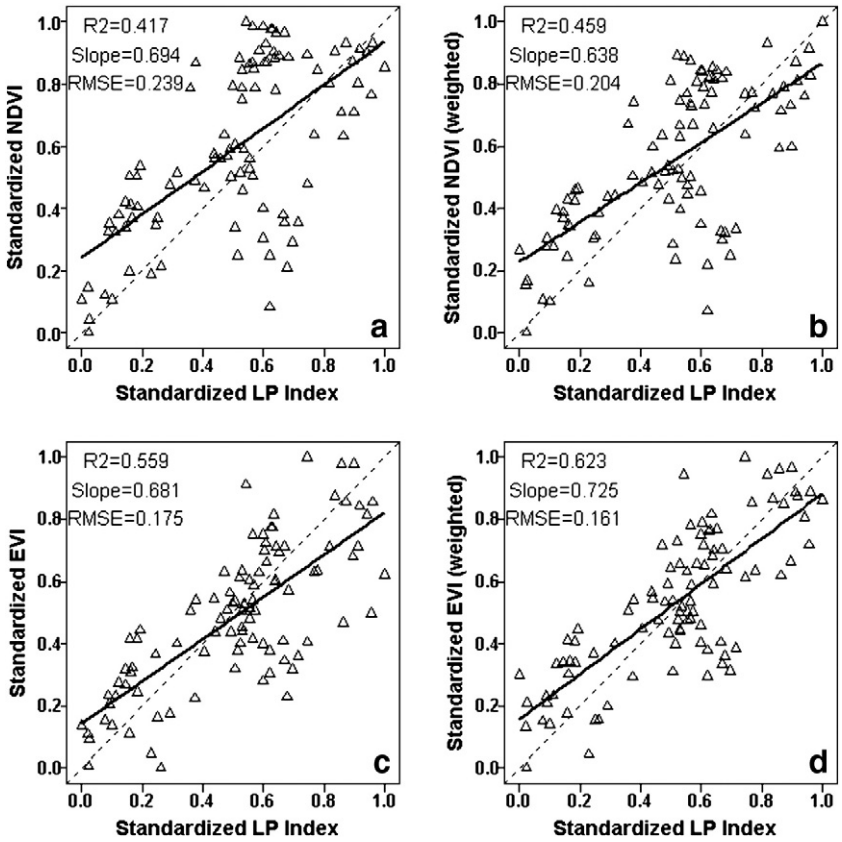


**Fig. 6.** Comparison of VI and landscape phenology (LP) index for MODIS pixel intersected fractions; triangles represent pairs of VI and LP index for the two study areas and for both 2008 and 2009; coefficients of determination (R2), slope, root mean square errors (RMSE), and 1:1 lines are provided; figures on the right hand side (b and d) are for pixel-by-pixel reflectance-weighted VI values; figures on the left hand side (a and c) are for VI values of complete MODIS pixels overlapped with the study areas; both VI and LP index were standardized using the formula: (observation − minimum)/(maximum − minimum).

coniferous LP indices showed general agreement (Table 3). SOS dates were closer to FBB dates of deciduous LP, but the coniferous FBB were consistently later than both deciduous LP and VI-derived SOS. EVI-based SOS was very close to deciduous FBB, with a maximum difference of only two days; yet NDVI-based SOS was different from deciduous FBB by as much as fifteen days. The large errors were mainly in 2009; for 2008, NDVI-based SOS appeared to be quite close to deciduous FBB dates, with a maximum difference of four days. Nonetheless, both SOS and FBB dates in 2009 were consistently earlier than 2008. The relationships between LSP time series with SOS dates, and LP time series with FBB dates are illustrated in Fig. 8. When compared with FBB dates of population phenologies (Table 4), all species appeared to have had earlier bud burst in 2009 than 2008, confirming what was seen from LSP and LP. In addition, according to EVI-based SOS, the south study area was predicted to be several days earlier than the north study area in both years. This was in agreement with FBB dates from the deciduous LP index. However, NDVI-based SOS predicted a contrary relationship for spring 2008.

FBB dates of population phenology also matched LSP predictions with less than seven days of absolute errors for EVI-based SOS measures. All species appeared to indicate an earlier spring onset in 2009. However, population phenologies did not show consistent regional difference between the two study areas for a common species. Moreover, phenology of populations with greater sample size appeared to be better predicted by LSP. As indicated in Fig. 9, when tested with all deciduous species, mean absolute errors notably decreased with the increase of population sample size, despite large variability at the lower end with populations of smaller sample size. When a population was represented with over 80 individuals, absolute errors fell below three days.

