IMPROVING FOREST GROWTH ESTIMATES USING A BAYESIAN NETWORK APPROACH

Y. T. Mustafa, PhD researcher
A. Stein, Professor
V. Tolpekin, Assistant Professor
Department of Earth Observation Science
Faculty of Geo-Information Science and Earth
Observation of the University of Twente (ITC)
Enschede 7500 AE, The Netherlands
mustafa@itc.nl; stein@itc.nl; tolpekin@itc.nl

ABSTRACT

Estimating the contribution of forests to carbon sequestration is commonly done by applying forest growth models. Such models inherently use field observations, such as leaf area index (LAI), whereas relevant information is also available from remotely sensed images. The purpose of this study is to improve the LAI estimated from the physiological principles predicting growth (3-PG) model by combining its output with LAI derived from Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) satellite imagery. A Bayesian network (BN) approach is proposed to take care of the different structure of the inaccuracies in the two data sources. It addresses the bias in the 3-PG model and the noise of the ASTER images. Moreover, the EM algorithm is introduced into BN to estimate missing the LAI ASTER data, since they are not available for long time series due to the atmospheric conditions. This paper shows that the outputs obtained with the BN were more accurate than the 3-PG estimate, as the root mean square error reduces to 0.46, and the relative error to 5.86%. We conclude that the EM-algorithm within a BN can adequately handle missing LAI ASTER values, and BNs can improve the estimation of LAI values. Ultimately, this method may be used as a predicting model of LAI values, and handling the missing data of ASTER images time series.

KEYWORDS: Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER), Bayesian Networks (BNs), EM-algorithm, leaf Area Index (LAI), physiological principles predicting growth (3-PG).

INTRODUCTION

Forests play a critical role in carbon sequestration (Wamelink et al., 2009), thus affecting the speed of climate change. Therefore, monitoring forest growth has received increasing attention (Bonan, 1993). An important parameter in observing forest growth is the leaf area index (LAI), defined as the ratio of leaf area to per unit ground surface area. It is a key biophysical variable influencing land surface photosynthesis, respiration, transpiration, leaf litterfall and energy balance (Bonan, 1993) . The LAI is estimated using process-based models, such as the Physiological Principles in Predicting Growth (3-PG) model, being a stand-level model of forest growth. This model developed by Landsberg and Waring (1997) and has been used as a point mode. A grid mode version of 3-PG model has been developed by Coops et al. (1998). Similarly, remote sensing imagery has been added to estimate LAI, for example from the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) satellite (Heiskanen, 2006; Ito et al., 2007).

Statistical methods have been used to estimate LAI, in particular Bayesian Networks (BNs) (Kalacska et al., 2005; Mustafa et al., 2011). A BN is a directed acyclic graph consisting of nodes and arcs, to represent variables and the dependencies between variables, respectively (Jensen and Nielsen, 2007). Gaussian Bayesian network (GBN) have improved LAI estimates by combining the 3-PG model output with MODIS images (Mustafa et al., 2011). Their approach relies on availability of satellite images. Remote sensing data, however, often contain gaps (missing values) due to atmospheric characteristics. Mustafa et al. (Subm.) integrated an Expectation Maximization algorithm with GBN to estimate missing satellite values.

The objective of this work is two-fold. First, we use the GBN model to estimate LAI values using the finer resolution satellite imagery ASTER with the point mode 3-PG model. Second, we develop a spatial GBN model to improve estimation of LAI by considering a grid mode 3-PG model bases on 15 m ASTER resolution.

ASTER imagery has been employed in this work because of its relatively high spatial resolution in the visible to near infrared bands, and high spectral resolution in the shortwave infrared bands. Provision of higher order data products, such as atmospherically corrected surface reflectance data, is increasing the applicability of ASTER images (Abrams, 2000). Furthermore, ASTER imagery may help to reduce the uncertainty of the estimated LAI by GBN.

MATERIALS AND METHOD

The 3-PG Model

The 3-PG model is a process-based stand-level model of forest growth. A full description of the 3-PG has been provided by Landsberg and Waring (1997). It requires few parameters and readily available site and climatic data as inputs. Its primary output variables are net primary production (NPP), the standing biomass in foliage, stem (i.e. all above-ground woody tissue) and roots, stem number, available soil water and evapotranspiration. It infers LAI (LAI_{3PG}), mean stem diameter at breast height, main stem volume, and mean annual increment. The 3-PG model has been modified by Coops et al. (1998) and called 3-PGS. They demonstrated the application of 3-PG to forested areas using remote sensing data. The 3-PGS produces spatially and temporally explicitly outputs at the scale of the input surfaces. Spatial outputs include variables such as above and below ground biomass, LAI, stem volume, and current annual increment. These models have been used in many areas like eastern Brazil, and British Columbia (Almeida et al., 2010; Coops et al., 2010; Landsberg et al., 2001). However, it is difficult to parameterize the models precisely e.g. to minimize the uncertainty of the model output.

Remote Sensing Imagery

The estimation of LAI by satellite remote sensing, in particular ASTER sensors, has been investigated in several studies at various spatial scales and environments (Heiskanen, 2006; Ito et al., 2007; Peng et al., 2003; Zheng G. and M., 2009). The ASTER instrument acquires surface data in the visible to near infrared (VNIR, three bands at 15 m/pixel), shortwave infrared (SWIR, six bands at 30 m/pixel), and thermal infrared (TIR, five bands at 90 m/pixel) wavelength regions of the electromagnetic spectrum (Abrams, 2000). Each ASTER scene captures a 60 × 60 km area.

