Bayesian Network Modeling for Improving Forest Growth Estimates

Yaseen T. Mustafa, *Member, IEEE*, Patrick E. Van Laake, and Alfred Stein

***Abstract*—Estimating the contribution of the 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 a relevant information is also available from remotely sensed images. This paper aims to improve the LAI estimated from the forest growth model [physiological principals predicting growth (3-PG)] by combining these values with the LAI derived from the Moderate Resolution Imaging Spectroradiometer (MODIS) satellite imagery. A Bayesian networks (BNs) approach addresses the bias in the 3-PG model and the noise of the MODIS images. A novel inference strategy within the BN has been developed in this paper to take care of the different structures of the inaccuracies in the two data sources. The BN is applied to the *Speulderbos* forest in The Netherlands, where the detailed data were available. This paper shows that the outputs obtained with the BN were more accurate than either the 3-PG or the MODIS estimate. It was also found that the BN is more sensitive to the variation of the LAI derived from MODIS than to the variation of the LAI 3-PG values. In this paper, we conclude that the BNs can improve the estimation of the LAI values by combining a forest growth model with satellite imagery.**

***Index Terms*—Bayesian networks (BNs), leaf area index (LAI), Moderate Resolution Imaging Spectroradiometer (MODIS), physiological principals predicting growth (3-PG) model.**

I. Introduction

FORESTS are traditionally important as sources of fuel, building materials, paper, fiber, and timber. More recently, their importance as major storehouses of carbon has been realized, as well as their capacity to exchange carbon dioxide (CO_2) between the vegetation and the atmosphere, which can affect the rate of climate change. Forest growth leads to carbon fixation, and thus, it may turn out to be an important way to diminish the rate of global warming. Modern silviculture may increasingly be able to regulate the rate of forest growth. The application of forest growth models could be helpful in managing the forests by quantifying the storage amounts of CO_2 and by comparing different management scenarios [1], [2].

A commonly measured variable in the forests is the leaf area index (LAI). For broadleaf canopies, the LAI is defined as the one-sided green leaf area per unit of ground area, whereas for coniferous canopies, it is defined as the hemisurface needle leaf area. In both instances, it is equal to the ratio of the total upper leaf surface of the vegetation divided by the surface area of the land on which the vegetation grows. The LAI values typically range from 3 to 15, whereas in some cases, the LAI is greater than 15 [3]. The LAI can be used to scale the measurements of photosynthesis, transpiration, and light interception from the leaf level to the canopy level. This can ultimately be used as an indicator of forest growth. A major challenge is to make a representative estimate of the LAI for large forest-covered areas.

Traditionally, three independent ways are distinguished to determine the LAI: by means of the measurements in the forest, by means of applying a forest growth model, and by means of a remote-sensing image. The LAI is directly or indirectly measured in the field. Direct methods normally involve destructive sampling of canopy elements, and in large complex forest, it may be impossible to collect sufficient samples to accurately characterize the structure. Indirect methods include both light interception instrumentation and hemispherical photography [4]. Such measurements, however, are time consuming, in particular, when obtaining a long time series. Process-based models have been used to estimate the LAI. Several of these models have been established recently, such as the physiological principals predicting growth (3-PG) model developed by Landsberg and Waring [1], the CABALA model [5], and the FORGRO model [2]. The 3-PG model, which we focus in this paper, is a process-based stand-level model of forest growth. It requires readily available site and climatic data as inputs, and it outputs the time course of stand development in a form that is familiar to the forest manager, as well as the LAI, biomass pools, stand water use, and available soil water. Moreover, the model output and several 3-PG model parameters (e.g., the average monthly root turnover rate, the fraction of the net primary production (NPP) to the roots, the fertility rating, the canopy conductance, and the ratio of soil organism biomass) contain uncertainties. Such error is multiplicative, and therefore, error in one parameter affects the determination of the next parameter in the modeling chain [6], [7]. Similarly, modern remote sensing can also provide the LAI estimates. For instance, the Moderate Resolution Imaging Spectroradiometer (MODIS) sensor provides eight-day global data sets of the LAI. Such images, however, have a relatively coarse resolution as compared to the size of the forest stand. MODIS imagery also has uncertainties as instrument noise exists during image acquisition, and atmospheric characteristics are constantly changing [8], [9].

Statistical methods combining images with forest models and observations may help to reduce uncertainties, thus attaining a

Manuscript received February 17, 2010; revised May 18, 2010; accepted July 5, 2010. Date of publication September 7, 2010; date of current version January 21, 2011.

The authors are with the Faculty of Geo-Information Science and Earth Observation, University of Twente (ITC), 7500 AA Enschede, The Netherlands (e-mail: mustafa@itc.nl; vanlaake@itc.nl; stein@itc.nl).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2010.2058581

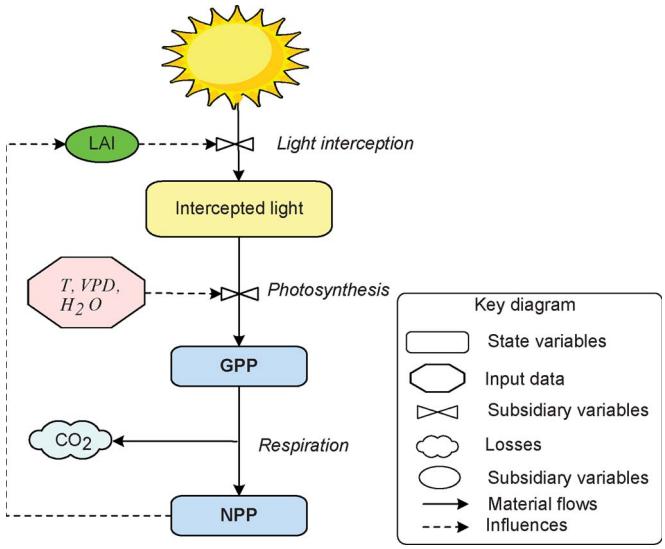

Fig. 1. Basic structure part of the 3-PG and the causal influences of its variables and processes. The symbols used stand for gross primary production (GPP), NPP, air temperature (T), vapor pressure deficit (VPD), water (H_2O), carbon dioxide (CO_2), and LAI.

more accurate LAI estimated value. Graphical models [particularly the Bayesian networks (BNs)] provide a promising way of dealing with this due to their ability to integrate different sources and to deal with uncertainty in a probabilistic way.

A BN is a probabilistic methodology for combining graphs and probabilities to express the relationships between variables [10]. In a BN, each node represents a variable that has a value, while the relations between the nodes are expressed in terms of conditional probabilities. The input data may influence any of the nodes. Once the data are obtained and assigned to the node, a Bayesian mechanism propagates their values to the subsequent nodes of the network. A BN can be used for the combination of information coming from separate sources with varying degrees of reliability. The BNs have been successfully applied to molecular biology, particularly in genetics and biotechnology [11]–[13], forensic sciences [14], computer science, image processing, and artificial intelligence studies [15].

The objective of this paper is to explore the possibilities of the BNs to improve LAI estimations by combining the 3-PG model output with MODIS images. Ultimately, this may be used as a predicting model of the LAI values. The study is applied to a forest area in The Netherlands (Speulderbos forest), where the detailed data are available for the calibration of the 3-PG model and for the validation of the BN.

A. 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 [1] and Sands and Landsberg [16], and a brief illustration of the model is shown in Fig. 1. It requires basic silvicultural and readily available site and climatic data as inputs and predicts as outputs the time course of the stand development in a form that is familiar to the forest manager. It further produces the LAI$(LAI_{3\mathrm{PG}})$, biomass pools, stand water use, and available soil water. The 3-PG model bridges the gap between conventional empirical models on the one hand and process-based carbon balance models on the other hand. It consists of five simple submodels: assimilation of carbohydrates; distribution of biomass between foliage, roots, and stems; determination of stem number; soil water balance; and conversion of biomass values into variables of interest to forest managers. It can be applied to plantations, i.e., even-aged relatively homogeneous forests. Of interest in this paper is the $LAI_{3\mathrm{PG}}$, being an important indicator of the vegetation status and a key parameter in process-based models to quantify the exchange of matter and energy flow between the vegetation and the atmosphere. The 3-PG model has been used to estimate and predict the LAI in different areas [7], [16].

The 3-PG model can be run for any number of years using monthly weather data for each year or monthly averages for the year. In this paper, we modified it to run for 16-day periods, matching with the temporal resolution of MODIS images. A major concern, however, is the quality of the output. Uncertainty in the LAI output of the 3-PG model exists as the LAI values may be modified by conditions or events that are difficult to model or to predict, such as droughts and flooding or pests and diseases. Several studies considered the uncertainty and variation of 3-PG model parameters [16], [17]. Esprey *et al.* [6] reported results from a sensitivity analysis of the 3-PG model, in particular, concerning the LAI.