## 5. Discussion and conclusions

Landscape phenology (LP) validated the overall accuracy of land surface phenology (LSP) in regards to the SOS dates estimation. In particular, MODIS EVI-based estimates of SOS dates using a logistic function (Zhang et al., 2003) predicted the full bud burst (FBB) dates of deciduous LP with less than two days of absolute error. NDVI-based SOS underperformed EVI with larger errors in tracing LP temporal changes. The FBB dates of deciduous LP marked the time when deciduous tree foliage began to significantly transform land surface reflectance. Therefore SOS dates (estimated using maximum curvature change) appear robust in capturing this critical spring phenology event underlying landscape greenness transitions. In addition, both deciduous and coniferous LP indices suggested that 2009 had an earlier spring onset than 2008. SOS dates from NDVI and EVI time series indicated the same temporal difference with LSP measures across both years. Hence, LSP successfully captured inter-annual variations of LP in the study region. This strongly suggests that LSP has the capability to monitor long-term LP change across multiple years in relation to global climate change. Further, our study has directly connected LSP SOS to conventional phenological events with explicit biological meaning.

Comparing LP in the two study areas, deciduous FBB occurred a few days earlier in the south study area for both years (2008 and 2009). This is reflected by EVI-based SOS dates, which also predicted earlier spring onset in the south for both years. However, NDVI-based SOS dates did not show a consistent regional difference across years. For large-scale studies, the ability of LSP to detect ecologically coherent regional variations has been well-tested (de Beurs & Henebry, 2005; White et al., 2005; Zhang et al., 2007). In our study,

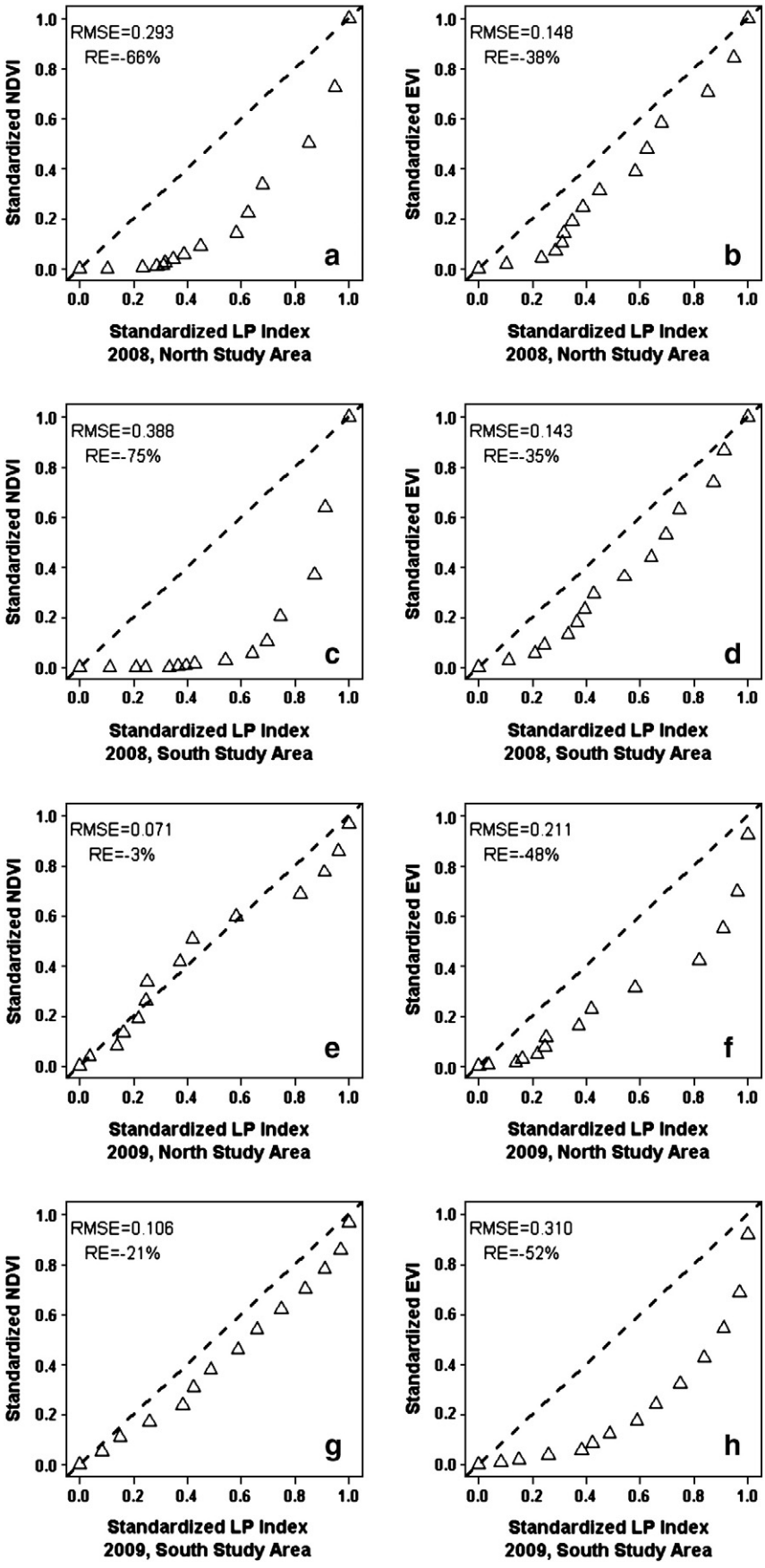


**Fig. 7.** Cross comparison of interpolated VI-based land surface phenology (LSP) against landscape phenology (LP) index; triangles represent pairs of LSP and LP index respective to the study area, observation year, and VI types; root mean square errors (RMSE), relative errors (RE), and 1:1 lines are provided; both LSP and LP index were standardized using the formula: (observation − minimum)/(maximum − minimum).

