A regional phenology model for detecting onset of greenness in temperate mixed forests, Korea: an application of MODIS leaf area index

Sinkyu Kang\textsuperscript{a,*}, Steven W. Running\textsuperscript{a}, Jong-Hwan Lim\textsuperscript{b}, Maosheng Zhao\textsuperscript{a}, Chan-Ryul Park\textsuperscript{b}, Rachel Loehman\textsuperscript{a}

\textsuperscript{a}Numerical Terradynamic Simulation Group, School of Forestry, University of Montana, Missoula, MT 59812, USA
\textsuperscript{b}Korea Forest Research Institute, Seoul, Republic of Korea

Received 3 September 2002; received in revised form 2 March 2003; accepted 16 March 2003

Abstract

A regional phenology model for detecting onset of vegetation greenness was developed using year 2001 MODIS land products in temperate mixed forests in Korea. The model incorporates a digital elevation model (DEM), moderate resolution imaging spectroradiometer (MODIS) landcover and leaf area index (LAI) products, and climate data from weather-monitoring stations. MODIS-based onset of greenness varied spatially and showed significant correlation with air temperature ($r^2 = 0.70$, $p < 0.001$). Our modeling methodology is to relate thermal summation calculated using the MODIS-based timing of onset with 30-year mean air temperature. Onset of greenness is expected to occur at values above the critical thermal summation threshold and is predicted to vary spatially. An algorithm for downscaling 8-day composite MODIS LAI product to a daily unit was introduced, and its predictability was validated using ground-measured onset of greenness. Two unknown parameters and the best regression were determined by iterative cross-validation. Minimal cross-validation errors between the predicted and MODIS-based timings of onset were found at a mean absolute error (MAE = 3.0 days) and bias (+ 1.6 days). The predicted onsets show good agreement with ground-measured onset of greenness (MAE = 2.5 days and bias = + 2.5 days). This study demonstrates the utility of MODIS land products as tools for detecting spatial variability in phenology across climate gradients.

Keywords: Onset of greenness; Phenology; MODIS; Mixed forest; Climate; Cross-validation

1. Introduction

In recent years, vegetation phenology has found new relevance within global climate change research, especially with increasing popularity and availability of satellite data (Myneni et al., 2002; Reed et al., 1994; White, Thornton, & Running, 1997) and in the development of regional phenology networks (Chuine, Belmonte, & Mignot, 2000; Kramer, 1994; Menzel & Fabian, 1999; Schwartz, 1998). Timing of onset and offset of vegetation greenness is of particular importance because of its considerable influence on the spatial and interannual variability of terrestrial carbon cycles (Kimball, Keyser, Running, & Saatchi, 1999; White, Running, & Thornton, 1999). Together with meteorological data, both satellite-based (White, Nemani, Thornton, & Running, 2002; White et al., 1997) and surface-based (Chuine, Belmonte, et al., 2000) phenology data have been used to develop several phenology models for predicting onset and/or offset of vegetation greenness. This study presents a phenology model for predicting regional variability of onset of greenness in temperate mixed forests using newly available satellite images from the moderate resolution imaging spectroradiometer (MODIS) sensor.

Following White et al. (2002), “onset of greenness” is the term for an aggregate phenological event, in contrast to “budburst” or “leafout”, which indicates a particular species-specific event. Nevertheless, modeling the aggregate phenological event follows principles of species-level leaf physiology because canopy is abstracted to a single large leaf in “big leaf” ecosystem models (e.g. BIOME-BGC; Running & Coughlan, 1988). In this context, both the timing of onset of greenness and budburst has been success-
fully modeled using cumulative thermal summation either with or without consideration of chilling requirements (Campbell & Norman, 1998; Cannell & Smith, 1983; Chuine, 2000; Chuine & Cour, 1999; Chuine, Cour, & Rousseau, 1999; Hanninen, 1994; Hanninen, Hakkinen, Hari, & Koski, 1990; Kramer, 1994; Murray, Cannel, & Smith, 1989; Valentine, 1983).

Budburst occurs at the termination of bud dormancy, a physiological state which is defined as a period of reduced growth rate with new or no cell divisions (Chuine & Cour, 1999). Bud dormancy has two different phases: rest and quiescence. Rest is the dormant period caused by growth arresting physiological conditions in the bud itself, and its phase length is determined by the summation of a rate of chilling with optimal and threshold temperatures (Hanninen & Backman, 1994; Kramer, 1994). Quiescence is defined as the period during which the buds remain dormant due to unfavorable environmental conditions and is broken when buds are exposed to forcing temperatures for a prolonged period (Kramer, 1994).