B. LAI Estimation Using MODIS Satellite Imagery

The estimation of the LAI from satellite imagery may serve as a proxy for field measurements at regional and global scales [18]. Many studies have been carried out to estimate the LAI values from different satellite sensors, in particular, using the MODIS sensors onboard the Earth Observing System Terra/Aqua platforms [8], [9], [19]. The MODIS sensor produces a standard suite of global products characterizing the vegetation cover, the LAI, and the fraction of absorbed photosynthetically active radiation (FPAR) at the 1-km spatial resolution based on observations and composited over an eight-day period. Morisette *et al.* [20] reviewed techniques employed in various countries to produce and evaluate the LAI products derived from satellite measurements, and Yang *et al.* [8] summarized the experience of several collaborating investigators on the validation of the MODIS LAI products. In this paper, we derived the LAI MODIS (LAI_M) from the normalized difference vegetation index (NDVI) MODIS product every 16 days at a 250-m spatial resolution, thus matching the temporal frequency of the 3-PG model and ground data. Hence, we refer to this as a modified MODIS LAI to distinguish it from the LAI MODIS product provided by NASA. Current techniques for estimating LAI often failed to provide consistent values. Estimating LAI from one satellite instrument MODIS is an ill-posed inversion problem because the number of unknowns is always larger than the available bands due to the nature of the Earth's complex environment. Furthermore, most LAI_M data products are not continuous in space and time because of a cloud contamination and an insufficient number of data points for retrieval. As a result, LAI_M products need significant improvements. For this reason, some methods for reducing

noise and for constructing high-quality MODIS time-series for further analysis have been formulated [21]. Qin *et al.* [22] estimated the LAI from remote-sensing data by taking advantage of the physical-model inversion. Xiao *et al.* [23] designed a temporally integrated inversion method to produce spatially and temporally continuous LAI products with relatively higher quality. These methods, however, have their own strengths and limitations. Uncertainty of the LAI remains, whereas the BNs provide a framework to combine the model output with the remotely sensed images to estimate the LAI values more precisely.

C. Bayesian Network

A BN is a network consisting of nodes linked with directed arcs (arrows) that allow us to carry out probabilistic reasoning. It is a mathematical methodology of combining graphs and probabilities to express relationships between variables identified by the nodes. Links in the network are configured as a directed acyclic graph, i.e., a graph without feedback [24]. Each node in a BN indicates a random variable, which is either discrete or continuous. The directed arcs linking the nodes are usually expressed in kinship terms such as the parents of a node (nodes with arrows pointing directly to this node), the children of a node (nodes with arrows pointing from this node), and the ancestors (parents at a higher level). These arcs are quantified by conditional probability tables (CPTs). Such networks can be designed using conventional scientific notions of cause and effect. For example, a node A with parents $B_1, \ldots, B_n$ indicates the existence of an arrow from each of the B_i's, $i = 1, \ldots, n$, to A. There exists a CPT of $\Pr(A|B_1, \ldots, B_n)$[10]. Conditional dependence and independence of nodes in a graphical model allows a substantial simplification of joint probabilities. For instance, two nodes A_1 and A_2 are conditionally independent, given a third set B, if all paths between A_1 and A_2 are separated by a node B [10], [24]

$$\Pr(A_1|A_2, B) = \Pr(A_1|B). \tag{1}$$

II. Model Description

A. Initialization of the BN

The main purpose of the BN, as used in this paper, is to enhance the correspondence between the LAI values from the 3-PG model and MODIS estimate. An important assumption underlying the inference within the BN is that biophysical models aim to describe a physical process characterized by competition for resources and seasonality in as much detail as possible. Complications in this process, however, may cause a lack of understanding of the process or its parameters to support such a detailed description. This results in an approximation of the physical process only. Forest biophysical models invariably iterate for a long time (typically 40–120 years) at a weekly or monthly resolution, resulting in hundreds of iterations over the lifetime of the trees in the forest. Small errors in a single iteration quickly accumulate to a large bias even if the biophysical process is well described. Satellite imagery, on the other hand, has an unknown error associated with it, for instance, due to a suboptimal parametrization of the numerical analysis routine or to atmospheric conditions. The BN takes the different structures of the data quality into consideration when updating the LAI values to input them into the next iteration of the 3-PG model.

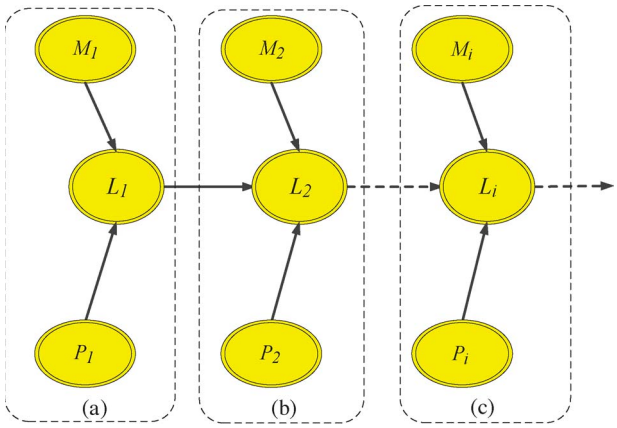

Fig. 2. BN for the i^{th} iterations ($i \geq 1$) every triple nodes M_i, L_i, and P_i indicates the network at time i and refers to the modified MODIS LAI, BN, and 3-PG modes, respectively. The double ellipses (nodes) indicate that the LAI is a continuous stochastic variable.

We start by building a simple network consisting of two nodes, namely, $LAI_{3\text{PG}}$ as the node P and LAI_M as the node M. To serve the objective of this paper, we introduce an intermediate node to combine M_1 and P_1, called L_1, where the subscript indicates the time step 1, taking values from both M_1 and P_1 [Fig. 2(a)]. This network contains two cliques of a child node with its parents, namely, $\{M_1, L_1\}$ and $\{P_1, L_1\}$. The connection type in Fig. 2 is a converging connection type, defined as a BN node with two or more parents, i.e., with two or more arrows pointing to it [24]. The double ellipses (nodes) indicate that LAI is a continuous stochastic variable and follow the normal distribution as illustrated in Section III-C.

The quantitative part of the BN is a joint probability distribution (*JPD*) written as a product of conditional probabilities that describe the dependences between the variables of the network. Given a set of variables $\mathbf{L} = \{L_1, \ldots, L_n\}$, each with its parents, the *JPD* of $\mathbf{L}$ is given by (Fig. 2)

$$\Pr(\mathbf{L}) = \prod_{i=1}^{n} \Pr\left(L_i|pa(L_i)\right) \tag{2}$$

where $pa(L_i)$ corresponds to the parent variables of L_i (i.e., $\{M_i, P_i, L_{i-1}\}$). A common type of a BN containing continuous variables is the Gaussian Bayesian Network (GBN) [25]–[27]. It is a BN where the *JPD* of $\mathbf{L}$ is a multivariate normal distribution $N(\mu, \Sigma)$. The joint density is given by

$$f(\mathbf{L}) = (2\pi)^{-n/2}|\Sigma|^{-1/2} \exp\left\{-\frac{1}{2}(\mathbf{L}-\mu)^T\Sigma^{-1}(\mathbf{L}-\mu)\right\} \tag{3}$$

where μ is the n-dimensional mean vector, Σ is the $n \times n$ covariance matrix with determinant $|\Sigma|$, and T denotes the transpose of a vector. The conditional probability distribution

of L_i that verifies the expression (2) is a univariate normal distribution with density

$$f(L_i|pa(L_i)) \sim N\left(\mu_{L_i} + \sum_{j=1}^{\#pa(L_i)} \beta_{ij}\left(pa(L_i)_j - \mu_{pa(L_i)_j}\right), \nu_{L_i}\right) \tag{4}$$

where μ_{L_i} is the expectation of L_i, β_{ij} is the regression coefficients of L_i on its parents, $\#pa(L_i)$ is the number of parents of L_i, and $\nu_{L_i} = \Sigma_{L_iL_i} - \Sigma_{L_ipa(L_i)}\Sigma^{-1}_{pa(L_i)}\Sigma^{T}_{L_ipa(L_i)}$ is the conditional variance of L_i, given its parents. Here, $\Sigma_{L_iL_i}$ is the unconditional variance of L_i, $\Sigma_{L_ipa(L_i)}$ is the covariances between L_i and its parents $pa(L_i)$, and $\Sigma_{pa(L_i)}$ is the covariance matrix of $pa(L_i)$. For more details of the GBN concept, we refer to [25] and [26], and for an illustrated example, we refer to [28].

The conditional distribution (4) is recovered by considering that

$$f\left(L_i|pa(L_i)\right) \sim N\left(\mu_{L_i} + \sum_{j=1}^{\#pa(L_i)} \beta_{ij} pa(L_i)_j, \nu_{L_i}\right). \tag{5}$$

Under the assumption of Gaussianity, the variable L_i is completely specified by its expectation μ and covariance matrix Σ, obtained from the conditional distribution.