**Table 3**
Start of spring season (SOS) dates in day of year (DOY) derived from MODIS NDVI and EVI time series; and full bud burst (FBB) dates for deciduous and coniferous landscape phenology (LP) indices in the two study areas for 2008 and 2009.

| | Year | SOS_NDVI | SOS_EVI | FBB_deci | FBB_coni |
|---|---|---|---|---|---|
| North | 2008 | 139 | 138 | 138 | 146 |
| Study area | 2009 | 122 | 133 | 135 | 142 |
| South | 2008 | 141 | 136 | 137 | 146 |
| Study area | 2009 | 116 | 132 | 131 | 143 |

the sensitivity of MODIS LSP to spatial variations of LP within a small region is partly validated. Since the two study areas were fairly close to each other, it is not surprising that satellite signals may not detect the differences for every case. This may also imply that EVI performed better than NDVI in regards to estimating spatially consistent SOS dates, perhaps due to the enhanced sensitivity to saturated greenness change (Huete et al., 2002), which may be contributed by evergreen canopies in the mixed forest.

We also found correspondence between LSP and population phenology of deciduous species for SOS and FBB dates. The absolute errors of LSP SOS from population phenology FBB were less than a week. Larger errors of LSP prediction were associated with species with relatively small sample size. Population phenologies of dominant deciduous species (large sample size, such as red maple, quaking aspen, and sugar maple) were predicted with mean absolute errors (for both years) less than three days. Although larger sample size is commonly associated with smaller errors, in the context of this study, such relationships may imply that LSP would be powerful in detecting species-level phenology provided that the species of interest are a primary contributor to the landscape greening process. However, population phenology of species present in both study areas did not show any regional differences as detected by LP, implying the importance of accounting for landscape heterogeneity. So we highly recommend that landscape characterization be employed for precisely aggregating phenological observations when deriving LP indices, so as to retain sensitivity to spatial variability. Nevertheless, field data collected merely for unrepresentative species in an area, or observations with insufficient sample size could generate large errors when used to validate LSP. Further, in areas with more homogeneous vegetation covers, LSP could potentially provide more accurate predictions of species-level phenology.

While SOS dates validation demonstrated the overall accuracy of LSP, comparison of the LP and LSP time series implied strong systematic difference. Initial stages of LP corresponded to very little LSP change; and the later stages of LP had stronger influence on LSP. This is explained by the field phenology protocol we employed in describing phenological development from buds to leaves. Before bud burst and leaves out, the effects of phenology on landscape reflectance change was minimal. However after leaves came out, LP started to contribute to LSP change in a significant manner. As LSP does not possess the ability to detect phenophases that do not incur land surface reflectance change, such temporal mismatch of the LP and LSP time series was anticipated. Moreover, the exponential relationship between LP and LSP suggested a generally accelerated influence of ground phenological development on satellite signals during the course of spring onset. This further implies that phenological stages of leaf development likely have a gradually increasing effect on land surface greening at the start of season.

To our knowledge, phenology datasets collected using traditional observer-based approaches in a highly intensive manner have not been available prior to this study. High density samples from dominant tree species significantly increase the representativeness of in situ

