Proceedings of the 7th International Conference on Functional-Structural Plant Models, Saariselkä, Finland, 9 - 14 June 2013. Eds. Risto Sievänen, Eero Nikinmaa, Christophe Godin, Anna Lintunen & Pekka Nygren. http://www.metla.fi/fspm2013/proceedings. ISBN 978-951-651-408-9.

Bayes trees and forests: combining precise empirical and theoretical tree models

Mikko Kaasalainen[1*], Ilya Potapov[1], Pasi Raumonen[1], Markku Åkerblom[1], Risto Sievänen[2] and Sanna Kaasalainen[3]

[1]*Department of Mathematics, Tampere University of Technology, P.O. Box 553, 33101, Tampere, Finland*
[2]*Finnish Forest Research Institute, Vantaa Research Unit, PL 18, FI-01301, Vantaa, Finland*
[3]*Department of Remote Sensing and Photogrammetry, Finnish Geodetic Institute, Geodeetinrinne 2, FI-02431 Masala, Finland*
**correspondence:* mikko.kaasalainen [at] tut.fi

Highlights: With the new analysis methods for TLS scans, there will be a growing and improving database of 3D descriptions of trees and forest stands. The attributes determining these descriptions can be represented as Bayesian probability distributions, with functional-structural models providing the prior information. These distributions can then be used to create versions of new realistic *Bayes forests*, where none of the trees are copied from data, but the structure of each is drawn from the data-based distributions. Repeated TLS measurements add a fourth dimension, time, to the mathematical modelling; in this way, we can simulate functional *4D Bayes forests*. As in the modelling of the 3D structure, forest models and regularities of growth and mortality are used as prior information; conversely, the accumulating data and modelling results improve the theoretical models.

Keywords: Tree models: empirical, tree models: theoretical, TLS, Bayesian inference

INTRODUCTION

A general method, based on terrestrial laser scanning (TLS) data, has recently been developed for producing precise 3D tree surface models that are automatic and fast to compute and record the topological and geometric properties of the tree (Raumonen et al. 2013 and this meeting). One of the core ideas behind the method is that practically any external attribute of a tree can be approximated accurately at will from a compact surface model of this type.

Once we have determined the models of a representative number of trees at different sites, we can construct well-defined statistical attributes from these as functions of species, age, etc. On forest stand level, we record the typical distribution of trees. Once the distribution functions (DFs) of tapering, branching frequency, branching angles, branch curving, tree positions, etc. have been defined, we can carry out a reversed process. Now we can draw new samples from these DFs and use them to construct new trees and forests with similar statistics; i.e., *3D and 4D Bayes forests*. In this process, we can use our prior knowledge of functional-structural (FS) tree models to ensure biological consistency. This can be carried out via the Bayesian approach to obtain a posterior distribution. The process can also be used to improve the theoretical models by detailed, precise, and comprehensive measurements.

PROBABILISTIC TREE MODELS

In the broadest sense, a tree (or, ultimately, a forest) can be described as a probabilistic entity. Its form and structure, specified by some parameters $\boldsymbol{x}=(x_1,...,x_N)$, has a probability $p(\boldsymbol{x};\boldsymbol{u})$ in some measurement space spanned by $\boldsymbol{u}$. Some components of $\boldsymbol{u}$ might describe, e.g., the local width h of a branch and its 3D direction $\boldsymbol{d}$ along its length s, resulting in a four-dimensional subspace $(h,\boldsymbol{d},s)$ of the $\boldsymbol{u}$-space. From a large number of scanned and analyzed trees, we can determine an experimental $p_{\mathrm{S}}(\boldsymbol{x})$ for a given species (possibly with different p-functions for different age groups or other identifiers, or by including them as components of $\boldsymbol{x}$). On the other hand, we can determine a theoretical $p_{\mathrm{M}}(\boldsymbol{x})$ from a large ensemble of functional-structural models. In our representation, two trees that are not clones of each other but have the same p (i.e., $\boldsymbol{x}$) are identical in the mathematical sense. We note here that, formally, two isolated stochastic FS models that have the same initial parameters do not necessarily produce identical 3D DFs at later stages due to self-shadowing that takes different evolutionary routes depending on the random choices during growth. In practice, identical environmental factors mostly produce statistically similar trees of the same species.