Following Fig. 2(a), we notice that, at the first time step, the marginal distribution of L_1, given its parents (M_1, P_1), yields the expectation and covariance as

$$E(L_1) = \beta_{L_1M_1}\mu_{M_1} + \beta_{L_1P_1}\mu_{P_1} + \mu_{L_1} \tag{6}$$

$$\Sigma_{L_1} = \begin{pmatrix} \sigma_{M_1M_1} & 0 & \sigma_{M_1L_1} \\ 0 & \sigma_{P_1P_1} & \sigma_{P_1L_1} \\ \sigma_{L_1M_1} & \sigma_{L_1P_1} & \sigma_{L_1L_1} \end{pmatrix}. \tag{7}$$

Here, $E(\cdot)$ refers to the expectation, and the subscript L_1 of Σ_{L_1} refers to the covariance matrix at the first time step. Both $\sigma_{L_1L_1}$ and $E(L_1)$ contribute to obtain LAI at the second time step in the network (Fig. 2). Hence, the marginal distribution of L_1 is

$$L_1 \sim N\left(\beta_{L_1M_1}\mu_{M_1} + \beta_{L_1P_1}\mu_{P_1} + \mu_{L_1}, \Sigma_{L_1}\right). \tag{8}$$

The estimated LAI values for P_1 are obtained from the ground truth values. At this stage, we assume that L_1 follows the same distribution as P_1 in order to initialize and implement the network with the truth values of the LAI. This assumption will be generalized when the network will be extended, as shown in the next section.

B. Updating of the BN

The previous section formulated the basis of a BN. Now, we extend the BN over multiple time steps. To do so, the value of L_1, after being derived from the BN, is inserted into the 3-PG model to update the LAI value at its next iteration [Fig. 2(b)]. This graphical representation is similar to Fig. 2(a), but it refers to the second time step. The arrow from L_1 to L_2 links the first time step to the second time step, indicating that the new LAI value in L_2 is conditionally determined by the previous LAI value L_1. This system can further be extended toward more than two time steps [Fig. 2(c)]. To deal with L_i, for $i \geq 2$, and the prior probability distribution of the LAI, the following equation applies:

$$L_i = (1-\alpha)\lambda M_i + \alpha P_i + (1-\lambda)(1-\alpha)L_{i-1} \tag{9}$$

where α and λ are the weighing values, defined as $\alpha = |(M_i - M_{i-1})/M_{i-1}|$ and $\lambda = |(P_i - P_{i-1})/P_{i-1}|$. They are proportional to the change in the LAI values obtained from the MODIS images and 3-PG output.

Expression (9) includes the BN output from the previous time step (L_{i-1}) to ensure that the LAI values in the new node are consistent with the LAI_M images, the LAI_{3PG}, and the BN output (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 α and λ addresses the deviation between the LAI field data (LAI_{FD}), LAI_{3PG} values, and LAI_M images at two consecutive time steps. Weighing these values as in (9) reduces the impact of large discrepancies between the LAI values of 3-PG and MODIS, as shown in Section III-E.

In Fig. 2(b), the distribution at node L_2 can be expressed as

$$f\left(L_2|M_2, L_1, P_2\right) \sim N\left(\mu_{L_2} + \beta_{L_2M_2}M_2 + \beta_{L_2L_1}L_1 + \beta_{L_2P_2}P_2, \nu_{L_2}\right). \tag{10}$$

Now, the expectation of L_2 can be found as

$$E(L_2) = \mu_{L_2} + \beta_{L_2M_2}\mu_{M_2} + \beta_{L_2L_1}\mu_{L_1} + \beta_{L_2P_2}\mu_{P_2} \tag{11}$$

and the covariance can be found as

$$\Sigma_{L_2} = \begin{pmatrix} \sigma_{M_2M_2} & 0 & 0 & \sigma_{M_2L_2} \\ 0 & \sigma_{L_1L_1} & 0 & \sigma_{L_1L_2} \\ 0 & 0 & \sigma_{P_2P_2} & \sigma_{P_2L_2} \\ \sigma_{L_2M_2} & \sigma_{L_2L_1} & \sigma_{L_2P_2} & \sigma_{L_2L_2} \end{pmatrix}. \tag{12}$$

Thus, the marginal distribution of L_2 is equal to

$$L_2 \sim N\left(\mu_{L_2} + \beta_{L_2M_2}\mu_{M_2} + \beta_{L_2L_1}\mu_{L_1} + \beta_{L_2P_2}\mu_{P_2}, \Sigma_{L_2}\right). \tag{13}$$

The previous expressions for the expectation and variance of the LAI_{BN} are considering two iterations. For n iterations, the network is as follows (Algorithm 1; Fig. 3).

1) For $i = 1$, it is assumed that the P_1 and L_1 values are the observed ground values. This assures that the BN starts with realistic values. The expectation and variance for L_1 are calculated from (6) and (7). The new LAI value is inserted into the 3-PG model to update the LAI value at the second time step. Next, the 3-PG model produces a new LAI value, represented as the node P_2 in the network.
2) For $i \geq 2$, we define L_i as in (9), and the distribution of $L_i|M_i, L_{i-1}, P_i$ is equal to

$$f\left(L_i|M_i, L_{i-1}, P_i\right) \sim N\left(\mu_{L_i} + \beta_{L_iM_i}M_i + \beta_{L_iL_{i-1}}L_{i-1} + \beta_{L_iP_i}P_i, \nu_{L_i}\right) \tag{14}$$

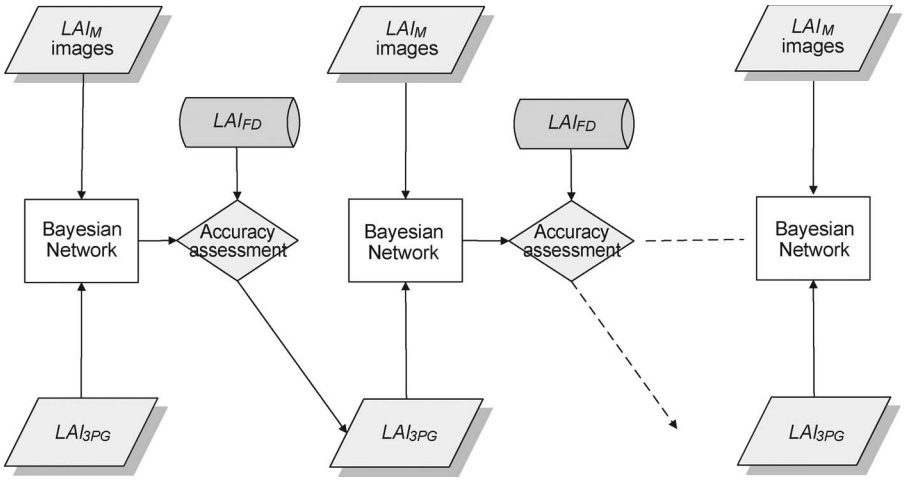

Fig. 3. Workflow process of improving the LAI value.

whereas the expectation and the variance of L_i after the network propagation are

$$E(L_i) = \mu_{L_i} + \beta_{L_iM_i}\mu_{M_i} + \beta_{L_iL_{i-1}}\mu_{L_{i-1}} + \beta_{L_iP_i}\mu_{P_i} \quad (15)$$

$$\Sigma_{L_i} = \begin{pmatrix} \sigma_{M_iM_i} & 0 & 0 & \sigma_{M_iL_i} \\ 0 & \sigma_{L_{i-1}L_{i-1}} & 0 & \sigma_{L_{i-1}L_i} \\ 0 & 0 & \sigma_{P_iP_i} & \sigma_{P_iL_i} \\ \sigma_{L_iM_i} & \sigma_{L_iL_{i-1}} & \sigma_{L_iP_i} & \sigma_{L_iL_i} \end{pmatrix}. \quad (16)$$

To obtain $\Sigma_{L_i}, i \geq 1$, the algorithm presented by Shachter and Kenley [26] is used.

Hence, the marginal distribution of L_i is

$$L_i \sim N\left(\mu_{L_i} + \beta_{L_iM_i}\mu_{M_i} + \beta_{L_iL_{i-1}}\mu_{L_{i-1}} + \beta_{L_iP_i}\mu_{P_i}, \Sigma_{L_i}\right). \quad (17)$$

New LAI values obtained as output from the BN L_i are compared with the LAI_{FD} to asses the agreement between them. This new BN output is inputted into the 3-PG model to update the LAI value, whereas the 3-PG model produces new LAI values, represented as a node (P_{i+1}) in the network.