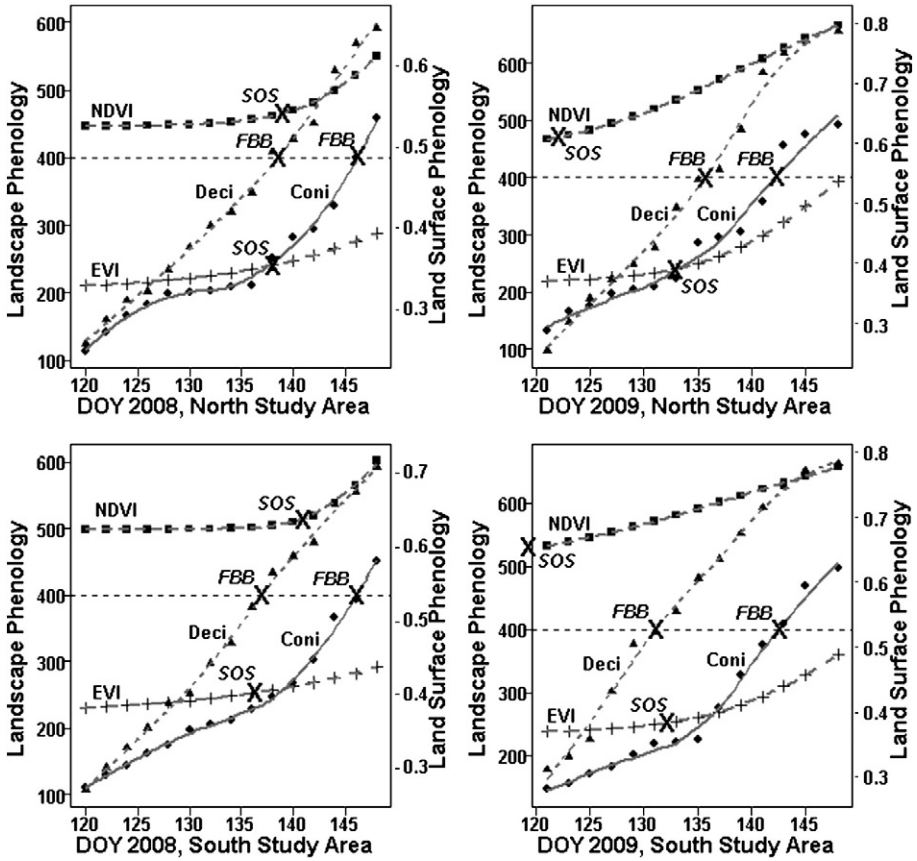


**Fig. 8.** Parallel comparison of interpolated (using logistic function, see Fig. 5) land surface phenology in VI ratio and landscape phenology indices for deciduous (Deci) and coniferous (Coni) covers respectively; full bud burst (FBB) and start of spring season (SOS) marked in correspondence with dates listed in Table 3; *x* axis unit in day of year (DOY); reference line for FBB (400 in field protocol) is provided.

**Table 4**
Full bud burst (FBB) dates in day of year (DOY) of dominant populations (sample size >20, provided in parenthesis after species name) in the two study areas for 2008 and 2009.

| | | Dominant population full bud burst (FBB) dates | | | | |
|---|---|---|---|---|---|---|
| North | Year | Balsam fir (21) | Basswood (24) | Red maple (83) | Sugar maple (173) | |
| Study | 2008 | 146 | 142 | 137 | 137 | |
| Area | 2009 | 142 | 139 | 135 | 135 | |
| South | Year | Balsam fir (99) | Quaking aspen (119) | Red maple (82) | Speckled alder (34) | White birch (23) |
| Study | 2008 | 146 | 136 | 137 | 135 | 143 |
| Area | 2009 | 142 | 128 | 136 | 131 | 138 |

phenology observations over a vegetated area. In order to derive spatially compatible LP indices for validating coarse-scale satellite-based LSP, such intensive survey is highly preferred. Also, the bi-daily frequency observation taken over a month-long period (which overlapped the time of growing season onset) provides opportunities to validate the entire springtime LSP time series, in addition to a single extracted phenological marker (e.g., SOS). As discussed previously, such comparison helps reveal detailed relationships between ground biophysical processes and satellite reflectance-based measures. Wu et al. (2006) pointed out that ideal up-scaling entails that the grain size of sampling or observation is smaller than the spatial or temporal dimension of the structures or patterns of interest, whereas the sampling extent is at least as large as the extent of the phenomenon under study. Our data exactly matched these preferences.

In this study, landscape scaling was critical in allowing a proper utilization of field data for satellite phenology validation. Through scaling, in situ phenology and landscape heterogeneity were effectively integrated to produce a LP index, at a spatial scale compatible with coarse resolution satellite measures. The nested hierarchical organization of an ecosystem provided a predefined scaling ladder to connect individual plant behaviors to processes at the landscape level. Such an individual–population–community–landscape (ecosystem patch) hierarchy worked effectively to fully integrate discrete in situ observations into a more holistic ecological context, through a practical stepwise procedure. The correspondence of these levels of ground-based phenology to satellite-pixel-based phenology depends on patch characteristics of the landscape under examination. In this study, 250 m LSP corresponds to landscape phenology given the smaller-sized communities. However, in areas with simpler vegetation, LSP may directly correspond to community phenology or even population phenology, which is a special case of landscape phenology. Therefore, application of this scaling model can be simplified or sophisticated according to the need to address vegetation covers with varying degrees of complexity. In a broader context, assigning phenological behaviors to population, community, and ecosystem levels may improve our understanding of plant phenology, which usually focus on individual plants or remotely sensed vegetation at two ends of a spatial scale continuum. The predefined hierarchical ecosystem organization offers an opportunity to assign phenological measures to respective levels within the structure, thus expanding the scope of phenological monitoring to integrated levels of ecosystems. Consequently, phenological studies can be investigated not only for plant individuals, but also for selected populations, plant communities, and landscape patches. Therefore, in the broader sense, this scaling model is not merely useful for generating landscape-level phenological indices for validating LSP, but also opens up a multi-level perspective to address phenological inquiries and associated monitoring and modeling tasks.