Estimation of the forest variables using optical remote sensing data has been based on empirical relationships formulated between the forest variables measured in the field and satellite data, often expressed in the form of spectral vegetation indices (SVI). Peng et al. (2003) compared twelve different vegetation indices (ranging from visible to shortwave infrared bands) with LAI and found that modified non-linear vegetation index (MNLI), simple ration (RS), and normalized vegetation index (NDVI) correlates best with LAI. Current techniques for estimating LAI often failed to provide consistent values. Furthermore, most LAI satellite data (LAI_{SAT}) products are not continuous in space and time because of a cloud contamination and an insufficient number of data points for retrieval.

Bayesian Networks

A Bayesian Network (BN) is a probabilistic graphical model that provides a graphical framework of complex domains with lots of inter-related variables (Jensen and Nielsen, 2007). Mustafa et al. (2011) designed a network to improve LAI estimation by combining LAI values derived from satellite images and estimated by the point mode 3-PG model. Figure 1(a) shows the graphical part of BN. The intermediate node (LAI_{BN}) represents the estimated LAI values of BN. Based on the continuous variation of LAI over time, it has shown in (Mustafa et al., 2011) that LAI follow normal distribution. The common type of a BN containing continuous variables is the GBN (Shachter and Kenley, 1989). A GBN is a BN where the joint probability distribution associated with its variables $\mathbf{LAI} = \{LAI_1, \dots, LAI_n\}$ is the multivariate normal distribution $N(\mu, \Sigma)$, given by $f(\mathbf{LAI}) = (2\pi)^{-n/2}|\Sigma|^{-1/2} exp\left\{-\frac{1}{2}(\mathbf{LAI} - \mu)^T \Sigma^{-1}(\mathbf{LAI} - \mu)\right\}$. Here μ is the *n*-dimensional mean vector, and Σ is the $n \times n$ positive definite covariance matrix with determinant $|\Sigma|$. The conditional probability distribution of the LAI_i represented by the LAI_{BN_i} as the variable of interest given its parentage, is the univariate normal distribution with density

$$f\left(LAI_{BN_i}|pa_i\right) \sim N\left(\mu_i + \sum_{j=1}^{\#pa_i} \beta_{ij}\left(pa_{ij} - \mu_{pa_{ij}}\right), v_i\right), \quad (1)$$

where μ_i is the expectation of LAI_{BN_i} at time i, the β_{ij} are a regression coefficients of LAI_{BN_i} on its parents, $\#pa_i$ is the number of parents of LAI_{BN_i}, and $v_i = \Sigma_i - \Sigma_{ipa_i}\Sigma_{pa_i}^{-1}\Sigma_{ipa_i}^T$ is the conditional variance of LAI_{BN_i} given its parents. Further, Σ_i is the unconditional variance of the LAI_{BN_i}, Σ_{ipa_i} are the covariances between LAI_{BN_i} and the variables pa_i, and Σ_{pa_i} is the covariance matrix of pa_i.

For $i > 2$, Mustafa et al. (2011) defined equation to identify the intermediate node of the GBN based on the contribution of all of the satellite images, 3-PG output, and previous GBN output as follows:

$$LAI_{BN_i} = \rho(1-\tau)\, LAI_{SAT_i} + \tau\, LAI_{3PG_i} + (1-\rho)(1-\tau)\, LAI_{BN_{i-1}}, \quad (2)$$

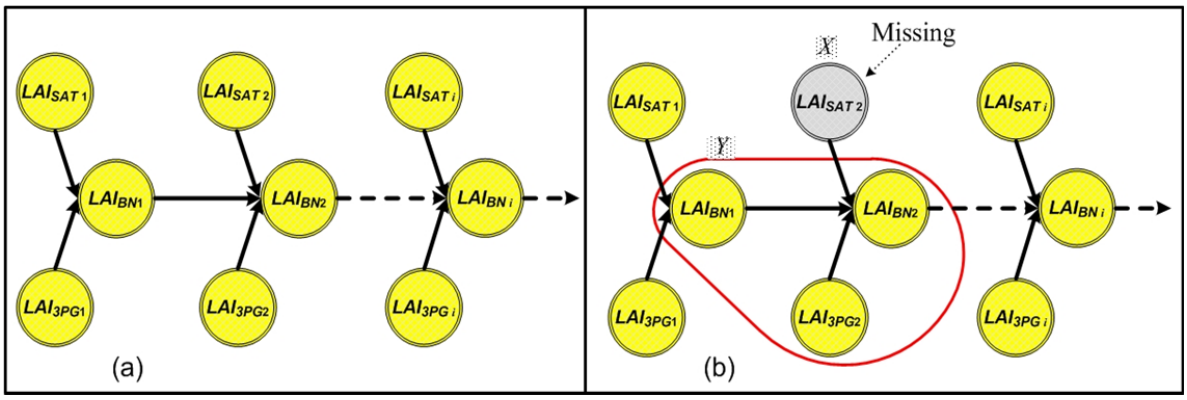

Figure 1. (a) The BN for i^{th} moments. Each moment consists of three nodes LAI_{3PG_i}, LAI_{BN_i} and LAI_{SAT_i} obtained from the 3-PG model, the BN, and satellite images, respectively; (b) BN with missing satellite observations. Y represents an observed data set consisting of three nodes LAI_{BN_2} , LAI_{BN_1} and LAI_{3PG_2} while X represents the variable LAI_{SAT_2} for which an observation is missing.