Historically, temperature dependence of onset of greenness or budburst has been modeled as following two methods. One-step onset models consider only a single-transition phase from bud dormancy to budburst (Cannell & Smith, 1983; Chuine et al., 1999; Hunter & Lechowicz, 1992; Valentine, 1983). Alternatively, two-step onset models identify two different phases during dormancy (rest and quiescence phases) and simulate two different transitions: winter rest to quiescence, and quiescence to budburst, catalyzed by the state of chilling and the state of forcing, respectively (Chuine, 2000; Hanninen, 1994; Kramer, 1994; Murray et al., 1989). Two-step models emphasize the role of chilling temperature in determining the timing of onset (Chuine, 2000; Chuine & Cour, 1999). Although there is some evidence that the degree of the state of chilling for the timing of onset is proportional to the preceding summer temperature (Balduman, Aitken, Harmon, & Adams, 1999), Chuine and Cour (1999) showed that performance of phenological models might not be significantly improved by considering this relationship.

Large-scale carbon cycling models require a regional-scale phenology model to capture mean phenological change related to average environmental conditions for each grid cell. Although there is a successful pilot study to scale phenology from local to regional level using species-specific phenological models based on data from regional phenology networks (Chuine, Cambon, & Comtois, 2000), satellite data provide an efficient means for monitoring regional or global-scale phenology (Chen & Cihlar, 1996; Moulin, Kergoat, Viyyo, & Dedieu, 1997; Myené et al., 2002; Myené, Kelling, Tucker, Asrar, & Nemenyi, 1997; Myené, Nemenyi, & Running, 1997; Running et al., 1994; Sellers et al., 1994; Settle & Drake, 1993; Spanner, Pierce, Running, & Peterson, 1990; Turner, Cohen, Kennedy, Fassnacht, & Briggs, 1999). Compared to species-specific phenology models, satellite data is especially useful to detect mean phenological change within each grid pixel, where diverse vegetation types are aggregated (White et al., 2002). Notably, satellite reflectance from NOAA-AVHRR has been used to estimate the timing of onset or offset of greenness because of its high temporal frequency, which can capture day-to-day variations in canopy status (Goward, Tucker, & Dye, 1985; Idieke, Ramge, & Kohlmaier, 1996; Reed et al., 1994; White et al., 1997).

In contrast to its predecessors (e.g. AVHRR), MODIS incorporates enhanced atmospheric correction, cloud detection, improved georeferencing, and the enhanced ability to monitor vegetation (Running et al., 1999, 1994). MODIS leaf area index (LAI) monitors seasonal variation in LAI at 1-km nadir resolution every 8 days (Knyazikhin, Martonchik, Myen, Diner, & Running, 1998) and provides ancillary information on landcover and cloud contamination of each pixel. These features promise greater utility of MODIS LAI for development of onset phenology models over its predecessors, which require extra image processing for masking clouds and classifying landcover (e.g. AVHRR; White et al., 2002). Moreover, if it is properly validated (Cohen & Justice, 1999; Myené et al., 2002), MODIS LAI provides a more physically meaningful threshold for defining onset of greenness (e.g. LAI>0) than NDVI, which uses a somewhat arbitrary threshold value from seasonal or cumulated NDVI curves (Chen, Tan, Schwartz, & Xu, 2000; Reed et al., 1994; White et al., 1997).

In the current study, we applied MODIS LAI to develop a simple one-step phenology model for predicting regional variability of onset of vegetation greenness in temperate mixed forests in Korea. The MODIS-based onset of greenness was detected and related with surface meteorological data to model critical threshold for onset of greenness, above which onset of greenness is predicted to occur. Empirical evidence indicates the threshold is spatially variable depending on local climate (Sorensen, 1983). Similarly, White et al. (1997) showed that the critical threshold for deciduous species is spatially variable depending on climatic gradients of soil temperature and daylight short-wave radiation within the conterminous US. In the current study, this critical threshold was assumed to be a function of regional gradients of specific climate variables. We also hypothesized that the control climate variable(s) can be detected, prior to model validation, through correlation with MODIS-based onset of greenness. As well, we introduce an algorithm for downscaling 8-day composite MODIS LAI to a daily unit, for enhanced temporal predictability, using seasonal trajectory patterns of MODIS LAI. Our regional phenology model and downscaling algorithm are validated using ground-measured onset of greenness data.