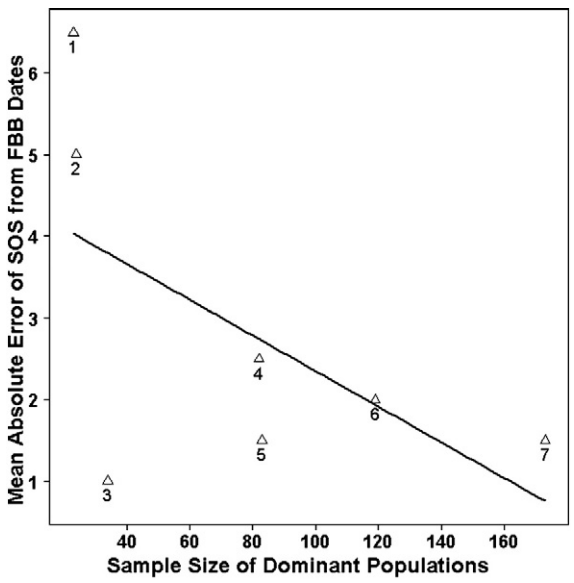


**Fig. 9.** Mean absolute error (in days) of start of spring season (SOS) dates (EVI-based) from full bud burst (FBB) of population phenologies decreasing with the increase in sample size (with larger variability at the low end); triangles represent mean absolute errors of SOS from FBB dates corresponding to sample size; deciduous species: 1—white birch (S), 2—basswood (N), 3—speckled alder (S), 4-red maple (S), 5-red maple (N), 6-quicking aspen (S), and 7-sugar maple (N).

We identified a number of drawbacks and limitations in our study. First, in order to collect data in an intensive manner, labor efficiency may be a major limitation for replicating this work. Fortunately, once such validation work is completed for an ecoregion, there may be less need to repeat within similar environments, as far as validating LSP. However, given the different phenological regimes in different biomes, we recommend similar studies be conducted in additional regions representing different phenology regimes. For instance, validating the accuracy of LSP in semiarid or Mediterranean climates, where SOS has proven more problematic than other temperate land regions (White et al., 2009), could be particularly useful. To facilitate conducting intensive phenological observations, analyses addressing monitoring efficiency will be provided in a forthcoming paper related to this project focusing on ground phenology patterns (Schwartz et al., in preparation). Secondly, our landscape scaling approach involved using high-resolution satellite imagery and relatively sophisticated techniques for landscape heterogeneity characterization. Such an approach is necessary for a mixed forest for which spatial heterogeneity cannot be ignored for precise scaling. But for more homogeneous vegetation covers, such as pure hardwood forests, grasslands, and croplands, simpler scaling approach may be sufficient. Thirdly, only dominant species in each community were used to generate community phenology and LP. Other species may either be minor in regards to lacking a sufficient sample size (for population phenology estimate) or being hardly noticeable in the field. More exhaustive survey of those species may help improve the accuracy of LP index, but will be difficult to achieve in the field. Further, coniferous species with unobservable spring phenology were also omitted. Yet all vegetation covers were included in LP index through the generalization of community composition into deciduous/coniferous end-members. Nevertheless, the omitted species posed a limitation to our LP derivation. Fourthly, in accounting for understory grass/shrub phenology in up-scaling, one shrub species (speckled alder) phenology was used based on its similarity with grass phenology measured in a

previous year (2007), and in relation to QuickBird detected grass/shrub land cover reflectance. This approach provisionally addressed the need to account for understory phenology, especially for areas with highly gapped canopies or grass/shrub cover. However, synchronized survey of understory phenology would be definitely desirable to increase the validation accuracy. Regarding satellite data, an additional limitation we noticed in our study is related to fitting LSP time series using 16-day MODIS VI data. Large absolute errors (up to fifteen days) were found with NDVI-based SOS and deciduous LP index FBB in 2009. Moreover, in comparing LP index and NDVI time series in 2009, the patterns with RMSE and relative errors departed from the consistency among other cases. This may be due to the corresponding fitted NDVI curves failing to capture the actual LP process. As such, based on comparisons discussed here and previously, we recommend that EVI be used as a preferred index for LSP derivation, though additional tests may be needed to evaluate the accuracy difference of NDVI and EVI based LSP in a more comprehensive manner.

According to our results, logistic model-calculated LSP measurements are relatively accurate in tracking LP processes and variations. The logistic model holds a biophysical inference in terms of being a "growth curve" model. Therefore, both the validation results and its conceptual coherency, encourage further application of this method as a primary technique for deriving phenological information from satellite data. Uncertainties still remain in relation to the 16-day temporal resolution of MODIS VI products, which was composited from daily data for cloud contamination removal. Analysis showed that 6–16 days temporal resolution would provide optimal estimates of LSP seasonal markers, yet higher temporal resolution is preferred to detect phenological changes occurring within the composite window period (Zhang et al., 2009). In our study, when actual acquisition dates were used, the temporal gaps of MODIS data were up to 23 days (DOY121–144 for 2008, DOY118–141 for 2009). Although the logistic function may reduce some uncertainties, ability to detect ground phenological transition dates may be unavoidably reduced. Therefore we remain convinced that temporal resolution of the MODIS MOD13Q1 product is a major source of error for deriving LSP estimates, and will compromise the accuracy of any LSP time series interpolation.