where τ and ρ are the weighing values, defined as $\tau = \left|(LAI_{SAT_i} - LAI_{SAT_{i-1}})/LAI_{SAT_{i-1}}\right|$ and $\rho = \left|(LAI_{3PG_i} - LAI_{3PG_{i-1}})/LAI_{3PG_{i-1}}\right|$. They are proportional to the change in the LAI values obtained from the satellite images and 3-PG output. Hence the conditional distribution of LAI_{BN_i} equals

$$LAI_{BN_i} \sim N\left(\mu_{LAI_{BN_i}} + \beta_{LAI_{BN_i}LAI_{BN_i}}\left(LAI_{SAT_i} - \mu_{LAI_{SAT_i}}\right) + \beta_{LAI_{BN_i}LAI_{BN_{i-1}}}\left(LAI_{BN_{i-1}} - \mu_{LAI_{BN_{i-1}}}\right) + \beta_{LAI_{BN_i}LAI_{3PG_i}}\left(LAI_{3PG_i} - \mu_{LAI_{3PG_i}}\right), \Sigma_{LAI_{BN_i}}\right) \quad (3)$$

For more details about a GBN of improving forest growth estimates using point mode 3-PG model and satellite images, we refer to (Mustafa et al., 2011).

In the current work, we modified the equation of the intermediate node in GBN model in order to model spatio-temporal $LAI_{BN_{ik}}$ value by contribution from all of the satellite images, 3-PG output and previous GBN output within two consecutive moments as follows:

$$LAI_{BN_{ik}} = \begin{cases} (1-\lambda)(1-\alpha)LAI_{SAT_{ik}} + \alpha\, LAI_{3PG_{ik}} + \lambda\,(1-\alpha)\, LAI_{BN_{(i-1)k}} & \text{if } \left|\frac{LAI_{SAT_{ik}} - LAI_{BN_{(i-1)k}}}{LAI_{BN_{(i-1)k}}}\right| \le \left|\frac{LAI_{SAT_{ik}} - LAI_{3PG_{ik}}}{LAI_{3PG_{ik}}}\right| \\ \lambda(1-\alpha)LAI_{SAT_{ik}} + \alpha\, LAI_{3PG_{ik}} + (1-\lambda)\,(1-\alpha)\, LAI_{BN_{(i-1)k}} & \text{Otherwise} \end{cases} \quad (4)$$

Here i is the image number (the moment number), $k = (n, m)$ is the pixel location, and $\alpha = \gamma\delta, \lambda = \omega\eta$, where γ, δ, ω, and η are the weights, defined as $\gamma = \left|\frac{LAI_{SAT_k} - LAI_{SAT_{k-1}}}{LAI_{SAT_{k-1}}}\right|$, $\delta = \left|\frac{LAI_{SAT_{ik}} - LAI_{SAT_{(i-1)k}}}{LAI_{SAT_{(i-1)k}}}\right|$, $\omega = \left|\frac{LAI_{3PG_k} - LAI_{3PG_{k-1}}}{LAI_{3PG_{k-1}}}\right|$, and $\eta = \left|\frac{LAI_{3PG_{ik}} - LAI_{3PG_{(i-1)k}}}{LAI_{3PG_{(i-1)k}}}\right|$. They are proportional to the spatial and temporal change in the LAI values obtained from the satellite images and 3-PG output.

Expression (4) includes the GBN output from the previous time step ($LAI_{BN_{(i-1)k}}$) to ensure that the LAI values in the new node are consistent with the LAI_{SAT} images, the LAI_{3PG}, and the LAI_{BN} at the previous iteration. This is based on the assumption that the LAI values do not sharply change in a short period of time. In fact, this choice for δ and η addresses the difference between the LAI field data (LAI_{FD}), LAI_{3PG} values, and LAI_{SAT} images at two consecutive time steps. Moreover, γ and ω consider the difference between the LAI_{3PG} values, and LAI_{SAT} images at two neighbor pixels. Weighing these values as in (4) reduces the impact of large discrepancies between the LAI values of 3-PG and satellite images.

EM-algorithm for Estimating Missing Values in a GBN

The Expectation Maximization (EM)-algorithm is a technique for estimating parameters of statistical models from incomplete data. The EM-algorithm is applicable for maximizing likelihoods. The EM-algorithm is formulated and applied to handle the problem of missing satellite data by estimating the missing parameters that are needed to implement a GBN approach (Mustafa et al., Subm.). In this section we give in brief the derivation of EM-algorithm with GBN model.

Consider missing data of satellite images at the i^{th} moment ($i > 1$) of the GBN as shown in Figure 1(b). The GBN output, LAI_{BN_i}, conditionally depends on three nodes (variables), i.e., LAI_{AST_i}, $LAI_{BN_{i-1}}$, and LAI_{3PG_i}, where LAI_{AST_i} is considered as a missing value. Let (X, Y) be the complete data set at the i^{th} moment of GBN, with observed (complete) data $Y = \{LAI_{BN_{i-1}}, LAI_{3PG_i}, LAI_{BN_i}\}$ and missing data $X = LAI_{SAT_i}$ (Figure 1(b)). For clarity, we re-name the variables in the GBN model as $y = LAI_{BN_i}$, $x = LAI_{SAT_i}$, $z = LAI_{BN_{i-1}}$, $w = LAI_{3PG_i}$, such that the GBN model's parameters named as μ_l, σ_l, and β_{yl}, where $l \in \{x, y, z, w\}$.
Hence expression (3) can be reformulated as:

$$f(y|x,z,w) \sim N\left(\mu_y + \beta_{yx}(x-\mu_x) + \beta_{yz}(z-\mu_z) + \beta_{yw}(w-\mu_w), \sigma_y^2\right). \tag{5}$$

The EM-steps to find new ML estimates for the parameters $\theta = (\mu_x, \Sigma_x)$ are as follows:

- Choose an initial setting for the parameters θ and name it as θ^{old}. These are guessed based on seasonal changes of LAI values that are obtained from satellite observations as:

$$\theta^{\text{old}} = \left(\mu^{\text{old}}, \sigma^{\text{old}}\right) = \begin{cases} \left(\mu_x - \left|\frac{\mu_{x_{i-2}} - \mu_{x_{i-1}}}{\mu_{x_{i-1}}}\right|, \sigma_x - \left|\frac{\sigma_{x_{i-2}} - \sigma_{x_{i-1}}}{\sigma_{x_{i-1}}}\right|\right) & \text{if} \quad \mu_{x_{i-2}} \le \mu_{x_{i-1}} \\ \left(\mu_x + \left|\frac{\mu_{x_{i-2}} - \mu_{x_{i-1}}}{\mu_{x_{i-1}}}\right|, \sigma_x + \left|\frac{\sigma_{x_{i-2}} - \sigma_{x_{i-1}}}{\sigma_{x_{i-1}}}\right|\right) & \text{Otherwise} \end{cases} \tag{6}$$

where μ_x, σ_x are the mean and the standard deviation values of the LAI_{SAT}, and obtained either for the period from September to February (non-growing season), or for the period from March to August (growing season). The determination of which period needs to obtain the μ_x, σ_x, is based on the occurrence of missing observation in that period. The $\left|\frac{\mu_{x_{i-2}} - \mu_{x_{i-1}}}{\mu_{x_{i-1}}}\right|$ and $\left|\frac{\sigma_{x_{i-2}} - \sigma_{x_{i-1}}}{\sigma_{x_{i-1}}}\right|$ are the relative changes of the mean and the standard deviation of the previous two LAI_{SAT} observations. Adding or subtracting these relative changes are based on the condition of increase or decrease the LAI_{SAT} during the period of non-growing or growing season.

- E-step: compute the expectation (with respect to the X data) of the likelihood function of the model parameters by including the missing variables as they were observed,

$$\begin{aligned} Q\left(\theta, \theta^{\text{old}}\right) &= E_X\left[\log f(Y, X|\theta) | Y, \theta^{\text{old}}\right] \\ &= \int \log f(Y, X|\theta)\, f\left(X|Y, \theta^{\text{old}}\right) dX = \int \log f(x, y, z, w|\theta)\, f\left(x|y, z, w, \theta^{\text{old}}\right) dx \end{aligned} \tag{7}$$

- M-step: compute the ML estimates of the parameters θ by maximizing the expected likelihood found during the E-step i.e., $\theta^{\text{new}} = \arg\max_\theta Q\left(\theta, \theta^{\text{old}}\right)$. Hence, by differentiation $Q\left(\theta, \theta^{\text{old}}\right)$ with respect to θ, and solve the differentiation equations for $\theta = (\mu_x, \sigma_x)$, the maximum values are found:

$$\mu_x^{\text{new}} = \frac{\sqrt[3]{\phi + \Psi}}{6\lambda} + \frac{2(-3\delta\lambda + \alpha^2)}{3\lambda\sqrt[3]{\phi + \Psi}} + \frac{\alpha}{3\lambda} \text{ and } \sigma_x^{\text{new}} = \sqrt{(\mu_x^{\text{new}})^2 - \frac{2C}{B}\mu_x^{\text{new}} + \frac{E}{B}}, \tag{8}$$

where $\phi = -36\,\delta\alpha\lambda + 108\,\eta\lambda^2 + 8\,\alpha^3$, $\Psi = 12\sqrt{3}\sqrt{4\,\delta^3\lambda - \delta^2\alpha^2 - 18\,\delta\alpha\lambda\eta + 27\,\eta^2\lambda^2 + 4\,\eta\alpha^3}\lambda$, $\lambda = AB$, $\alpha = 2\,AC + DB$, $\delta = AE + B^2 + 2\,DC$, $\eta = CB + dE$, $A = 4\,\frac{a^2\Omega\beta_{yx}^2}{\sigma_y^2}$, $B = 4\,a^2$, $C = 2\,ab\Omega$, $E = 2\,a\Omega + b^2\Omega$, $D = 2\,\frac{ab\Omega\beta_{yx}^2}{\sigma_y^2} - 4\,\frac{a^2\Omega(y - \mu_y - \beta_{yz}(z-\mu_z) - \beta_{yw}(w-\mu_w))\beta_{yx}}{\sigma_y^2}$.

Here $\Omega = V\sqrt{\frac{\pi}{a}}e^{\left(-c+\frac{b^2}{4a}\right)}$, $V = \frac{\sqrt{\sigma_y^2+\beta_{yx}^2(\sigma^{\text{old}})^2}}{\sqrt{2\pi}\sqrt{\sigma_y^2(\sigma^{\text{old}})^2}}$, $a = \frac{1}{2}\left(\frac{1}{(\sigma^{\text{old}})^2}+\frac{\beta_{yx}^2}{\sigma_y^2}\right)$, $b = \left(\frac{\left(y-\mu_y+\beta_{yx}\mu^{\text{old}}-\beta_{yz}(z-\mu_z)-\beta_{yw}(w-\mu_w)\right)\beta_{yx}}{\sigma_y^2}+\frac{\mu^{\text{old}}}{(\sigma^{\text{old}})^2}\right)$ and $c = \frac{b^2}{4a}$.

- Check for convergence of θ^{new} values. If $\left|\theta^{\text{new}}-\theta^{\text{old}}\right| \le \varepsilon$ is not satisfied, then let $\theta^{\text{old}} \leftarrow \theta^{\text{new}}$, and the algorithm returns to E-step, where ε is the stop criterion which has been selected to be 10^{-5}.