2. Methodology

The one-step phenology model relates meteorological variables with the MODIS-driven timing of onset and
predicts spatial and temporal variability in the onset at a regional scale. MODIS LAI was used to derive the timing of onset and was correlated with meteorological data from a regional network of ground weather-monitoring stations in Korea (Fig. 1c). Unknown model parameters were determined through a process of cross-validation. Normalization of MODIS LAI was applied to eliminate seasonal effects of evergreen vegetation and deciduous understory species within mixed forest pixels (Eq. (1)). Temporal downscaling from 8-day to daily intervals was implemented (Eq. (2)); its predictability was validated using ground measurements.

2.1. MODIS data products

The MODIS LAI 8-day product and landcover are available from the Earth Observing System data gateway (EOS, 2003) at 1-km nadir resolution. The operational MODIS LAI algorithm ingests up to seven atmosphere-corrected surface spectral bidirectional reflectance factors and outputs the most probable values for pixel LAI (Myneni et al., 2002). An 8-day composite method is applied for MODIS LAI to eliminate cloud contamination, which means that a given scene represents the maximum LAI for a consecutive 8-day period. For detailed information on the MODIS LAI product, refer to: Knyazikhin et al. (1998) for the theoretical basis of the algorithm, Knyazikhin et al. (1999) for the implementation aspects, and Myneni et al. (2002) for the recent validation results.

In the current study, MODIS LAI 8-day composite scenes from 3/4/2001 to 11/30/2001 were downloaded, reprojected into a geographic projection using reprojtool, a reprojection software (J. Glassy and P. Thornton, NTSG, University of Montana), and then merged for further analysis using ArcInfo (ArcInfo v.7.2, ESRI) (Fig. 1a). MOD15A2 quality control (QC) flags were used to detect cloud contamination in each pixel (Fig. 1b). Only cloud-free pixels, identified as those with QC flags of 0 and 4 (Myneni et al., 2002), were used to develop the phenology model; these designations indicate the use of the empirical backup method and radiative transfer method to calculate FPAR and LAI, respectively. Landcover classification was made using MODIS UMD (University of Maryland) landcover from 1/1/2001. In Fig. 1c, The UMD landcover (18 classes) was simplified to five classes for graphic enhancement, which shows mixed forest as the dominant landcover type in Korean montane forests.

2.2. Study area and site selection

Our study region, Korea (Republic of), is located in Northeast Asia. The total land area of 98,500 km² and approximate latitudinal and longitudinal ranges are N33°–N37° and E125°–E130°, respectively (Fig. 1). Forest covers approximately 65% of the country area with equivalent composition of deciduous and evergreen needleleaf species (Korea Forest Service, 1997). Most forests are distributed in rugged mountain areas because of intensive landuse elsewhere. Pine species are dominant on south-facing slopes, but oak trees dominate north-facing slopes. Because the projected length of slope reach is generally less than 1 km (e.g. 200 m in eastern montane areas; Kang, 2001), we expect that most Korean forests will be classified as mixed forest at the 1-km resolution of MODIS landcover (Fig. 1c). With respect to seasonal LAI, understory vegetation is likely to have an influence on early spring, late fall, and winter LAI. In general, evergreen understory (e.g. Sasa borealis) is broadly distributed along the southern coasts, and deciduous understory is dominant within inland forests.

Fig. 1. Geographic location of selected study sites and sample MODIS products: (a) MOD15A2 8-day composite LAI (m² m⁻²) and (b) Quality control flags in late April (4/23/2001–4/30/2001); (c) Land cover (MOD12) and the sites. In (c), dots indicate model development sites and stars denote validation sites with ground measurement of onset of greenness.
It has been observed that understory vegetation composition is aspect-dependent (Lee, Yoo, Oh, Shim, & Kang, 1999) and that the species reach their maximum LAI at 1.3 m$^2$/C0 under deciduous canopy (Kang, 2001). Within a given 1-km pixel, therefore, the classification of mixed forest represents a combination of deciduous and evergreen overstory and deciduous and/or evergreen understory. We applied the following filtering rules for site selection: (1) the site is located within a 10-km radius of a Korean national weather station (NWS); (2) the nearest NWS has 30-year climatological records including air temperature, precipitation, and daylength; and (3) the site is a mixed forest, at least, with a minimum area of 25 km$^2$. A total of 35 mixed forest sites were selected for the study (Fig. 1a). Mean LAI was calculated using LAI values from cloud-free pixels (QC flag: 0 or 4) for 5 × 5-pixel (25 km$^2$) windows from around each of the 35 study sites.