In summary, we demonstrated that intensive field data and adequate landscape scaling are two keys to validating satellite-based LSP. Should validation work be pursued in other regions, these two elements are critical. Further, MODIS LSP provides generally reliable measurements of LP for the temperate mixed forest. Continued efforts are still needed to further improve the synergy among field observation and remote sensing of vegetation phenology. We especially suggest that such efforts be directed towards validating LSP using similar methods to our study in additional biomes of interest.

## Acknowledgements

We would like to acknowledge those who contributed to the field campaigns and phenology data collection: Rachael Dearing, Audrey Fusco, Alan Halfen, Jonathan Hanes, Joshua Hatzis, Patricia O'Kane, and Virginia Seamster. Jonathan Hanes also contributed to field work organization and data entry. Robert B. Cook and Suresh-Kumar Santhana-Vannan from ORNL DAAC provided valuable help in using MODIS data. Changshan Wu provided constructive advice on data analysis. NASA supplied the QuickBird and IKONOS images. We also thank six anonymous reviewers for their helpful comments. This project was supported by the National Science Foundation under grants BCS-0649380 and BCS-0703360.

## References

Ahrends, H., Brügger, R., Stöckli, R., Schenk, J., Michna, P., Jeanneret, F., et al. (2008). Quantitative phenological observations of a mixed beech forest in northern Switzerland with digital photography. *Journal of Geophysical Research*, *113*, G04004.

Burrows, S., Gower, S., Clayton, M., Mackay, D., Ahl, D., Norman, J., et al. (2002). Application of geostatistics to characterize Leaf Area Index(LAI) from flux tower to landscape scales using a cyclic sampling design. *Ecosystems*, *5*, 667–679.

Cheng, P., Toutin, T., Zhang, Y., & Wood, M. (2003). QuickBird-geometric correction, path and block processing and data fusion. *Earth Observation Magazine*, *12*, 24–30.

Clayton, M., & Hudelson, B. (1995). Confidence intervals for autocorrelations based on cyclic samples. *Journal of the American Statistical Association*, *90*, 753–757.

Cleland, E. E., Chuine, I., Menzel, A., Mooney, H. A., & Schwartz, M. D. (2007). Shifting plant phenology in response to global change. *Trends in Ecology & Evolution*, *22*, 357–365.

de Beurs, K. M., & Henebry, G. M. (2004). Land surface phenology, climatic variation, and institutional change: Analyzing agricultural land cover change in Kazakhstan. *Remote Sensing of Environment*, *89*, 497–509.

de Beurs, K. M., & Henebry, G. M. (2005). Land surface phenology and temperature variation in the International GeosphereBiosphere Program high-latitude transects. *Global Change Biology*, *11*, 779–790.

de Beurs, K. M., & Henebry, G. M. (2008). Northern annular mode effects on the land surface phenologies of northern Eurasia. *Journal of Climate*, *21*, 4257–4279.

de Beurs, K. M., & Henebry, G. M. (2010). Spatio-temporal statistical methods for modeling land surface phenology. In I. L. Hudson, & M. R. Keatley (Eds.), *Phenological research: Methods for environmental and climate change analysis* (pp. 177–208). Dordrecht: Springer.

DeFries, R., Hansen, M., & Townshend, J. (2000). Global continuous fields of vegetation characteristics: A linear mixture model applied to multi-year 8 km AVHRR data. *International Journal of Remote Sensing*, *21*, 1389–1414.

Doktor, D., Bondeau, A., Koslowski, D., & Badeck, F. -W. (2009). Influence of heterogeneous landscapes on computed green-up dates based on daily AVHRR NDVI observations. *Remote Sensing of Environment*, *113*, 2618–2632.

ERDAS Inc. (2008). *ERDAS imagine field guide.* Atlanta, Georgia: ERDAS Inc.

Ewers, B., Mackay, D., Gower, S., Ahl, D., Burrows, S., & Samanta, S. (2002). Tree species effects on stand transpiration in northern Wisconsin. *Water Resources Research*, *38*, 1103.

Fischer, A. (1994). A model for the seasonal variations of vegetation indices in coarse resolution data and its inversion to extract crop parameters. *Remote Sensing of Environment*, *48*, 220–230.

Fisher, J. I., & Mustard, J. F. (2007). Cross-scalar satellite phenology from ground, Landsat, and MODIS data. *Remote Sensing of Environment*, *109*, 261–273.

Fisher, J. I., Mustard, J. F., & Vadeboncoeur, M. A. (2006). Green leaf phenology at Landsat resolution: Scaling from the field to the satellite. *Remote Sensing of Environment*, *100*, 265–279.

Friedl, M., Henebry, G., Reed, B., Huete, A., White, M., Morisette, J., et al. (2006). Land surface phenology. A community white paper requested by NASA. ftp://ftp.iluci.org/Land_ESDR/Phenology_Friedl_whitepaper.pdf Last accessesd May 15, 2010.