Algorithm 1: BN Implementation
input: Iteration number (N), LAI_M, $LAI_{3\mathrm{PG}}$.
Output: Updated values of the forest growth parameter "intermediate BN node."
Initial step: for $i \longleftarrow 1$;
$L_1node \longleftarrow f(M_1node, P_1node)$;
$L_1data \longleftarrow P_1data \longleftarrow fielddata$;
after the BN propagates
$L_1 \sim N(\beta_{L_1M_1}\mu_{M_1} + \beta_{L_1P_1}\mu_{P_1} + \mu_{L_1}, \Sigma_{L_1})$;
Update the 3-PG model with L_1 (which represents P_2 in the network);
for $i \leftarrow 2$ **to** N **do**
  $L_inode \longleftarrow f(M_inode, L_{i-1}node, P_inode)$;
  calculate L_i data as (9);
  after the BN propagates
  $L_i \sim N(\mu_{L_i} + \beta_{L_iM_i}\mu_{M_i} + \beta_{L_iL_{i-1}}\mu_{L_{i-1}} + \beta_{L_iP_i}\mu_{P_i}, \Sigma_{L_i})$
  Statistical validation;
  if $L_i \approx LAI_{\mathrm{FD}}$ **then**
    the BN output is close to field observation;
  else
    Update the 3-PG model with L_i (which represents P_{i+1} in the network);
  end
end

C. Sensitivity of the BN

Several methods have been proposed to measure the accuracy of the LAI derived from satellite imagery. Verger *et al.* [29] evaluated the performance of the LAI derived from multisource images and their mutual agreement. Uncertainty in the LAI values derived from the MODIS images is introduced by factors like atmospheric variation and functioning of the sensor. Similarly, the reliability of the 3-PG estimates depends on the accuracy of its input parameters. A full modeling of the uncertainty is beyond the frame of this paper. However, we assess the sensitivity of the BN when the input values LAI_M and $LAI_{3\mathrm{PG}}$ are varied in steps of 0.25 units within the interval $[-1, 1]$. The sensitivity is also assessed when LAI_M and $LAI_{3\mathrm{PG}}$ values are varying simultaneously.

III. BN Implementation: The Speulderbos Forest in the Netherlands

A. Study Area

The Speulderbos forest is located at $52°15'08''$ N, $05°41'25''$ E (Fig. 4), 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. Dominant species in the neighborhood of the Douglas fir stand are the Japanese Larch (*Larix kaempferi*), Beech (*Fagus sylvatica*), Scotch Pine (*Pinus sylvestris*), and Hemlock (*Tsuga spp*) [30], [31].

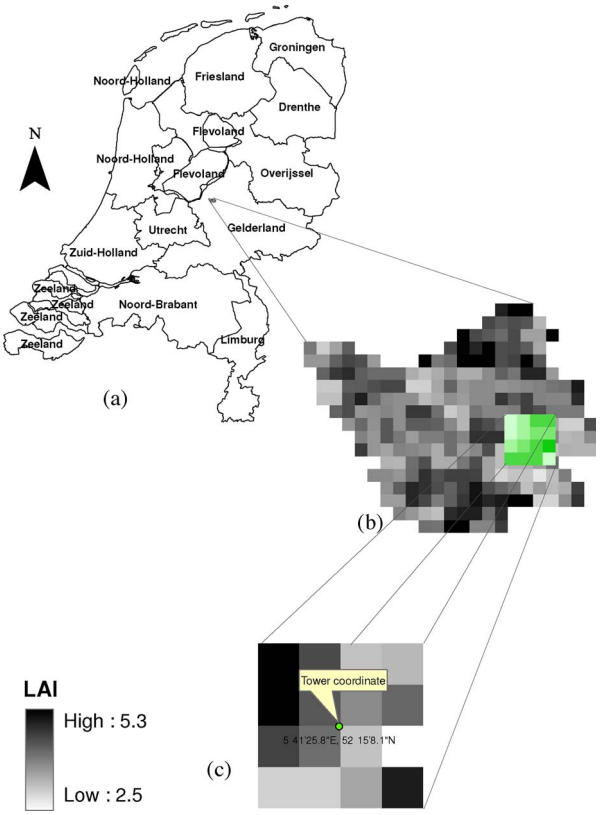

Fig. 4. Map of The Netherlands with the MODIS images of the derived LAI of the study area (acquired on August 13, 2007) with a 250-m spatial resolution. (a) Map regions of The Netherlands. (b) Speulderbos forest. (c) Study area around the tower (tower coordinate $52°15'08''$ N, $05°41'25''$ E).

The forest floor is covered with a needle layer on which little vegetation is present. A trunk space and a crown layer can be discriminated, whereas the trunk space with a few branches and needles reaches a height of approximately 10 m. The single-sided LAI varies between 8 and 11 throughout the year [32]. 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 [33].

B. Data Description

The study is primarily based on the MODIS TERRA land satellite collection 5 data and fieldwork measurements from July 2007 to September 2009.

1) Ground Data: The ground data were collected at the observation tower of the Speulderbos forest, which is equipped with a weather station and various scientific instruments. The following data were captured as needed by the 3-PG model:

1) *climate data*: 16-day mean temperature, solar radiation, rainfall, vapor pressure deficit, and frost days;
2) *site factors*: site latitude, maximum available water stored in the soil, and soil fertility rating;
3) *initial conditions*: stem, root, and foliage biomass; stocking; and soil water at some time;
4) *3-PG parameters*: parameters characterizing the modeled species.

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. The PAR data are acquired from the Speulderbos forest using four sensors placed at the tower. Two are located at the top of the tower to record the PAR outside the canopy forest in two opposite directions: an upward looking sensor to measure the incoming sun radiation and a downward looking sensor to measure the reflected radiation from the canopy. The other two sensors are located below the forest canopy to record the PAR inside the forest. These sensors are also located in two opposite directions: an upward looking sensor to measure the radiation penetration through the canopy coming from the sun and a downward looking sensor to record the reflected radiation from the soil and the forest components. The PAR data are recorded every minute during daytime. The calculation of the LAI from the PAR data is based on the relationship between the leaf area and the light transmittance, described by the Beer–Lambert model [34]. This model utilizes the light canopy transmittance in the Beer–Lambert equation to assess the monthly LAI of a tropical deciduous forest. The daily average LAI_{FD} has been computed from the recorded PAR data after the LAI calculated at every minute of the solar noon. Furthermore, to verify the spatial variation of the LAI, measurements were randomly taken around the tower using LAI-2000 Plant Canopy Analyzer (Li-COR) between the days 274 and 287 of 2008 since they were significantly correlated ($R^2 = 0.74$). As shown by the studies in the area [32], we adjusted the ground LAI observations of Douglas fir trees in Speulderbos by a clumping factor.

Based on the agreement of the temporal resolution of the MODIS images and 3-PG model, the LAI_{FD} measurements have been computed by averaging the LAI daily average every 16 days, which consolidated a set of 50 LAI_{FD} values and which were used to validate the BN output.

The LAI measurements by ground techniques, however, are affected by random errors and bias [4], [35], [36]. The measurements of the PAR are usually concentrated around the solar noon to avoid high solar angles relative to the zenith. Hence, an uncertainty of the PAR measurement arises as to whether the solar angles should be expressed relative to a normal to the inclined surface or to the solar zenith angle.

2) MODIS Data: The low spatial resolution of the LAI MODIS product, however, limits the utility at local scales, as in this paper. To increase the resolution to 250 m, we derived the LAI from the 250-m NDVI MODIS product (MOD13Q1) by establishing the relation between the NDVI and the FPAR. A stronger relation exists between the FPAR and the NDVI than between the LAI and the NDVI [37], [38]. A linear relationship between the NDVI and the FPAR is present, which is approximately linear for green vegetation [39], [40]. This relationship is sensitive to soil background, irradiance quality, and canopy structure.

We used the FPAR values from the LAI/FPAR 8-day L4 global 1 km (MOD15A2) and the NDVI values from the vegetation indices 16-day L3 global 250 m (MOD13Q1) as well. We found a linear relationship between the NDVI and the FPAR

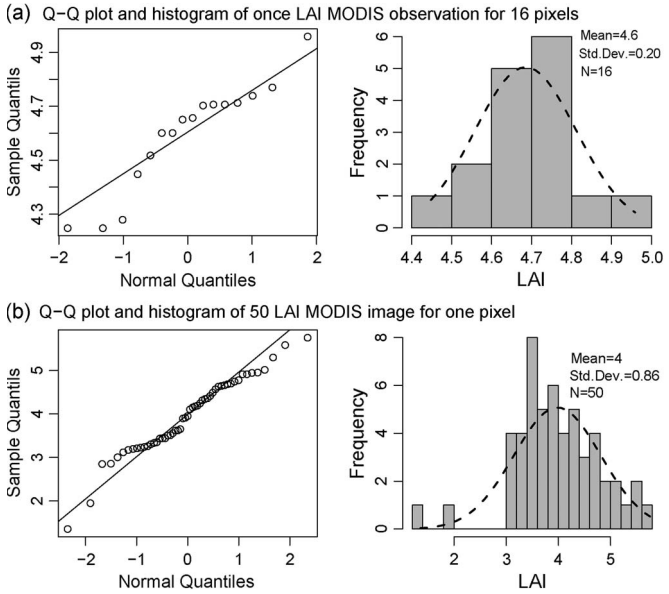

Fig. 5. Two normal Q–Q plots and histograms with the normal curve of LAI_M. (a) Sixteen pixels covering the study area. The Q–Q plot and histogram of one LAI MODIS observation for 16 pixels. (b) Series of a single pixel followed from the 50 LAI_M images during 50 observation times. The Q–Q plot and histogram of the 50 LAI MODIS images for one pixel.

with $R^2 = 0.71$, the linear relationship was defined by comparing the 16-day FPAR composites with the 16-day MODIS NDVI products during the time period (July 2007–September 2009), and we calculated the 250-m LAI as [41]

$$LAI = -\frac{\ln(1 - FPAR)}{k} \tag{18}$$

where FPAR is the FPAR at 250-m spatial resolution derived from the 250-m NDVI based on the relationship equation between them and k is the extinction coefficient. For the illustrated example in the next section, the modified MODIS LAI data are used for two years and two months.