### 2.3. Satellite detection of onset of greenness

An example of seasonal MODIS LAI (Fig. 2a) shows considerably smaller midsummer LAI values than expected. To correct this probable error, we made a modification to seasonal LAI by ignoring decreased LAI during leaf expansion periods: if the LAI for a particular 8-day period was less than the LAI for the preceding period, that LAI was assigned the value from the preceding period (Fig. 2b). For deciduous species, an LAI>0 theoretically means that budburst has already occurred. Our study sites, however, include evergreen species that return a non-zero minimum LAI during winter, and minimum LAI is likely dependent on both evergreen overstory and/or understory species within a mixed forest pixel. Seasonal LAI was normalized to eliminate the effect of evergreen species on the minimum LAI and to establish an objective rule for detecting the timing of onset where the maximum values of seasonal LAI differ by site (White et al., 1997). The minimum and maximum LAI of the $i^{th}$ site ($\text{LAIM}_{\text{min},i}$ and $\text{LAIM}_{\text{max},i}$) in 2001 were used for the normalization (Eq. (1)).

We assume that onset of greenness is expected to occur when the normalized LAI increases over an onset threshold ($\text{LAI}_{\text{on},i}$). We expect (1) that the normalization is not sufficient to eliminate the effect of deciduous understory species; and (2) that these species affect surface reflectance very early in the growing season before growth of overstory canopy occurs. This signal increases MODIS LAI slightly, and may cause early detection of onset of greenness for tree species within a mixed forest pixel. In addition, the MODIS compositing method, which adopts maximum LAI for an 8-day period, may also underestimate the timing of onset. To minimize these potential sources for error we used a positive onset threshold ($\text{LAI}_{\text{on},i}$), in which three different onset thresholds (0.1, 0.2, and 0.3) were applied, and the best threshold was determined by cross-validation (Table 1).

\[
\text{LAI}_{\text{on},i} = \frac{\text{LAIM}_{\text{max},i} - \text{LAIM}_{\text{min},i}}{\text{LAIM}_{\text{max},i} - \text{LAIM}_{\text{min},i}}
\]

\[
\text{OnsetS}_i = \text{Onset8}_i + 7 \times \left( 1 - \frac{\Delta \text{LAI}_{\text{on},i} - \Delta \text{LAI}_{\text{on},\text{min}}}{\Delta \text{LAI}_{\text{on},\text{max}} - \Delta \text{LAI}_{\text{on},\text{min}}} \right)
\]

The onset week (8-day MODIS LAI period) was identified when the normalized LAI exceeded the onset threshold. The date of onset of greenness was then determined using two alternative timings of onset (Eq. (2)). For a simple case, the beginning of the detected onset week was chosen as the timing of onset (Onset8). A scaled timing of onset was also chosen from the onset week, depending on the rate of change of the normalized LAI. We hypothesized that onset starts earlier in the onset week, as the onset-week LAI$_{on,i}$ is larger than the LAI$_{on,i}$ from the previous week. The scaled onset of the $i^{th}$ site (designated OnsetS$_i$) is defined by a linear interpolation using the Onset8$_i$ and an increase in LAI$_{on,i}$ ($\Delta \text{LAI}_{\text{on},i}$) in the onset week and the maximum and the minimum increase of the LAI$_{on,i}$ ($\Delta \text{LAI}_{\text{on},\text{max},i}$ and $\Delta \text{LAI}_{\text{on},\text{min},i}$) during the leaf expansion period. If $\Delta \text{LAI}_{\text{on},i}$
equals $\Delta LAI_{\text{in}, \text{max}}$, the OnsetS$_i$ is the beginning of the onset week (Onset8$_i$); if $\Delta LAI_{\text{in}, i} = \Delta LAI_{\text{in}, \text{min}}$, the OnsetS$_i$ is the end of the onset week (Onset8$_i$ + 7). Terminology and definitions of variables used in the onset model are listed in Table 1.

2.4. Meteorological and geographic data

Daily meteorological data for 2001 and 30-year mean (1971–2000) data were downloaded from the Korean Meteorological Administration (http://www.kma.go.kr/) for the selected 35 national weather stations. The data include daily average air temperature, precipitation, and daylight hours. For all but three of these sites, we orographically corrected the NWS meteorological data. Air temperature was corrected for orographic effects using a temperature lapse rate of $-5^\circ\text{C/km}$. No orographic correction was, however, made on precipitation and daylength. A digital elevation model (DEM) of this study region was downloaded from the HYDRO 1-km Asia DEM at EROS Data Center (EROS, 2003) and then resampled to a 5-km DEM. Weather station and site elevations were determined from 1- and 5-km DEMs, respectively, and elevation differences between NWS and study sites were used for the orographic corrections.