Ganguly, S., Friedl, M., Tan, B., Zhang, X., & Verma, M. (2010). Land surface phenology from MODIS: Characterization of the Collection 5 global land cover dynamics product. *Remote Sensing of Environment*, *114*, 1805–1816.

Goward, S. N., Markham, B., Dye, D. G., Dulaney, W., & Yang, J. L. (1991). Normalized difference vegetation index measurements from the Advanced Very High-Resolution Radiometer. *Remote Sensing of Environment*, *35*, 257–277.

Hargrove, W. W., Spruce, J. P., Gasser, G. E., & Hoffman, F. M. (2009). Toward a national early warning system for forest disturbances using remotely sensed canopy phenology. *Photogrammetric Engineering & Remote Sensing*, *75*, 1150–1156.

Heinsch, F., Zhao, M., Running, S., Kimball, J., Nemani, R., Davis, K., et al. (2006). Evaluation of remote sensing based terrestrial productivity from MODIS using regional tower eddy flux network observations. *IEEE Transactions on Geoscience and Remote Sensing*, *44*, 1908–1925.

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, *83*, 195–213.

IPCC (2007). *Climate change 2007: Synthesis report.* Geneva: Intergovernmental Panel on Climate Change.

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

Kassnacht, K., & Gower, S. (1997). Interrelationships among the edaphic and stand characteristics, leaf area index, and aboveground net primary production of upland forest ecosystems in north central Wisconsin. *Canadian Journal of Forest Research*, *27*, 1058–1067.

Kathuroju, N., White, M. A., Symanzik, J., Schwartz, M. D., Powell, J. A., & Nemani, R. R. (2007). On the use of the Advanced Very High Resolution Radiometer for development of prognostic land surface phenology models. *Ecological Modelling*, *201*, 144–156.

Khanduri, V. P., Sharma, C. M., & Singh, S. P. (2008). The effects of climate change on plant phenology. *Environmentalist*, *28*, 143–147.

Liang, L., & Schwartz, M. (2009). Landscape phenology: An integrative approach to seasonal vegetation dynamics. *Landscape Ecology*, *24*, 465–472.

Lieth, H. (Ed.). (1974). *Phenology and seasonal modeling.* New York: Springer-Verlag.

Lloyd, D. (1990). A phenological classification of terrestrial vegetation cover using shortwave vegetation index imagery. *International Journal of Remote Sensing*, *11*, 2269–2279.

Lu, D., Moran, E., & Batistella, M. (2003). Linear mixture model applied to Amazonian vegetation classification. *Remote Sensing of Environment*, *87*, 456–469.

Martin, L. (1965). *Physical geography of Wisconsin.* Madison: University of Wisconsin Press.

Meier, U. (2001). *Growth stages of mono-and dicotyledonous plants: BBCH monograph*, 2nd ed. Berlin/Braunschweig: Federal Research Centre for Agriculture and Forestry.

Menzel, A., Sparks, T. H., Estrella, N., Koch, E., Aasa, A., Ahas, R., et al. (2006). European phenological response to climate change matches the warming pattern. *Global Change Biology*, *12*, 1969–1976.

Morisette, J. T., Privette, J. L., & Justice, C. O. (2002). A framework for the validation of MODIS land products. *Remote Sensing of Environment*, *83*, 77−96.

Morisette, J. T., Richardson, A. D., Knapp, A. K., Fisher, J. I., Graham, E. A., Abatzoglou, J., et al. (2009). Tracking the rhythm of the seasons in the face of global change: Phenological research in the 21st century. *Frontiers in Ecology and the Environment*, *7*, 253−260.

Morellato, L. C. P., Camargo, M. G. G., Neves, F. F. D., Luize, B. G., Mantovani, A., & Hudson, I. L. (2010). The influence of sampling method, sample size, and frequency of observations on plant phenological patterns and interpretation in tropical forest trees. In I. L. Hudson, & M. R. Keatley (Eds.), *Phenological research: Methods for environmental and climate change analysis* (pp. 99−121). Dordrecht: Springer.

Qu, J., Gao, W., Kafatos, M., Murphy, R., & Salomonson, V. (2006). Earth science satellite remote sensing vol. 1: Science and instruments. Beijing: Tsinghua University Press and Berlin/Heidelberg: Springer-Verlag.

Reed, B. C., Brown, J. F., Vanderzee, D., Loveland, T. R., Merchant, J. W., & Ohlen, D. O. (1994). Measuring phenological variability from satellite imagery. *Journal of Vegetation Science*, *5*, 703−714.

Reed, B., Schwartz, M., & Xiao, X. (2009). Remote sensing phenology: Status and the way forward. In A. Noormets (Ed.), *Phenology of ecosystem processes: Applications in global change research* (pp. 231−246). Dordrecht: Springer.