For the calculation details of integration EM-algorithm with GBN approach, we refer to (Mustafa et al., Subm.).

Study Area Description

The Speulderbos forests reserves are located between 52°15'08.1" N, 05°41'25.8" E, (Figure 2), covering 2390 ha, near the village of Garderen, The Netherlands. A climate station is placed within a dense 2.5 ha Douglas-fir (*Pseudotsuga menziesii*) stand planted in 1962. The tree density is 785 trees ha^{-1}. The tree height in 1995 was approximately 22 m, and it has increased to 32 m in 2006. The single sided LAI varies between 8 and 11 throughout the year(Steingröver and Jans, 1995). The stand is surrounded by a larger forested area of approximately 50 km^2. The nearest edge is at a distance of 1.5 km southeast from the site. A small clearing of 1 ha is situated to the north of the stand (Van Wijk et al., 2001). For this work, we consider the area around the tower which is of 2.5 ha Douglas-fir trees only (Figure 2).

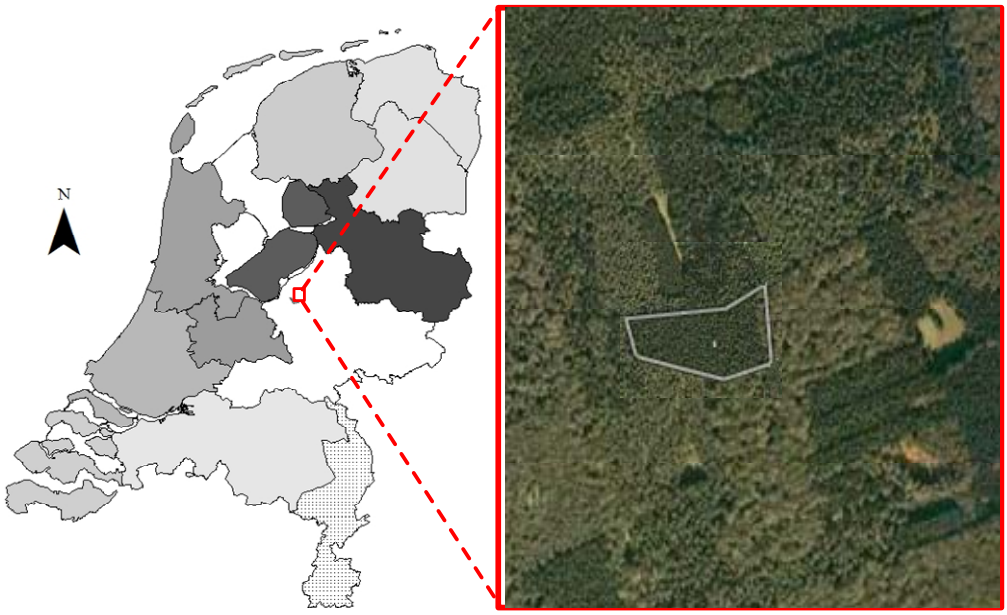

Figure 2. Location of the study area in the Speulderbos forest

Data Description

The study is primarily based on the ASTER surface reflectance products (AST_07) data and fieldwork measurements from June 3, 2010 to October 16, 2010.

***Ground data*:** They were collected at the observation tower of the Speulderbos forest, which is equipped with a weather station and various scientific instruments. The required data for 3-PG model are climate data (16-day mean temperature, solar radiation, rainfall, vapor pressure deficit, and frost days), site factors (site latitude, maximum available water stored in the soil, and soil fertility rating), initial conditions (stem, root, and foliage biomass; stocking; and soil water at some time), and 3-PG parameters. The parameters values which are used in this work are obtained and used by Waring and McDowell (2002). The spatial datasets were required to run the grid mode 3-PG model in this work contained soil data as the stored water in the soil and the grid location at 15 m resolution same as ASTER

resolution. In addition, the LAI was measured in the ground to validate the estimates of the LAI_{BN} values. The LAI is indirectly measured from the canopy transmission by the inversion of the measurements of the photosynthetically active radiation (PAR) above and below the canopy. These LAI values are adjusted by a clumping factor. For more details about calculating LAI at the Speulderbos forest, we refer to (Mustafa et al., 2011). Furthermore, spatial measurements of the LAI were surveyed in June 2010. A total 15 measurements were randomly taken at the study area using LAI-2000 Plant Canopy Analyzer (Li-COR). These measurements are considered as an effective LAI ($eLAI_{FD}$). The ground points of the spatial LAI measurements were located using a handheld GPS device (Magellan Meridian Platinum). According to the manufacturer the device should have an accuracy of 7m for 95% of time. Furthermore, the GPS measurements were averaged over several minutes in order to enhance the accuracy. The spatial $eLAI_{FD}$ values used to establish the empirical relationship with ASTER data. We used effective LAI due to the fact that it is more closely associated with the nature of information gathered by above canopy remote sensors (Chen et al., 2004). Chen and Cihlar (1996) suggested using effective LAI for radiation interception consideration because $eLAI_{FD}$ is less variable and easier to measure than LAI, and because it has better correlation to the satellite vegetation indices than does LAI.