The tree probability $p(\boldsymbol{x})$ can comprise a number of sub-probabilities. A probability may be a product of independent, lower-dimensional probability functions. These can be purely morphological; e.g., of the form $p(\{x_i\};\boldsymbol{d},h,s)=p[\{x_i\};\{\boldsymbol{d}_i(s)\},\{h_i(s)\},s]$, where $\{\boldsymbol{d}_i(s)\}$ and $\{h_i(s)\}$ are sets of some basis functions modified by the set of parameters $\{x_i\}$. On the other hand, the sub-probabilities can describe stochastic processes, which is the natural way of describing the growth of a tree. The sub-probabilities are then Markovian in character, specifying the probability of a length element to differ from the previous one in various ways (with either independent or joint attributes; i.e., one- or multidimensional probabilities). The elements are similar in both the observational and theoretical models; e.g., cylindrical in our basic laser-scanning model and the LIGNUM FS model (Sievänen et al. 2008). For example, $\Delta\boldsymbol{d}$ can be taken to describe the difference between $\boldsymbol{d}$ of two adjacent elements. Then the probability distribution of this difference would be given by a $p(\Delta\boldsymbol{d})$. A model for p linear in some parameters $(x_i,...,x_j)$ is of the form

$$p(x_i,\dots,x_j;\Delta\boldsymbol{d})=\sum_{k=i}^{j}x_k f_k(\Delta\boldsymbol{d}),$$

where $f_k(\Delta\boldsymbol{d})$ are some given basis functions. Another linear possibility is to discretize the $\Delta\boldsymbol{d}$-space into bins whose occupation probabilities are given by x_k.

The problem of determining the probabilities or statistical profiles $p(\boldsymbol{x})$ from either empirical data or synthetic models is that of determing the underlying distribution function from a set of samples (Kaasalainen 2008). This can be carried out, in the sense of the Radon transform, by least-squares fitting the parameters $\boldsymbol{x}$ such that the cumulative marginal distributions (CDFs) of $p(\boldsymbol{x};\boldsymbol{u})$ in various directions in the measurement space of $\boldsymbol{u}$ best match the corresponding CDF values $\boldsymbol{y}$ defined by the given set of sampled $\boldsymbol{u}$. In the case of linear models such as the $p(\Delta\boldsymbol{d})$ above, this is a linear least-squares problem of the matrix form $\boldsymbol{y}=A\boldsymbol{x}$.

Once we have determined a DF, we can resample it to create statistically equivalent new trees at will (Fig. 1). This is simple for the one-dimensional case via its cumulative DF; for multi-dimensional DFs we can use sprinkle algorithms (Kaasalainen 2008) or Markov Chain Monte Carlo methods.

BAYESIAN MODELS

Next, we introduce the concept of a Bayesian tree (or forest). To make the experimental and theoretical tree statistics p_S and p_M compatible in the Bayesian sense, we write a probability distribution that can smoothly incorporate any given degree of prior information from p_M. The desired posterior distribution can be obtained by a metaprobabilistic approach: we define the probability P of a tree to have a certain kind of statistical profile p represented by the parameters $\boldsymbol{x}$. Thus, the Bayesian principle yields

$$P(\boldsymbol{x})\propto P_S(\boldsymbol{x})P_M(\boldsymbol{x})$$

(without normalization as we are only interested in the shape of $P(\boldsymbol{x})$ in $\boldsymbol{x}$-space). The role of $P_M(\boldsymbol{x})$ is defined by some "tightening" or weight parameters that can be separate for each x_i. For zero weights, $P_M(\boldsymbol{x})$ is unity; i.e., a uniform distribution without any prior preferences for $\boldsymbol{x}$. For weights approaching infinity, $P_M(\boldsymbol{x})$ approaches the Dirac-delta distribution around its centre point $\boldsymbol{x}_0$.

If the models describing the DFs are linear in the parameters $\boldsymbol{x}$ in the form $\boldsymbol{y}=A\boldsymbol{x}$ above, we obtain a simple formula for the centre point or the maximum a posteriori estimate $\hat{\boldsymbol{x}}$ of the final distribution $P(\boldsymbol{x})$, if it as well as the distributions $P_M(\boldsymbol{x})$ and $P_S(\boldsymbol{x})$ are taken to be Gaussian:

$$\hat{\boldsymbol{x}}=Q^{-1}(E_0^{-1}\boldsymbol{x}_0+A^T E_1^{-1}\boldsymbol{y});\qquad Q=(E_0^{-1}+A^T E_1^{-1}A),$$

where E_0 and E_1 are, respectively, the covariance matrices of $P_M(\boldsymbol{x})$ and $P_S(\boldsymbol{x})$. The smaller their (diagonal) elements are, the more tightly the marginal probabilities of the components of $\boldsymbol{x}$ are constrained.

We can describe the probabilities in both 3D and 4D. In the 3D-case, we consider a snapshot of a tree, specifying its momentary form. The straightforward DF product approach above is then simplest to apply to essentially isolated trees. In the time-dependent 4D-case, we examine the history of trees; i.e., the shape probability as a function of time. The probabilities are then determined by a number of stochastic as well as deterministic processes. This is necessary especially for Bayes forests, where the collective history (competition etc.) should be taken into account. The growth rules of FS models, for example, are then rendered as prior DFs. The theoretical FS models not only serve as prior probabilities in the Bayesian sense; they also help to design the parameter space in which we describe the probabilities. To create prior DFs from

FS models, we can use stochastic FS processes and deterministic FS models with stochastic conditions initially and during the growth.