Richardson, A. D., Jenkins, J. P., Braswell, B. H., Hollinger, D. Y., Ollinger, S. V., & Smith, M. L. (2007). Use of digital webcam images to track spring green-up in a deciduous broadleaf forest. *Oecologia*, *152*, 323−334.

Roberts, D., Gardner, M., Church, R., Ustin, S., Scheer, G., & Green, R. (1998). Mapping chaparral in the Santa Monica Mountains using multiple endmember spectral mixture models. *Remote Sensing of Environment*, *65*, 267−279.

Robin, J., Dubayah, R., Sparrow, E., & Levine, E. (2007). Monitoring start of season in Alaska with GLOBE, AVHRR, and MODIS data. *Journal of Geophysical Research*, *113*, G01017.

Schwartz, M. D. (1997). Spring Index models: An approach to connecting satellite and surface phenology. In H. S. Lieth, & M. D. Schwartz (Eds.), *Phenology of seasonal climates I* (pp. 23−38). Netherlands: Backhuys Publishers.

Schwartz, M. D. (Ed.). (2003). *Phenology: An integrative environmental science*. Dordrecht: Kluwer Academic Publishers now Springer.

Schwartz, M. D., Ahas, R., & Aasa, A. (2006). Onset of spring starting earlier across the Northern Hemisphere. *Global Change Biology*, *12*, 343−351.

Schwartz, M. D., & Hanes, J. (2009). Intercomparing multiple measures of the onset of spring in eastern North America. *International Journal of Climatology, Early View*. doi:10.1002/joc

Schwartz, M. D., Liang, L., & Hanes, J. M. (in preparation). High-resolution Spatial and Temporal Phenological Measurements.

Schwartz, M. D., & Reed, B. C. (1999). Surface phenology and satellite sensor-derived onset of greenness: An initial comparison. *International Journal of Remote Sensing*, *20*, 3451−3457.

Schwartz, M. D., Reed, B. C., & White, M. A. (2002). Assessing satellite-derived start-of-season measures in the conterminous USA. *International Journal of Climatology*, *22*, 1793−1805.

Seong, J., Mulcahy, K., & Usery, E. (2002). The sinusoidal projection: A new importance in relation to global image data. *The Professional Geographer*, *54*, 218−225.

Soudani, K., le Maire, G., Dufrene, E., Francois, C., Delpierre, N., Ulrich, E., et al. (2008). Evaluation of the onset of green-up in temperate deciduous broadleaf forests derived from Moderate Resolution Imaging Spectroradiometer (MODIS) data. *Remote Sensing of Environment*, *112*, 2643−2655.

Van Buskirk, J., Mulvihill, R. S., & Leberman, R. C. (2009). Variable shifts in spring and autumn migration phenology in North American songbirds associated with climate change. *Global Change Biology*, *15*, 760−771.

West, N., & Wein, R. (1971). A plant phenological index technique. *Bioscience*, *21*, 116−117.

White, M. A., de Beurs, K., Didan, K., Inouye, D., Richardson, A., Jensen, O., et al. (2009). Intercomparison, interpretation, and assessment of spring phenology in North America estimated from remote sensing for 1982–2006. *Global Change Biology*, *15*, 2335−2359.

White, M. A., Hoffman, F., Hargrove, W. W., & Nemani, R. R. (2005). A global framework for monitoring phenological responses to climate change. *Geophysical Research Letters*, *32*, L04705.

White, M. A., Thornton, P. E., & Running, S. W. (1997). A continental phenology model for monitoring vegetation responses to interannual climatic variability. *Global Biogeochemical Cycles*, *11*, 217−234.

Wu, J. (1999). Hierarchy and scacling: Extrapolating information along a scaling ladder. *Canadian Journal of Remote Sensing*, *25*, 367−380.

Wu, J., Jones, K. B., Li, H., & Loucks, O. L. (2006). *Scaling and uncertainty analysis in ecology: Methods and applications*. Dordrecht: Springer.

Wu, C. S., & Murray, A. T. (2003). Estimating impervious surface distribution by spectral mixture analysis. *Remote Sensing of Environment*, *84*, 493−505.

Zhang, X. Y., Friedl, M. A., & Schaaf, C. B. (2009). Sensitivity of vegetation phenology detection to the temporal resolution of satellite data. *International Journal of Remote Sensing*, *30*, 2061−2074.

Zhang, X. Y., Friedl, M. A., Schaaf, C. B., Strahler, A. H., Hodges, J. C. F., Gao, F., et al. (2003). Monitoring vegetation phenology using MODIS. *Remote Sensing of Environment*, *84*, 471−475.

Zhang, X. Y., Hodges, J. C. F., Schaaf, C. B., Friedl, M. A., Strahler, A. H., & Gao, F. (2001). Global vegetation phenology from AVHRR and MODIS data. *Proceedings of the International Geoscience and Remote Sensing Symposium*, *5*, 2262−2264.

Zhang, X. Y., Tarpley, D., & Sullivan, J. T. (2007). Diverse responses of vegetation phenology to a warming climate. *Geophysical Research Letters*, *34*, L19405.