2.5. Model development

Our approach to regional phenological modeling is a simple one-step model relating meteorological variables to the MODIS-driven timing of onset. We therefore ignore chilling requirements for the timing of onset. Selection of the meteorological variables was accomplished by examining the Pearson correlation coefficient between 30-year climate variables and the MODIS-based timing of onset in 2001. Air temperature ($T_{30}$) was the only meteorological variable significantly correlated with the timing of onset ($0.61–0.70$, $p < 0.01$) and was therefore retained in the model (Table 2).

Onset of greenness occurs when the state of forcing ($F$) is greater than the critical state of forcing ($F^*$) (Eq. (3)) (Chuine, 2000). The state of forcing ($F$) is defined as the summation of mean air temperature ($T_c$) over a base temperature ($T_a$) from the first day of the year (DOY) (Eq. (4)). The critical state of forcing ($F^*$) is calculated using a linear regression model of 30-year mean air temperature ($T_{30}$) (Eq. (5)). Two unknown coefficients in the regression model ($\alpha$ and $\beta$) were determined by relating $T_{30}$ with $F^*$, based on the timing of onset detected from satellite data (Onset_Modis) for 2001 (Eq. (6)). This parameterization process will be described in detail in the following section. Using the regression model, we can estimate the spatial variability of $F^*$ using gradients of $T_{30}$ at a regional scale; therefore, we can predict the timing of onset using daily mean air temperature for a given year.

$$F_i \geq F_i^*$$  \hspace{1cm} (3)

$$F_i = \sum_{\text{Onset}_i} \frac{T_{a,i}}{\text{DOY}=1} (T_{a,i} \geq T_c)$$  \hspace{1cm} (4)

$$F_i^* = \alpha T_{30,i} + \beta$$  \hspace{1cm} (5)

$$F_i^* = \sum_{\text{Onset}_\text{Modis}} \frac{T_{a,i}}{\text{DOY}=1} (T_{a,i} \geq T_c)$$  \hspace{1cm} (6)
Table 2
Summary of MODIS-based onset of greenness (day of year) and correlations with environmental variables

<table>
<thead>
<tr>
<th></th>
<th>Onset8</th>
<th>OnsetS</th>
<th>Onset8</th>
<th>OnsetS</th>
<th>Onset8</th>
<th>OnsetS</th>
</tr>
</thead>
<tbody>
<tr>
<td>LAIn,C 0.1</td>
<td>107</td>
<td>113</td>
<td>113</td>
<td>117</td>
<td>115</td>
<td>117</td>
</tr>
<tr>
<td>LAIn,C 0.2</td>
<td>24</td>
<td>25</td>
<td>24</td>
<td>23</td>
<td>24</td>
<td>23</td>
</tr>
<tr>
<td>LAIn,C 0.3</td>
<td>119</td>
<td>126</td>
<td>127</td>
<td>131</td>
<td>127</td>
<td>131</td>
</tr>
<tr>
<td>LAIn,C 0.5</td>
<td>95</td>
<td>102</td>
<td>103</td>
<td>108</td>
<td>103</td>
<td>108</td>
</tr>
</tbody>
</table>

Correlation

<table>
<thead>
<tr>
<th></th>
<th>Onset8</th>
<th>OnsetS</th>
<th>Onset8</th>
<th>OnsetS</th>
<th>Onset8</th>
<th>OnsetS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Latitude</td>
<td>0.54 (0.001)**</td>
<td>0.60 (0.000)**</td>
<td>0.50 (0.002)**</td>
<td>0.49 (0.003)**</td>
<td>0.51 (0.002)**</td>
<td>0.46 (0.006)**</td>
</tr>
<tr>
<td>DEM</td>
<td>0.23 (0.19)</td>
<td>0.29 (0.09)</td>
<td>0.41 (0.02)*</td>
<td>0.41 (0.01)*</td>
<td>0.23 (0.19)</td>
<td>0.38 (0.02)*</td>
</tr>
<tr>
<td>T30</td>
<td>-0.61 (0.000)**</td>
<td>-0.70 (0.000)**</td>
<td>-0.70 (0.000)**</td>
<td>-0.69 (0.000)**</td>
<td>-0.61 (0.000)**</td>
<td>-0.66 (0.000)**</td>
</tr>
<tr>
<td>P30</td>
<td>-0.30 (0.08)</td>
<td>-0.25 (0.15)</td>
<td>0.11 (0.53)</td>
<td>0.14 (0.43)</td>
<td>-0.08 (0.64)</td>
<td>0.02 (0.89)</td>
</tr>
<tr>
<td>RH30</td>
<td>-0.07 (0.70)</td>
<td>-0.09 (0.63)</td>
<td>-0.08 (0.66)</td>
<td>-0.91 (0.60)</td>
<td>0.04 (0.83)</td>
<td>-0.11 (0.53)</td>
</tr>
</tbody>
</table>