C. LAI Frequency Distribution

Two tests of normality (Shapiro–Wilk and Lilliefors) were applied in order to verify the frequency distribution of the LAI values in each one of the MODIS images and the 3-PG model.

1) Modified MODIS LAI: The normality tests of the LAI values are taken in terms of spatial and temporal resolution. The tests were first applied to a single LAI_M image for the whole study area (16 pixels with 250-m spatial resolution). Second, the tests were applied to the pixels at one location at successive LAI_M images during the full study period (July 2007 until September 2009 at a 16-day temporal resolution). The Q–Q plots and the histograms of the LAI_M values [Fig. 5(a) and (b)] show some deviations from normality, but according to the tests of normality, the LAI_M data do not deviate from a normal distribution [Table I(a)].

2) LAI Output of the 3-PG Model: A test of normality has next been applied to the 16 LAI output values from the 3-PG model that was run at the same spatial and temporal resolution as the LAI_M images. Although Fig. 6(a) and (b) shows that these values deviate from normality, these values were not significantly different from normality [Table I(b)] according to the implemented tests.

TABLE I
TWO NORMALITY TESTS. (a) NORMALITY TESTS FOR THE LAI_M. COLUMN (A) IS COMPOSED OF 16 PIXELS COVERING THE STUDY AREA, AND COLUMN (B) IS A SERIES OF A SINGLE PIXEL FOLLOWED FROM 50 LAI_M IMAGES DURING 50 OBSERVATION TIMES. (b) NORMALITY TESTS FOR THE LAI_{3PG}. COLUMN (A) IS COMPOSED OF 16 LAI_{3PG} SIMULATED VALUES, AND COLUMN (B) IS A SERIES OF 50 OUTPUT VALUES DURING THE TIME SERIES (50 ITERATIONS)

(a)

Test		(A)	(B)
Shapiro-Wilk	P-value	0.09626	0.1714
	Statistic (W)	0.9049	0.9684
Lilliefors	P-value	0.1072	0.3825
	Statistic (D)	0.1948	0.0883

(b)

Test		(A)	(B)
Shapiro-Wilk	P-value	0.99	0.039
	Statistic (W)	0.98	0.89
Lilliefors	P-value	0.98	0.21
	Statistic (D)	0.086	0.02

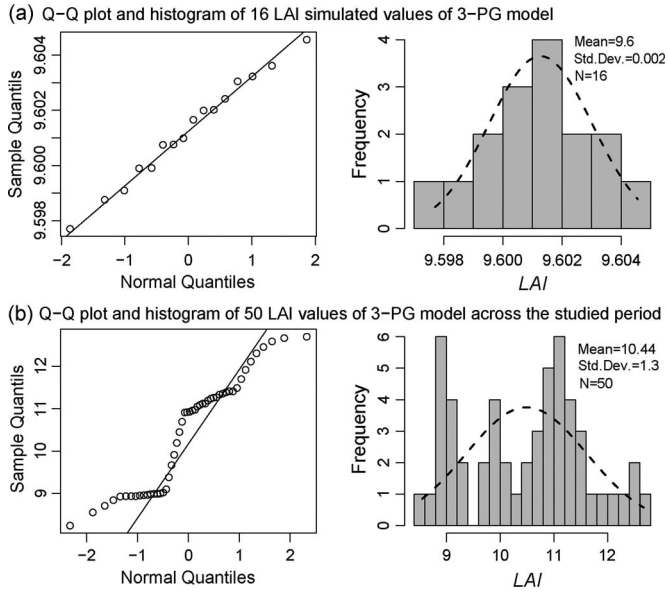

Fig. 6. Normal Q–Q plots and histograms with the normal curve of the LAI_{3PG}. (a) Sixteen LAI_{3PG} simulated values. The Q–Q plot and histogram of the 16 LAI simulated values of the 3-PG model. (b) Series of 50 LAI_{3PG} during the full study period. The Q–Q plot and histogram of the 50 LAI values of the 3-PG model across the studied period.

D. Example Application

We run the BN during 50 iterations, corresponding to a time period of 26 months, and we consider a homogeneous area of 1 km^2 around the climate station in the Speulderbos forest. Such an area is covered with 16 LAI values of the modified MODIS LAI at 250-m spatial resolution. We observe the LAI output values from the 3-PG model, which was executed simultaneously in the BN for 50 iterations, also producing 16 values per iteration for 16-day periods. The MODIS and the 3-PG model were thus harmonized. The first two steps of the BN are presented here as explained in Algorithm 1, and the reminder

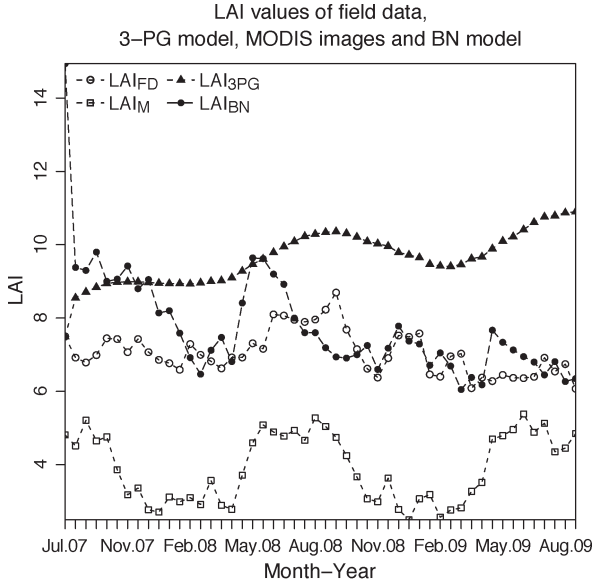

Fig. 7. LAI distribution of the Speulderbos forest obtained from four sources, namely, the field data, the 3-PG model, the MODIS images, and the BN, during the period July 2007–September 2009.

iterations are similar to the second iteration. The results are shown in Fig. 7 as follows.

1) First Iteration: The first iteration initializes the BN in its three nodes, namely, M_1, L_1, and P_1. The data of nodes P_1 and L_1 are the same and are acquired from the field data. Hence, $P_1 = L_1 \sim N(7.47, 1.04)$, and $M_1 \sim N(4.81, 7 \times 10^{-3})$.

The conditional distribution of $L_1|M_1, P_1$ is equal to

$$\begin{aligned}\hat{\mu}_{L_1|M_1,P_1} &= \beta_{L_1M_1} M_1 + \beta_{L_1P_1} P_1 + \hat{\mu}_{L_1} \\ &= 2.77 \times 10^{-15} M_1 + 1 \times P_1 + \hat{\mu}_{L_1} \end{aligned} \quad (19)$$

where $\beta_{L_1M_1}$ and $\beta_{L_1P_1}$ are the regression coefficients relating L_1 to M_1 and P_1, respectively. The regression coefficients are close to zero or one at this first iteration step, but they will deviate from zero and one in the subsequent iterations as the bias in the 3-PG model leads to an accumulation of error. The marginal distribution of L_1 is hence found by implementing the network with the data in (6) and (7). We thus have

$$E(L_1) = 7.44 \times 10^{-15}(4.81) + 7.47 + 7.47 = 14.94 \quad (20)$$

$$\Sigma_{L_1} = \begin{pmatrix} 0.007 & 0 & 1.7 \times 10^{-17} \\ 0 & 1.04 & 1.04 \\ 1.7 \times 10^{-17} & 1.04 & 2.08 \end{pmatrix}. \quad (21)$$

Then, the distribution of node L_1 after BN propagation is equal to $L_1 \sim N(14.94, 2.08)$.