***Remotely sensed data*:** Three cloud free ASTER scenes of level 2 surface reflectance product (AST_07) were acquired from June 3, to October 16, 2010. Each scene was rectified using 10 GCPs collected from 1: 25 000 topographic map sheets with a final Root Mean Square Error (RMS) of about 0.5 pixels. The VNIR bands of all images were resampled to a pixel size of $15m^2$ using nearest-neighbor resampling. A nonlinear relationship has been found between the calculated NDVI from ASTER images ($NDVI_{AST}$) and the spatial measurements of $eLAI_{FD}$ ($eLAI_{FD} = 0.74e^{2.11\,NDVI_{AST}}$, R^2=0.80, Figure 3). The LAI is calculated based on the empirical relationship and it has been considered as the LAI values from ASTER images (LAI_{AST}). This LAI_{AST} used in GBN model as the input source of the satellite imagery.

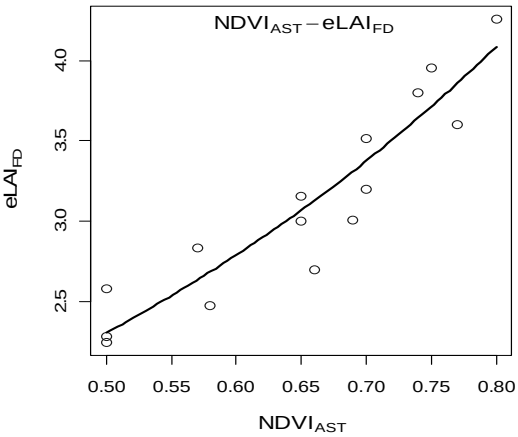

Figure 3. The relationship between $NDVI_{AST}$ and $eLAI_{FD}$.

IMPLEMENTATION

The GBN is applied to the Speulderbos forest in The Netherlands where the LAI_{FD} is available as a time series from June 3, 2010 until October16, 2010. Only three ASTER images are available during the time study, whereas five images are missing due to the cloud coverage. The EM-algorithm is applied to estimate five missing LAI_{AST} and used as an input into the GBN. The implementation of this work includes two cases. First, implementation of the GBN of point mode 3-PG model for the full period of four months after estimating missing LAI_{AST} values using the EM-algorithm. Second, is to implement GBN of grid mode 3-PG model for one ASTER data.

RESULTS

GBN Performance with Point Mode 3-PG Model

Figure 4 shows the LAI values estimated from the GBN for a period of four months, along with the LAI derived from the 3-PG model, LAI ground observations, and the LAI ASTER that were found from the empirical relationship between $NDVI_{AST}$ and $eLAI_{FD}$. The accuracy of the LAI 3-PG output is tested using the RMSE and the relative error (RE) rate with respect to the LAI ground observation. We found an RMSE of 1.40 and an RE of 18.74% (Table 1). The 3-PG overestimated the LAI values across the studied period. The LAI_{AST} shows a big difference with respect to the LAI_{FD}, with an RMSE and an RE of 3.08 and 45.58%, respectively (Table 1). The missing LAI_{AST} values are estimated with significant values such that no big deviation noticed with respect to the non-missing LAI_{AST} values. They are indicated in Figure 4 as a black square symbols. For the four month period, we notice that the

Table 1. Mean, RMSEs, and REs for the various way to estimate the LAI.

	Mean	RMSE	RE rate
LAI_{FD}	6.68		
LAI_{AST}	3.61	3.08	45.82%
LAI_{3PG}	7.88	1.40	18.74%
LAI_{BN}	6.51	0.46	5.86%

combination of the LAI_{AST} values and LAI_{3PG} in a GBN reduces the RMSE to 0.41 and the RE to 5.86%.

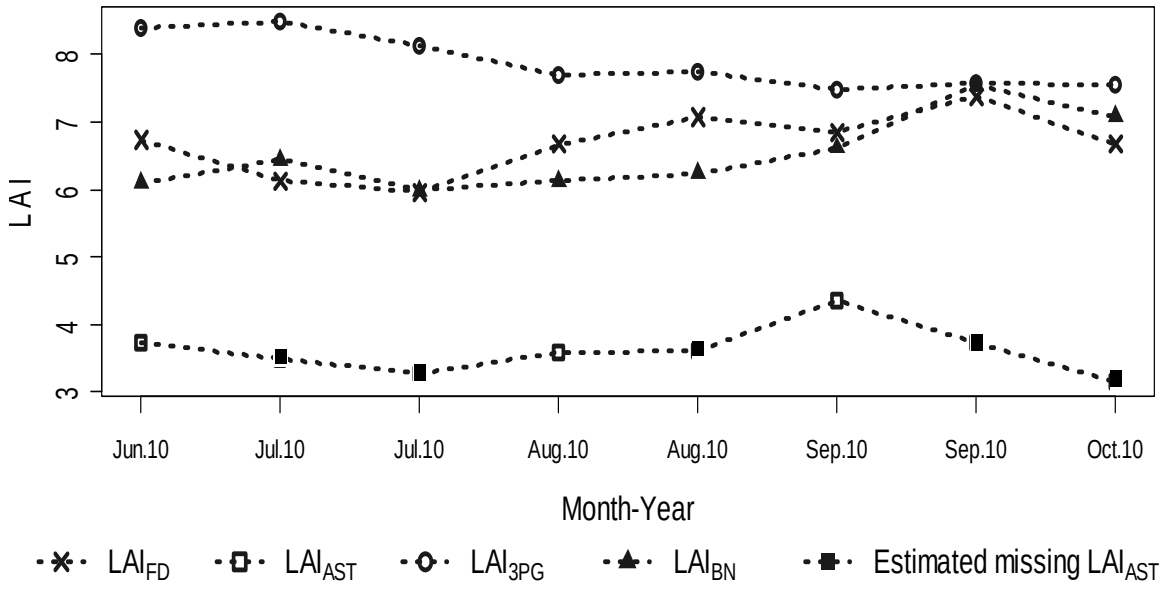

Figure 4. LAI distribution of the Douglas-fir trees obtained from four sources, namely, the field data, the 3-PG model, the ASTER data, and the GBN.