Data in parentheses are significance level of correlation.

DEM = Digital elevation model (m); T30 = 30-year mean air temperature (°C); P30 = 30-year mean annual precipitation (mm year⁻¹); RH30 = 30-year mean annual daylight hours (hours).

* Correlation is significant at the 0.05 level (two-tailed).
** Correlation is significant at the 0.01 level (two-tailed).

2.6. Model parameterization

Two unknown parameters, a critical threshold (LAIn,C) and a base temperature (Tc), were determined to complete model development. Although in previous studies, Tc has been arbitrarily defined as 0 or 5 °C (Kramer, 1994; White et al., 1997), we determined Tc and the critical threshold for onset detection (LAIn,C) through a cross-validation process. Using this method, one observation is withheld, the models are developed from the remaining observations, and a prediction is generated for the withheld observation. Error statistics are calculated from the cross-validation runs; a best regression model with minimal bias and mean absolute error (MAE) is determined by averaging regression coefficients generated by the cross-validation (Deutsch & Journel, 1992; Kang, Kim, & Lee, 2002; White et al., 1997). We used the iterative cross-validation method to determine both the parameters and coefficients of the regression model. In the iterative cross-validation, LAIn,C varied from 0.1 to 0.3 and Tc varied from −5 to 10 °C.

2.7. Model validation

The onset model was validated using two sets of ground measurements for onset of greenness. Length of bud and subsequent leaf development were measured at two long-term ecological study sites. The northern Gyebang site (data period: 1996–2001) is cooler (T30 = 7.1 °C) than the southern Backwoon site (data period: 2000; 11.3 °C). The onset was observed at several plots across different slopes and elevations at each site. Although these ground measurements are limited compared to other studies (Chuine, Campbell, et al., 2000; Chuine, Cambon, et al., 2000; Menzel & Fabian, 1999; White et al., 1997), we consider these data sufficient for validation of the regional phenology model because the data were obtained across climatic gradients and over plot-scale observations. Because our phenology model estimates mean onset of diverse vegetation types within a grid pixel, but not the onset of specific species, we used a weighted average timing of onset by species dominance as the representative timing of onset at each site.

3. Results

3.1. Detection of the timing of onset

The MODIS-based onsets are spatially variable, ranging from 23 to 25 days depending on the applied onset threshold (LAIn,C) and the alternative detection methods of Onset8i and OnsetSi. The timings of onset are significantly different (p = 0.02, n = 35) among the cases in Table 2. OnsetSi occurs significantly later (111.6) than Onset8i (111.8), but the range of distribution is similar in the both cases (Table 2). Later onset dates occurred with increasing latitude (p = 0.006, n = 35) and decreasing 30-year mean air temperature (T30; p < 0.0001, n = 35). Annual precipitation (P30) and radiation hours (RH30) did not show significant correlation with timings of onset. The correlation analysis indicates that annual average air temperature acts as a forcing variable in determining the spatial variability in the timing of onset in our study region, as found in several previous studies (Campbell & Norman, 1998; Chuine & Cour, 1999; Valentine, 1983; White et al., 1997).

3.2. Model development and parameterization

Among the climatological variables, only T30 was used for developing the regression model (Eq. (5)). The site-specific critical state of forcing (F*) was calculated (Eq. (6)) using the MODIS-based timing of onset and daily mean air temperature for 2001. OnsetSi showed significantly higher F* than Onset8i. For example, when LAIn,C (0.1) and Tc (0
°C) are applied, the mean $F^*$ of OnsetS$_i$ (471 °C, degree days) is significantly higher ($p < 0.0001$) than that of Onset8$_i$ (397 °C). $T_{30}$ was positively related with and explained 58% and 41% variations in $F^*$ of OnsetS$_i$ and Onset8$_i$, respectively (Fig. 3). These results indicate that our phenology model captures a known trend: warmer sites require higher critical states of forcing ($F^*$) than cooler sites.