2) Second Iteration: The new LAI value L_1 is inserted into the 3-PG model to calculate the development of the forest over the next time step. This produces the LAI value at the next period. At this time step, the network consists of four nodes, namely, M_2, L_1, P_2, and L_2. The probability distribution of the new P_2 after receiving an updated LAI and after running the 3-PG model is $P_2 \sim N(15.39, 0.0006)$, whereas $M_2 \sim N(4.51, 0.013)$, $L_1 \sim N(14.94, 2.08)$, and $L_2 \sim N(4.77, 0.016)$ according to (9). Therefore, the expectation and variance of $L_2|M_2, P_2, L_1$ are

$$\begin{aligned}\hat{\mu}_{L_2|M_2,L_1,P_2} &= \hat{\mu}_{L_2} + \beta_{L_2M_2} M_2 + \beta_{L_2L_1} L_1 + \beta_{L_2P_2} P_2 \\ &= 4.77 + 0.99 \times M_2 - 0.05 \times L_1 + 0.06 \times P_2 \end{aligned} \quad (22)$$

and the covariance can be found as

$$\Sigma_{L_2} = \begin{pmatrix} 0.013 & 0 & 0 & 0.013 \\ 0 & 2.08 & 0 & 0.88 \\ 0 & 0 & 0.0001 & 1.2 \times 10^{-5} \\ 0.013 & 0.8 & 1.2 \times 10^{-15} & 0.029 \end{pmatrix}. \quad (23)$$

Therefore, the value of LAI (L_2) is distributed as $L_2 \sim N(9.41, 0.029)$.

By observing the difference of this last value with the LAI field observation, we get $|7.49 - 9.41| = 1.92$, whereas the difference between the 3-PG LAI output, when the 3-PG is executed alone, and the LAI_{FD} is equal to $|7.49 - 8.59| = 1.1$.

At this time step, the LAI_{BN} is still far from the validation data LAI_{FD}, whereas after a few iterations, LAI_{BN} converges more rapidly than $LAI_{3\mathrm{PG}}$ (see Fig. 7). Later iterations are similar to the second iteration.

The BN has been implemented using the C++ code. Initially, the 3-PG model produces 16 LAI values that are combined with the LAI_M values in the BN, which introduces the expectation and variance of the LAI.

TABLE II
MEAN VALUES (MEAN), RMSEs, AND REs FOR THE VARIOUS WAYS TO ESTIMATE THE LAI

	Mean	RMSE	Relative error rate
LAI_{FD}	7.05		
LAI_M data	3.9	3.2	44.1%
$LAI_{3\mathrm{PG}}$ data	9.6	2.7	37.4%
LAI_{BN} data	7.7	1.5	14.7%

E. Results

Fig. 7 shows the LAI values estimated from the BN for a period of 26 months, along with the LAI derived from MODIS images, the 3-PG model, and ground observations. The accuracy of the LAI 3-PG output is tested using the root mean square error (RMSE) and the relative error (RE) rate with respect to the LAI ground observation. We found an RMSE of 2.7 and an RE of 37.4%. On average, the $LAI_{3\mathrm{PG}}$ was 9.6, while the LAI_{FD} is 7.05 (Table II). The 3-PG overestimated the LAI values across the studied period.

The modified MODIS LAI shows a high deviation with respect to the LAI_{FD}, with an RMSE and an RE of 3.2 and 44.1%, respectively (Table II). Overall, for this paper, the average LAI_M of 3.9 indicates an underestimation of the LAI values with respect to the LAI_{FD}. Conversely, as shown in Section III-D, after the first iteration, the BN yields more accurate LAI estimates than the 3-PG model and the MODIS images (Table II). For the 26-month period, we notice that the combination of the MODIS image and the 3-PG in a BN reduces the RMSE to 1.5 and the RE to 14.5%.

Fig. 8 shows the output of the sensitivity test implemented by independently varying the $LAI_{3\mathrm{PG}}$ and LAI_M values. As shown in the figure, the BN is more sensitive to the variations

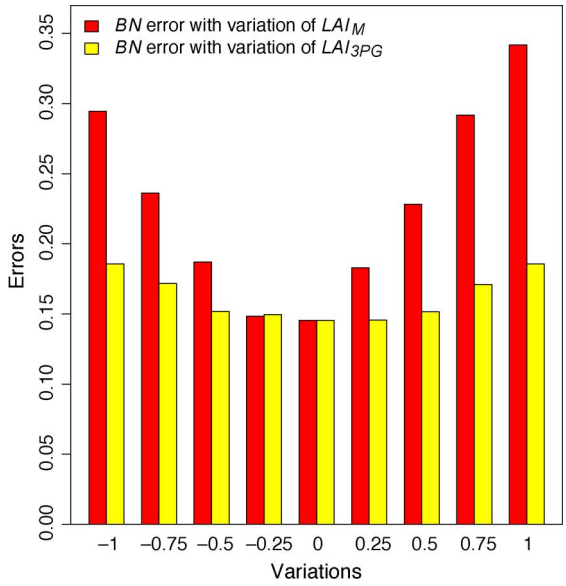

Fig. 8. Sensitivity test of the BN after independently varying the $LAI_{3\mathrm{PG}}$ and LAI_M values in steps of 0.25 units within the interval $[-1, 1]$ across the studied period applied in the Speulderbos forest.

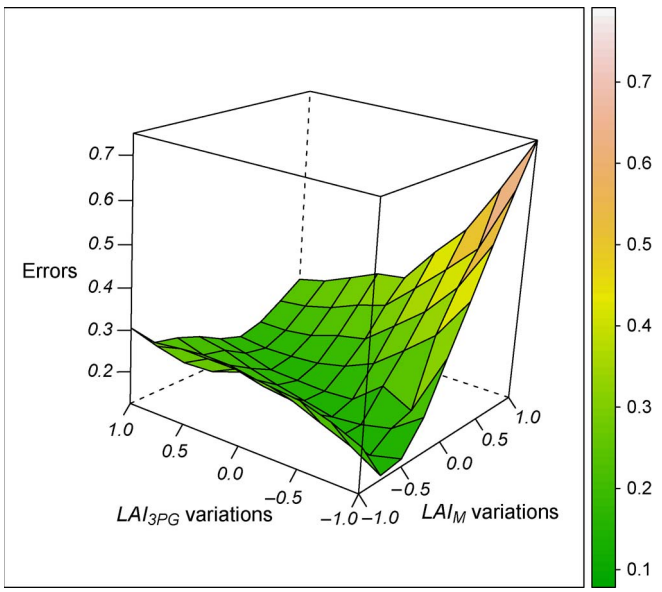

Fig. 9. Sensitivity test of the BN after simultaneously varying the $LAI_{3\mathrm{PG}}$ and LAI_M values in steps of 0.25 units within the interval $[-1, 1]$ over the studied period applied in the Speulderbos forest.

in the LAI_M than to the variations of the $LAI_{3\mathrm{PG}}$ values. Hence, for an added variation of the LAI to the MODIS LAI, the resulting error of the BN is larger than when adding the same variation to the 3-PG estimates. Furthermore, as shown in Fig. 9, large errors of the LAI are obtained when the LAI_M and the $LAI_{3\mathrm{PG}}$ are simultaneously varied. We found a maximum error of 0.74 with $LAI_M + 1$ and $LAI_{3\mathrm{PG}} - 1$.

IV. Discussion

In this paper, a methodology is presented to improve estimates of forest growth based on the manipulation of the LAI values using a BN that combines the LAI output of the biophysical 3-PG model with the LAI derived from satellite imagery. We used the BN to infer the updated LAI values from those predicted by the 3-PG model and the values reported by the modified MODIS LAI. As shown in Fig. 7, the deviation of the BN output and ground measurement is lower than the deviation between the 3-PG model output and the ground measurement, indicating that the LAI output of the BN is more accurate than that of the 3-PG model output alone and is closer to the ground measurement. The high LAI values from the 3-PG (Fig. 7) are due to an uncorrected simulation. Moreover, a sensitivity test of the 3-PG's parameters could help the BN in improving the LAI estimates. The 3-PG is executed using a specific set of species and site parameters from 1962, when the forest was planted, to 2007, when we started implementing the BN, to generate the initial conditions of the foliage, root, and stem biomass. These initial conditions were used as inputs into the 3-PG model and were combined with the MODIS images when using the BN.

The strength of the presented methodology lies in the use of two sources of information and a combination of these in a BN to improve the estimate of the LAI values. It addresses two long-standing issues in biophysical modeling and satellite image analysis. In biophysical modeling, uncertainties in the model and its parameters lead to an accumulation of error over time (multiple iterations), which give the model output a bias. This bias can be reduced by refinement and proper parameterization of the model, but often, the available data of the biophysical system are not accurate enough to refine the model or to reduce the uncertainty in the parameters. In satellite remote sensing, there are varying instrument and environmental conditions that lead to errors, which are difficult to correct unless detailed local terrain and atmosphere conditions are available. The combination of these two sources of information provides a mechanism to reduce the bias in the biophysical model by integrating a data source that has no long-term bias but an error that is evaluated in the inference mechanism of the BN.

A major contribution of this paper lies in the combination of remote-sensing images with the 3-PG model within a BN. As a strategy for the consideration of these two products in a BN node, we resorted to the mathematical formulation in (9). From this equation, we can identify the intermediate node of the BN based on the contribution of all of the MODIS images, 3-PG output, and previous BN output. However, to account for the uncertainty in both MODIS images and 3-PG model, weighing factors are introduced. This new expression also includes the BN output of the previous iteration due to the fact that the LAI values have no large changes within a relatively short period of time.