Spatial GBN Performance with Grid Mode 3-PG Model

In this work, a full spatial LAI field dataset for the study area is not available at the ASTER resolution. The spatial output of the GBN is averaged over all pixels and compared with the LAI_{FD} calculated from the PAR data at the tower. Therefore, the accuracy of the spatial GBN is assessed using the absolute error of the averaged spatial GBN with respect to the LAI ground observation at the tower. The absolute error between LAI_{FD} and the LAI_{AST}, LAI_{3PG}, and LAI_{BN} are 3.02, 1.65, and 0.64, respectively. The absolute error between LAI_{BN} and LAI_{FD} is lower than the absolute error between LAI_{3PG} and LAI_{FD}. This indicates that the output of the GBN is more accurate than the 3-PG model output. Figure 5 (a), (b), and (c) show the spatial LAI values of ASTER images, grid mode 3-PG model, and spatial GBN output of the study area.

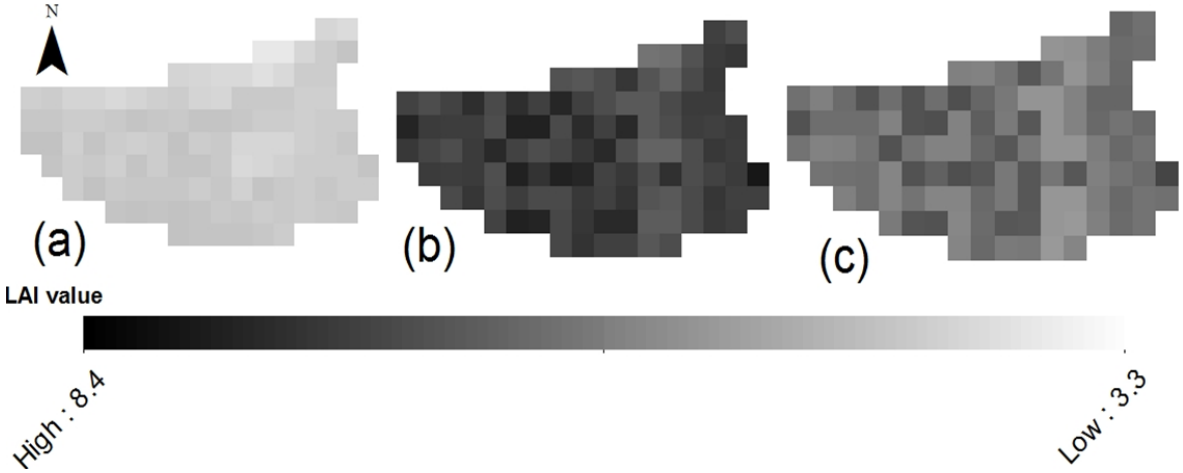

Figure 5. Spatial LAI values of the Speulderbos forest obtained from one observation at June 3, 2010. (a) LAI ASTER, (b) LAI output of grid mode 3-PG model, and (c) LAI estimated by the spatial GBN.

DISCUSSION AND CONCLUSION

In this work, the GBN of the point mode 3-PG model with the EM-algorithm is applied to the Douglas-fir trees in the Speulderbos forest to estimate the missing LAI ASTER and improve LAI values during the period of four months (June until October, 2010). Moreover, the spatial estimation of LAI with 15m resolution is achieved for one ASTER image acquired in June 2010, by modifying the GBN model of the grid mode 3PG model.

Our results show that the missing LAI_{AST} is estimated with a small deviation among the non-missing LAI_{AST} values. From the results we observed that the deviation of the GBN output and LAI_{FD} is lower than the deviation between the 3-PG model output and the LAI_{FD}, indicating that the LAI output of the GBN is more accurate than that of the 3-PG model output alone and is closer to the LAI_{FD}.

A major contribution of this work is to apply the methods of (Mustafa et al., Subm.; Mustafa et al., 2011) to finer spatial resolution images (ASTER). It shows the applicability of their methods with different satellite imagery. Moreover, the GBN is modified to infer 15m spatial estimation of LAI using a grid mode 3-PG model. It addresses the spatial LAI values pixel by pixel of ASTER image and 3-PG model output. As a strategy for the consideration of these two products, ASTER and 3-PG model, in a spatial GBN node, we resorted to the mathematical formulation in (4). From this equation, we can identify the intermediate node of the spatial GBN based on the contribution of the ASTER images, 3-PG output, and previous spatial GBN output. However, to account for the uncertainty in both ASTER images and 3-PG model, weighing factors are introduced. This new expression also includes the spatial GBN output of a previous moment due to the fact that the LAI values have no large changes within a relatively short period of space and time.

A drawback of this work is that the spatial values of the LAI ground measurements of the study area are not available. This was a restriction during validation with respect to the spatial GBN and that was also the reason of applying the spatial GBN for one ASTER observation. However, this study can be applied to a small or a big area whenever the appropriate field data are available.

Also, some other issues require further work. For instance, the time period of this study needs to be extended to include at least one year. This may show a better explanation of the model in terms of LAI seasonal change. Also, ground measurements should be collected for the study area as much as possible to the spatial and temporal resolution as the ASTER images to use as a validation of the spatial GBN model. Further, the modified version of GBN needs to be applied for more than one satellite image. This may be addressed by developing the EM-algorithm with the GBN of (Mustafa et al., Subm.) by integrating the EM-algorithm with the spatial version of the GBN to estimate the spatial missing LAI values of satellite images.