Iterated cross-validation was implemented to determine two unknown parameters, LAIn$_{n,c}$ and $T_c$, and two coefficients of the regression model (Eq. (5)). MAE and bias were produced for error statistics for the cross-validation. The minimal MAE (3.0 days) was found with LAIn$_{n,c}$ (0.1) and $T_c$ (0 °C), and the minimal bias ($+1.4$ days) was with LAIn$_{n,c}$ (0.3) and $T_c$ ($-4$ °C), respectively (Fig. 4 and Table 3). Because bias was not sensitive to the applied LAIn$_{n,c}$ and $T_c$, we used MAE as a major error indicator and, therefore, parameterized LAIn$_{n,c}$ and $T_c$ as 0.1 and 0 °C, respectively. Finally, coefficients of the regression model were obtained by averaging the coefficients from the iterated cross-validation runs using LAIn$_{n,c}$ (0.1) and $T_c$ (0 °C) (Eqs. (7) and (8)).

$$F^*_i = 31.5 T_{30,i} + 48.2 \quad \text{for Onset8}_i \quad (7)$$

$$F^*_i = 34.5 T_{30,i} + 89.2 \quad \text{for OnsetS}_i \quad (8)$$

The critical state of forcing ($F^*_i$) at the $i$th site is estimated with the above equations, using $T_{30,i}$. The timing of onset is predicted using Eqs. (3) and (4). When the predicted timing of onset was compared to the MODIS-based onset, OnsetS$_i$ fitted better (MAE = 3.0 days) than Onset8$_i$ (MAE = 4.0 days) and occurred significantly later ($p < 0.0001$; 118 DOY) than Onset8$_i$ (114 DOY).

### 3.3. Validation

The predicted timings of onset generally corresponded well with independent field observations for all years except 1998. An extraordinary warm winter during late 1997 and early 1998 resulted in earlier onset than in other years, but our model could not capture the effect of the anomalous winter temperatures on the timing of subsequent onset events. The OnsetS$_i$ of the two validation sites was better fitted to the field observation with MAE (2.5 days) and bias ($2.5, +2.5$ days) than Onset8$_i$ (4.1, and $-4.1$ days), respectively (Fig. 6a and b). Annual variability in observed timings of onset at the Gyebang site ranged between OnsetS$_i$ and Onset8$_i$ values for the site for all years except 1998 (Fig. 6c). Although the Buckwoon site, which has a more diverse species composition than Gyebang, showed large interplot variability in observed timings of onset, the mean timing of

### Table 3

<table>
<thead>
<tr>
<th>LAIn$_{n,c}$</th>
<th>Onset8</th>
<th>OnsetS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Onset8</td>
<td>OnsetS</td>
</tr>
<tr>
<td>0.1</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>0.2</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>0.3</td>
<td>3.5</td>
<td>3.8</td>
</tr>
</tbody>
</table>

LAIn$_{n,c}$ and $T_c$ are critical onset thresholds and base temperature, respectively.
onset occurred between the OnsetS and Onset8 predictions for the site.

4. Discussion and conclusion

This study demonstrates the utility of MODIS land products as powerful tools in the construction of regional phenology models for temperate mixed forests. We consider our modeling process as one of a number of diverse validation efforts for the MODIS LAI product (Cohen & Justice, 1999; Gower, Kucharik, & Norman, 1999; Myneni et al., 2002; Reich, Turner, & Bolstad, 1999; Running et al., 1999; Turner et al., 1999). This study suggests an objective method for detecting spatial variability in the timing of onset of greenness across climate gradients using MODIS LAI. We found that spatial variability in onset was successfully explained by temperature gradients. Among many other climate variables, temperature could be identified, prior to model validation, as a controlling variable by regional correlation analysis using MODIS-based onsets. Iterative cross-validation runs applied by White et al. (1997) were extended to simultaneously determine two unknown parameters and two model coefficients. Although the model developed in this study is specific to our study region (Korea) and forest type (mixed temperate), we believe that our approach can be easily applied to other regions where meteorological data are available.