Regarding the influence of the input sources to the BN accuracy (Fig. 7), we found out that the BN is more sensitive to the variation in the LAI_M than to the variation in the $LAI_{3\mathrm{PG}}$ (Fig. 8). This indicates that LAI_M has a significant influence on the BN. Hence, the BN proposed in this paper gives more weight to LAI_M(9), but it is sensitive as well to $LAI_{3\mathrm{PG}}$ when they vary simultaneously (Fig. 9). Prices LAI_M as an input may help to get prices LAI_{BN} as an output. Improving the accuracy of the method when the input sources are less sensitive is an area for future studies.

Also, some other issues require further work. For instance, the spatial resolution of the MODIS images (250 m) is insufficient in representing the spatial variability and distribution

of trees within the forest. This brings additional uncertainty into the LAI estimation, which could be reduced by using finer resolution satellite imagery such as ASTER at 15–30-m resolution. Such spatial refinement would be particularly useful in assessing the variability in the reflectance and, thus, in the associated LAI values in smaller forest patches. Furthermore, the uncertainties in the example application (Section III-D) that may result by the neighborhood of the 16 pixels have an impact on the resulting output of the BN. This may be addressed with a subpixel remote-sensing approach. Lastly, the multisource image fusion technique could improve the BN output so that the LAI is estimated more frequently.

In addition, the seasonal changes of the LAI_{BN} (Fig. 7) are not clear and smooth, whereas the modified LAI_M shows a good agreement with the expected seasonal changes of the LAI through the year. As shown in Fig. 7, the LAI_{BN} increased rapidly in May 2008 and 2009, indicating that the LAI_{BN} during the growing season follows the modified LAI_M due to the weighing equation (9). The LAI_{BN} after a few iterations, however, drops, showing an unexpected seasonal growth. This inappropriate BN output is due to the fact that the LAI_{BN} is affected by the high LAI_{3PG} values. Apparently, the LAI_{BN} estimates reduce the uncertainties of the 3-PG output. Furthermore, the BN needs a long time series until the LAI_{BN} gets close to the LAI_{FD} and represents the changes in the seasonal growth. Fig. 7 shows that the LAI_{BN} started far from the LAI_{FD} and converged to the LAI_{FD} only after more than seven months, corresponding to 13 iterations.

Regarding the applicability of our approach, the proposed BN requires satellite imageries and field data to estimate the LAI. The remote-sensing data are often not available or of unknown quality due to atmospheric characteristics such as the presence of clouds and aerosols. Likewise, ground observations may be partly or entirely absent as some areas are difficult to reach or the instruments for field survey are too expensive for the national forest survey institutes. The use of the BNs, however, can be useful in dealing with this problem, which will be considered in further research.

In this paper, we selected a homogeneous forest area. This assumption has been verified after a significant correlation was found between the measured LAI at and around the tower $(R^2 = 0.74)$. Homogeneity, however, rarely occurs within the forests. Nonhomogeneity of the forests is difficult to address when extracting the biophysical parameters from remote-sensing images, particularly for relatively small areas. This may prohibit the extension of our method toward a more general applicability.

Spatial data are becoming increasingly important for forest vegetation management and decision making. Better estimations of the biophysical parameters provide forest managers information on forest growth, which may help in getting a better understanding of the forest and which ultimately can serve as an informative factor in climate change. The forest canopy data (e.g., vegetation parameters like LAI) play a major role in the simulation of the surface energy balance and, therefore, weather and climate prediction. It may also be useful for the forest scientists and the Intergovernmental Panel on Climate Change to apply it in directly or indirectly assessing the carbon stored in the forest, which is one of the important current issues in climate change mitigation and adaptation strategies.

V. Conclusion

This paper has presented a BN for improving the LAI estimation by combining the 3-PG output with that derived from remote-sensing imagery. Some concepts of the BN are introduced and summarized. A new equation (9) is defined according to the relative error of the MODIS images and the 3-PG output. We have illustrated the framework by implementing the BN for the Speulderbos forest in The Netherlands. The study leads to the following conclusions.

1) A BN is able to integrate the LAI values that come from different sources into a single reasoning framework.
2) By using the satellite images, the output of the 3-PG forest growth model is improved. It closely matches the mean forest growth estimate, and it substantially reduces both the RMSE and the RE.

References

[1] J. J. Landsberg and R. H. Waring, "A generalized model of forest productivity using simplified concepts of radiation-use efficiency, carbon balance and partitioning," *For. Ecol. Manage.*, vol. 95, no. 3, pp. 209–228, Aug. 1997.

[2] G. M. J. Mohren and J. R. Van de Veen, "Forest growth in relation to site conditions. Application of the model FORGRO to the solling spruce site," *Ecol. Model.*, vol. 83, no. 1/2, pp. 173–183, Dec. 1995.

[3] E. Schulze, "Plant life forms and their carbon, water and nutrient relations," in *Encyclopedia of Plant Physiology*, vol. 12B, O. L. Lange, C. B. Osmond, P. S. Nobel, and H. Ziegler, Eds. Berlin, Germany: Springer-Verlag, 1982, pp. 615–676.

[4] I. Jonckheere, S. Fleck, K. Nackaerts, B. Muys, P. Coppin, M. Weiss, and F. Baret, "Review of methods for in situ leaf area index determination: Part I. Theories, sensors and hemispherical photography," *Agric. For. Meteorol.*, vol. 121, no. 1/2, pp. 19–35, Jan. 2004.

[5] D. C. Mummery and M. Battaglia, "Significance of rainfall distribution in predicting eucalypt plantation growth, management options, and risk assessment using the process-based model CABALA," *For. Ecol. Manage.*, vol. 193, no. 1/2, pp. 283–296, May 2004.

[6] L. J. Esprey, P. J. Sands, and C. W. Smith, "Understanding 3-PG using a sensitivity analysis," *For. Ecol. Manage.*, vol. 193, no. 1/2, pp. 235–250, May 2004.

[7] L. Fontes, J. Landsberg, J. Tomé, M. Tomé, C. A. Pacheco, P. Soares, and C. Araujo, "Calibration and testing of a generalized process-based model for use in Portuguese *Eucalyptus* plantations," *Can. J. For. Res.*, vol. 36, no. 12, pp. 3209–3221, Dec. 2006.

[8] W. Z. Yang, B. Tan, D. Huang, M. Rautiainen, N. V. Shabanov, Y. Wang, J. L. Privette, K. F. Huemmrich, R. Fensholt, I. Sandholt, M. Weiss, D. E. Ahl, S. T. Gower, R. R. Nemani, Y. Knyazikhin, and R. B. Myneni, "MODIS leaf area index products: From validation to algorithm improvement," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 7, pp. 1885–1898, Jul. 2006.

[9] W. Yang, D. Huang, B. Tan, J. C. Stroeve, N. V. Shabanov, Y. Knyazikhin, R. R. Nemani, and R. B. Myneni, "Analysis of leaf area index and fraction of PAR absorbed by vegetation products from the terra MODIS sensor: 2000–2005," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 7, pp. 1829–1842, Jul. 2006.

[10] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Mateo, CA: Morgan Kaufmann, 1988.

[11] S. Imoto, T. Higuchi, T. Goto, K. Tashiro, S. Kuhara, and S. Miyano, "Combining microarrays and biological knowledge for estimating gene networks via Bayesian networks," in *Proc. IEEE Comput. Soc. Bioinformatics Conf.*, 2003, pp. 104–113.

[12] C. J. Needham, J. R. Bradford, A. J. Bulpitt, and D. R. Westhead, "A primer on learning in Bayesian networks for computational biology," *PLoS Comput. Biol.*, vol. 3, no. 8, pp. 1409–1416, Aug. 2007.

[13] A. V. Werhli and D. Husmeier, "Reconstructing gene regulatory networks with Bayesian networks by combining expression data with multiple sources of prior knowledge," *Stat. Appl. Genet. Mol. Biol.*, vol. 6, no. 1, pp. 1–45, May 2007.

[14] F. Taroni, C. Aitken, P. Garbolino, and A. Biedermann, *Bayesian Networks and Probabilistic Inference in Forensic Science*. Hoboken, NJ: Wiley, Feb. 2006, ser. Statistics in Practice.