From the present work, we conclude that the GBN model can be applied with different satellite images. Moreover, spatial estimation of LAI can be done with the modified version of the GBN, such that the deviation of the averaged LAI output of spatial GBN and LAI ground measurement is less than the deviation between the averaged LAI output of 3-PG and LAI ground measurement.

REFERENCES

Abrams, M., 2000. The Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER): data products for the high spatial resolution imager on NASA's Terra platform. *International Journal of Remote Sensing 21* (5): 847-859.

Almeida, Auro C., Anders Siggins, Thiago R. Batista, Chris Beadle, Sebastião Fonseca, and Rodolfo Loos, 2010. Mapping the effect of spatial and temporal variation in climate and soils on Eucalyptus plantation production with 3-PG, a process-based growth model. *Forest Ecology and Management 259* (9): 1730-1740.

Bonan, G., 1993. Importance of leaf area index and forest type when estimating photosynthesis in boreal forests. *Remote Sensing of Environment 43* (3): 303-314.

Chen, Jing M., and Josef Cihlar, 1996. Retrieving leaf area index of boreal conifer forests using Landsat TM images. *Remote Sensing of Environment 55* (2): 153-162.

Chen, Xuexia, Lee Vierling, Eric Rowell, and Thomas DeFelice, 2004. Using lidar and effective LAI data to evaluate IKONOS and Landsat 7 ETM+ vegetation cover estimates in a ponderosa pine forest. *Remote Sensing of Environment 91* (1): 14-26.

Coops, N. C., R. A. Hember, and R. H. Waring, 2010. Assessing the impact of current and projected climates on Douglas-Fir productivity in British Columbia, Canada, using a process-based model (3-PG). *Canadian*

Journal of Forest Research 40 (3): 511-524.
Coops, N. C., R. H. Waring, and J. J. Landsberg, 1998. Assessing forest productivity in Australia and New Zealand using a physiologically-based model driven with averaged monthly weather data and satellite-derived estimates of canopy photosynthetic capacity. *Forest Ecology and Management 104* (1-3): 113-127.
Heiskanen, J., 2006. Estimating aboveground tree biomass and leaf area index in a mountain birch forest using ASTER satellite data. *International Journal of Remote Sensing 27* (5-6): 1135-1158.
Ito, E., S. Lim, S. Pol, B. Tith, P. Pith, S. Khorn, A. Tani, M. Kanzaki, T. Kaneko, Y. Okuda, and M. Araki, 2007. *Use of ASTER optical indices to estimate spatial variation in tropical seasonal forests on the west bank of the Mekong river, Cambodia*.
Jensen, F. V., and Thomas Dyhre Nielsen, 2007. *Bayesian networks and decision graphs*. 2nd ed. New York: Springer.
Kalacska, M., A. Sanchez-Azofeifa, T. Caelli, B. Rivard, and B. Boerlage, 2005. Estimating leaf area index from satellite imagery using Bayesian networks. *IEEE Transactions on Geoscience and Remote Sensing 43* (8): 1866-1873.
Landsberg, J. J., K. H. Johnsen, T. J. Albaugh, H. L. Allen, and S. E. McKeand, 2001. Applying 3-PG, a simple process-based model designed to produce practical results, to data from loblolly pine experiments. *Forest Science 47* (1): 43-51.
Landsberg, J. J., and R. H. Waring, 1997. A generalised model of forest productivity using simplified concepts of radiation-use efficiency, carbon balance and partitioning. *Forest Ecology and Management 95* (3): 209-228.
Mustafa, Y. T., V. Tolpekin, and A. Stein, Subm. The application of Expectation Maximization algorithm to estimate missing values in Gaussian Bayesian network modeling for forest growth. *IEEE Transactions on Geoscience and Remote Sensing*.
Mustafa, Y. T., P. E. Van Laake, and A. Stein, 2011. Bayesian Network Modeling for Improving Forest Growth Estimates. *IEEE Transactions on Geoscience and Remote Sensing 49* (2): 639-649.
Peng, Gong, Pu Ruiliang, G. S. Biging, and M. R. Larrieu, 2003. Estimation of forest leaf area index using vegetation indices derived from Hyperion hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing 41* (6): 1355-1362.
Shachter, Ross D., and C. Robert Kenley, 1989. Gaussian influence diagrams. *Management Science 35* (5): 527-550.
Steingröver, E.G., and W.W.P. Jans. 1995. Physiology of forest-grown Douglas-fir trees, effects of air pollution and drought. Wageningen University & Research Centre.Wageningen, The Netherlands.
Van Wijk, M.T., S.C. Dekker, W. Bouten, W. Kohsiek, and G.M.J. Mohren, 2001. Simulation of carbon and water budgets of a Douglas-fir forest. *Forest Ecology and Management 145* (3): 229-241.
Wamelink, G. W. W., H. J. J. Wieggers, G. J. Reinds, J. Kros, J. P. Mol-Dijkstra, M. van Oijen, and W. de Vries, 2009. Modelling impacts of changes in carbon dioxide concentration, climate and nitrogen deposition on carbon sequestration by European forests and forest soils. *Forest Ecology and Management 258* (8): 1794-1805.
Waring, RH, and N McDowell, 2002. Use of a physiological process model with forestry yield tables to set limits on annual carbon balances. *Tree Physiology 22*: 179–188.
Zheng G., and Moskal L. M., 2009. Retrieving Leaf Area Index (LAI) Using Remote Sensing: Theories, Methods and Sensors. *Sensors 9* (4): 2719-2745.