The Bayesian approach stabilizes the synthetic realizations of TLS-based DFs by, e.g., removing unrealistic outliers or biases due to selection effects in data. The FS models are also used to determine the best ways to select and express the actual tree attributes we should use in our virtual modelling. On the other hand, experimental data help to improve theoretical models.

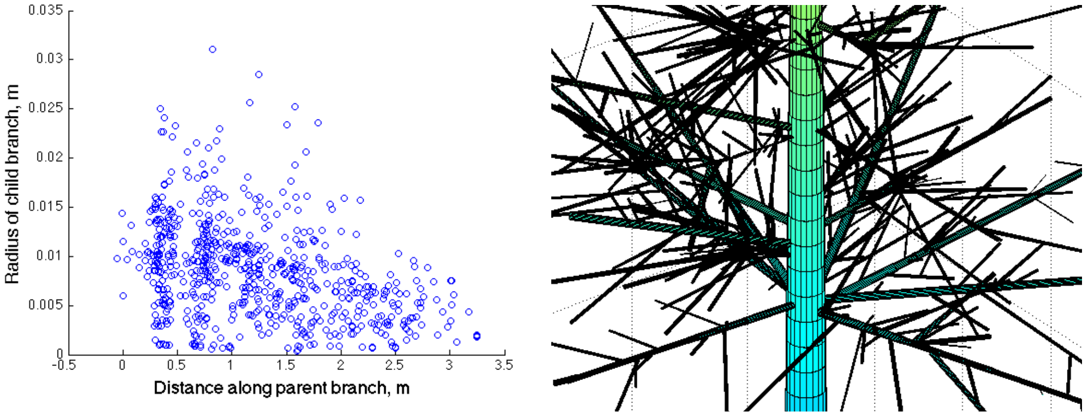

Figure 1. Left: a 2D-plot of the measured second-order branch statistics of one tree. Right: a closeup of the basic geometry (location, direction, and width) of synthetic branch generation by drawing samples from DFs corresponding to the data.

MEASUREMENTS

The development of a Bayes forest model requires continuous assimilation of measured data to the system. We measure and collect new TLS and other information from forests, including hyperspectral lidar (HSL) data for augmenting the structure data by information on source material and condition (Hakala et al. 2012). The sample sites and trees are chosen to acquire representative quantities, and to test a number of statistical hypotheses. For example, to which degree do the parameters $\boldsymbol{x}$ of presumably similar trees correlate? What are the typical deviations given by the experimental covariance matrix E_1? How much prior information needs to be introduced in practice? FS models and observations are used to determine the most appropriate parameter space for $\boldsymbol{x}$ and the measurement space for $\boldsymbol{u}$ best representing the various attributes of trees, including branch hierarchies. The preliminary results are reported in this meeting.

In 4D Bayes forest modelling, the DFs of the 3D Bayes forest have the additional dimension of time, based on the repeated observations of sample trees. Consequently, the realizations of these DFs for each Bayes forest include behaviour in time: litter production, growth, etc. As in the 3D Bayes forest, the prior constraints from the process-based studies (incorporated in LIGNUM) are important. A number of process descriptions are used as prior information sources taking into account the effects of, e.g., the competition for light, space, and resources. This information can be incorporated into the Bayes-forest model as, e.g., spatial competition indices and simple carbon-balance rules of foliage. With the biologically consistent prior components, the end product is a statistically and biologically realistic model of a forest and its processes in time.

LITERATURE CITED

Hakala T, Suomalainen J, Kaasalainen S, Chen Y. 2012. Full Waveform Hyperspectral LiDAR for Terrestrial Laser Scanning. *Optics Express.* **20:** 7119-7127.

Kaasalainen M. 2008. Dynamical tomography of gravitationally bound systems. *Inverse Problems and Imaging.* **2**: 527-549.

Raumonen P, Kaasalainen M, Åkerblom M, Kaasalainen S, Kaartinen H, Vastaranta M, Holopainen M, Disney M, Lewis P. 2013. Fast Automatic Precision Tree Models from Terrestrial Laser Scanner Data. *Remote Sensing*. **5**: 491-520.

Sievänen R, Perttunen J, Nikinmaa E, Kaitaniemi P. 2008. Toward extension of a single tree functional structural model of Scots pine to stand level: effect of the canopy of randomly distributed, identical trees on development of tree structure. *Functional Plant Biology*. **35**: 964-975.