[15] O. Pourret, P. Naïm, and B. Marcot, *Bayesian Networks: A Practical Guide to Applications*. Hoboken, NJ: Wiley, 2008, ser. Statistics in Practice.
[16] P. J. Sands and J. J. Landsberg, "Parameterisation of 3-PG for plantation grown *Eucalyptus* globulus," *For. Ecol. Manage.*, vol. 163, no. 1–3, pp. 273–292, Jun. 2002.
[17] A. Almeida, J. Landsberg, and P. Sands, "Parameterisation of 3-PG model for fast-growing *Eucalyptus* grandis plantations," *For. Ecol. Manage.*, vol. 193, no. 1/2, pp. 179–195, May 2004a.
[18] M. Kalacska, A. Sanchez-Azofeifa, T. Caelli, B. Rivard, and B. Boerlage, "Estimating leaf area index from satellite imagery using Bayesian networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 8, pp. 1866–1873, Aug. 2005.
[19] P. Yang, R. Shibasaki, W. Wu, Q. Zhou, Z. Chen, Y. Zha, Y. Shi, and H. Tang, "Evaluation of MODIS land cover and LAI products in cropland of North China Plain using in situ measurements and Landsat TM images," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 10, pp. 3087–3097, Oct. 2007.
[20] J. T. Morisette, F. Baret, J. L. Privette, R. B. Myneni, J. E. Nickeson, S. Garrigues, N. V. Shabanov, M. Weiss, R. A. Fernandes, S. G. Leblanc, M. Kalacska, G. A. Sanchez-Azofeifa, M. Chubey, B. Rivard, P. Stenberg, M. Rautiainen, P. Voipio, T. Manninen, A. N. Pilant, T. E. Lewis, J. S. Iiames, R. Colombo, M. Meroni, L. Busetto, W. B. Cohen, D. P. Turner, E. D. Warner, G. W. Petersen, G. Seufert, and R. Cook, "Validation of global moderate-resolution LAI products: A framework proposed within the CEOS land product validation subgroup," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 7, pp. 1804–1817, Jul. 2006.
[21] B. He, "A simple data assimilation method for improving the MODIS LAI time-series data products based on the object analysis and gradient inverse weighted filter," *Chin. Opt. Lett.*, vol. 5, no. 6, pp. 367–369, Jun. 2007.
[22] J. Qin, S. L. Liang, X. W. Li, and J. D. Wang, "Development of the adjoint model of a canopy radiative transfer model for sensitivity study and inversion of leaf area index," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 7, pp. 2028–2037, Jul. 2008.
[23] Z. Q. Xiao, S. L. Liang, J. D. Wang, J. L. Song, and X. Y. Wu, "A temporally integrated inversion method for estimating leaf area index from MODIS data," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 8, pp. 2536–2545, Aug. 2009.
[24] F. V. Jensen and T. D. Nielsen, *Bayesian Networks and Decision Graphs*, 2nd ed. New York: Springer-Verlag, 2007, ser. Information Science and Statistics.
[25] D. Geiger and D. Heckerman, "Learning Gaussian networks," in *Proc. 10th Annu. Conf. Uncertainty Artif. Intell.*, R. L. de Mántaras and D. Poole, Eds., 1994, pp. 235–243.
[26] R. D. Shachter and C. R. Kenley, "Gaussian influence diagrams," *Manage. Sci.*, vol. 35, no. 5, pp. 527–550, May 1989.
[27] S. Sullivant, "Algebraic geometry of Gaussian Bayesian networks," *Adv. Appl. Math.*, vol. 40, no. 4, pp. 482–513, May 2008.
[28] E. Castillo and U. Kjaerulff, "Sensitivity analysis in Gaussian Bayesian networks using a symbolic-numerical technique," *Reliab. Eng. Syst. Saf.*, vol. 79, no. 2, pp. 139–148, Feb. 2003.
[29] A. Verger, F. Baret, and M. Weiss, "Performances of neural networks for deriving LAI estimates from existing CYCLOPES and MODIS products," *Remote Sens. Environ.*, vol. 112, no. 6, pp. 2789–2803, Jun. 2008.
[30] Z. Su, W. J. Timmermans, C. van der Tol, R. Dost, R. Bianchi, J. A. Gomez, A. House, I. Hajnsek, M. Menenti, V. Magliulo, M. Esposito, R. Haarbrink, F. Bosveld, R. Rothe, H. K. Baltink, Z. Vekerdy, J. A. Sobrino, J. Timmermans, P. van Laake, S. Salama, H. van der Kwast, E. Claassen, A. Stolk, L. Jia, E. Moors, O. Hartogensis, and A. Gillespie, "EAGLE 2006-Multi-purpose, multi-angle and multi-sensor in-situ and airborne campaigns over grassland and forest," *Hydrology and Earth System Sciences*, vol. 13, no. 6, pp. 833–845, 2009.
[31] G. Lemoine, D. Hoekman, and H. van Leeuwen, "Separability of agricultural fields and forest stands using multi-frequency polarimetric SAR data of the Flevoland site," in *Proc. IGARSS. Remote Sensing: Global Monit. Earth Manage.*, Jun. 3–6, 1991, vol. 2, pp. 681–684.
[32] E. Steingröver and W. Jans, "Physiology of forest-grown Douglas-fir trees, effects of air pollution and drought," Wageningen Univ. Res. Centre, Wageningen, The Netherlands, 1995.
[33] M. T. Van Wijk, S. C. Dekker, W. Bouten, W. Kohsiek, and G. M. Mohren, "Simulation of carbon and water budgets of a Douglas-fir forest," *For. Ecol. Manage.*, vol. 145, no. 3, pp. 229–241, May 2001.
[34] J. M. Maass, J. M. Vose, W. T. Swank, and A. Martinez-Yrizar, "Seasonal changes of leaf area index (LAI) in a tropical deciduous forest in West Mexico," *For. Ecol. Manage.*, vol. 74, no. 1–3, pp. 171–180, Jun. 1995.
[35] N. C. Coops, M. L. Smith, K. L. Jacobsen, M. Martin, and S. Ollinger, "Estimation of plant and leaf area index using three techniques in a mature native *Eucalypt* canopy," *Austral Ecol.*, vol. 29, no. 3, pp. 332–341, Jun. 2004.
[36] K. R. Whitford, I. J. Colquhoun, A. R. G. Lang, and B. M. Harper, "Measuring leaf area index in a sparse *Eucalypt* forest: A comparison of estimates from direct measurement, hemispherical photography, sunlight transmittance and allometric regression," *Agric. For. Meteorol.*, vol. 74, no. 3/4, pp. 237–249, May 1995.
[37] R. B. Myneni, F. G. Hall, P. J. Sellers, and A. L. Marshak, "The interpretation of spectral vegetation indexes," *IEEE Trans. Geosci. Remote Sens.*, vol. 33, no. 2, pp. 481–486, Mar. 1995.
[38] R. B. Myneni, S. Hoffman, Y. Knyazikhin, J. L. Privette, J. Glassy, Y. Tian, Y. Wang, X. Song, Y. Zhang, G. R. Smith, A. Lotsch, M. Friedl, J. T. Morisette, P. Votava, R. R. Nemani, and S. W. Running, "Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data," *Remote Sens. Environ.*, vol. 83, no. 1/2, pp. 214–231, Nov. 2002.
[39] S. D. Prince and S. N. Goward, "Global primary production: A remote sensing approach," *J. Biogeogr.*, vol. 22, no. 4/5, pp. 815–835, Jul. 1995.
[40] D. C. Steinberg and S. Goetz, "Assessment and extension of the MODIS FPAR products in temperate forests of the eastern United States," *Int. J. Remote Sens.*, vol. 30, no. 1, pp. 169–187, Jan. 2009.
[41] J. M. Norman, W. P. Kustas, and K. S. Humes, "Source approach for estimating soil and vegetation energy fluxes in observations of directional radiometric surface temperature," *Agric. For. Meteorol.*, vol. 80, no. 2–4, p. 297, Jul. 1996.

Yaseen T. Mustafa (M'10) received the M.Sc. degree in mathematics science from the University of Duhok, Duhok, Iraq. He is currently working toward the Ph.D. degree in the Faculty of Geo-Information Science and Earth Observation, University of Twente, Enschede, The Netherlands.

Patrick E. Van Laake received the M.Sc. degree in agricultural engineering from the Wageningen Agricultural University, Wageningen, The Netherlands, and the Ph.D. degree in earth and atmospheric sciences from the University of Alberta, Edmonton, AB, Canada.

He was an Assistant Professor in geo-information for sustainable forest management with the International Institute for Geo-Information Science and Earth Observation, Enschede, The Netherlands. He is currently with the United Nations Development Programme, Ha Noi, Vietnam, working on sustainable forest management and climate change. His main research interest is in the biophysical modeling of tropical forests and carbon balance assessment under scenarios of forest degradation.

Alfred Stein received the M.S. degree in mathematics and information science, with a specialization in applied statistics, from the Eindhoven University of Technology, Eindhoven, The Netherlands, and the Ph.D. degree in agricultural and environmental science from the Wageningen Agricultural University, Wageningen, The Netherlands.

He is currently a Professor of mathematical and statistical methods for geodata with the Faculty of Geo-Information Science and Earth Observation, University of Twente, Enschede, The Netherlands, where he heads the Department of Earth Observation Science and where he is also a Professor of stochastic image analysis. He is the Chief Editor of the *International Journal of Applied Earth Observation and Geoinformation*. His main interest is in spatial statistics and image mining, including spatial data quality. Applications emerge from a range of agricultural, urban, and environmental fields.