In spite of the relatively simple structure of our one-step, temperature-based phenology model, model predictions were very close matched with field observations (MAE = 2.5 days and bias = +2.5 days). The scaled onset (OnsetS) was significantly later than \( p < 0.0001 \), but better fitted to the field observations than the Onset8 dates, demonstrating that the 8-day MOD15A2 LAI composite product can be successfully partitioned into daily temporal units using the downscaling process suggested in this study (Eq. (2)). Although use of the MODIS daily surface reflectance could provide the enhanced temporal detection of the MODIS-based onsets at a daily time step, this advantage is constrained by daily cloud contamination and increased computing loads. Moreover, surface reflectance or its derivative (e.g., NDVI) does not always correspond to an identical LAI value, but shows statistical distribution depending on canopy physiological status (Myneni et al., 2002). We therefore consider that the temporal downscaling method compensates for the coarse temporal units of MODIS LAI and that the use of MODIS LAI makes data processing more straightforward than using daily surface reflectance.

One of the critical problems in using satellite data to detect onset of greenness is identifying the onset threshold of the satellite-based LAI or NDVI (Chen et al., 2000; Reed et al., 1994; White et al., 1997). In this study, three onset threshold levels \( \text{LAI}_{n,C} = 0.1, 0.2, \) and 0.3 were applied and selected although cross-validation runs. The 0.2 and 0.3 thresholds were less predictive than the 0.1 threshold because higher thresholds fall within the range of already high increases in the rate of the normalized LAI (Fig. 2b) and, therefore, overestimate the onset. Alternatively, the smaller threshold of 0.1 is vulnerable to potential contamination by clouds (Cihlar, Du, & Latifovic, 2001; Cihlar, Latifovic, Chen, & Li, 1999), landcover misclassification or misregistration (Cihlar, 1996; Myneni et al., 2002), and deciduous understory vegetation, resulting in slight increases in LAI in early spring. We solved this problem by using only highest-quality pixels with QC flags of 0 and 4 by averaging 5 × 5 pixels and by constraining model development sites to areas with mixed forest landcover. We recognize, however, that this approach is sometimes insufficient to eliminate cloud contamination, as found in the unreasonably low summer LAI in Fig. 2.

The cross-validation error was higher at warm sites than cool sites in both cross-validations and in comparison with field observations (Figs. 5 and 6). The warm sites, distributed in southern Korea, are characterized by extreme variability in precipitation. Although annual precipitation was...
not significantly correlated with the MODIS-based onset at a regional scale, precipitation might be an influential factor at a subregional scale, as suggested by Zhao, Fu, Yan, and Wen (2001). Zhao et al. found positive correlations between temperature and NDVI and negative correlations between precipitation and NDVI across China during the spring season. They observed a reversal in these relationships in northern China, demonstrating that onset is more dependent upon water availability than air temperature in some regions. In addition, we suggest that the cross-validation error may be related to our basic assumption: the presence of a linear relationship between the critical state of forcing ($F^*$) and air temperature ($T_{30}$) because the mean slope between $F^*$ and $T_{30}$ becomes almost flat over 12 °C (Fig. 3). This result is interesting because southern Korea is known as a transition boundary where the broadleaf species within the mixed forest transits from northern deciduous to southern evergreen broadleaf species, which restricts the application of a regional phenology model across different functional vegetation types. We suggest, however, that the study area should be large or topographically variable enough to have sufficient ranges of the climatic variability and MODIS-based onsets, which promises a presence of a significant linear relationship between the critical state of forcing ($F^*$) and air temperature ($T_{30}$).

This study was an initial step in constructing and validating a regional phenology model for Korean mixed forests, which will be used as a phenology module in a satellite-based model of forest carbon cycling (Kang, 2001). Together with the estimated surface meteorological and climatological data (Daly, Neilson, & Phillips, 1994; Kang et al., 2002; Thornton, Running, & White, 1997), this phenology model will be used to predict retrospective temporal and spatial variability of the onset of greenness in Korean temperate forests for better understanding of the interactions between climatic variability and the terrestrial carbon cycle (White et al., 1999).

Acknowledgements

The authors wish to thank Ramakrishna Nemani and Alana Oakins for their constructive comments on our early draft of this manuscript. We thank Youngee Cho for her kind help with our editing process. This research was funded by NASA grant NAS5-31368. Sinkyu Kang was partly supported by a postdoctoral fellowship program from Korea Science and Engineering Foundation (KOSEF). Jong-Hwan Lim was supported by the “Forest biodiversity and ecosystem change” project from Korea Forest Research Institute.

References


Korea Forest Service (1997). Statistical yearbook of forestry. Seoul, Repub-
lic of Korea (In Korean).


Reich, P., Turner, D., & Bolstad, P. (1999). An approach to spatially-dis-


Running, S. W., Justice, C. O., Salomonson, V., Hall, D., Barker, J., Kauf-


