# Estimation of stand density using aerial images: a forestry application of the Boolean model

Lauri Mehtätalo[1] and Jari Vauhkonen[2]

[1]School of Computing and [2]School of Forest Sciences
University of Eastern Finland, Joensuu campus

6.6.2014 / NordStat2014, Turku.

# Outline

# Motivation

- Aircraft-mounted laser scanners (ALS) are used to collect three-dimensional data on forest canopies.
- We consider individual tree detection approaches for such data (c.f. Virpi's presentation on an area-based approach)
  - → Crown areas for those trees that were detectable on the image.
  - → The proportion of the image area coverted by tree crowns (canopy closure, *cc*).
- Small trees below dominant trees remain undetected. The consequent detectability of a tree on the ALS image is a function of its size.
- We present and evaluate model-based estimators for stand density and detectability in this context.

# Motivation

- Aircraft-mounted laser scanners (ALS) are used to collect three-dimensional data on forest canopies.
- We consider individual tree detection approaches for such data (c.f. Virpi's presentation on an area-based approach)
  - → Crown areas for those trees that were detectable on the image.
  - → The proportion of the image area coverted by tree crowns (canopy closure, *cc*).
- Small trees below dominant trees remain undetected. The consequent detectability of a tree on the ALS image is a function of its size.
- We present and evaluate model-based estimators for stand density and detectability in this context.

# Motivation

- Aircraft-mounted laser scanners (ALS) are used to collect three-dimensional data on forest canopies.
- We consider individual tree detection approaches for such data (c.f. Virpi's presentation on an area-based approach)
  - → Crown areas for those trees that were detectable on the image.
  - → The proportion of the image area coverted by tree crowns (canopy closure, *cc*).
- Small trees below dominant trees remain undetected. The consequent detectability of a tree on the ALS image is a function of its size.
- We present and evaluate model-based estimators for stand density and detectability in this context.

# Motivation

- Aircraft-mounted laser scanners (ALS) are used to collect three-dimensional data on forest canopies.
- We consider individual tree detection approaches for such data (c.f. Virpi's presentation on an area-based approach)
  - → Crown areas for those trees that were detectable on the image.
  - → The proportion of the image area coverted by tree crowns (canopy closure, *cc*).
- Small trees below dominant trees remain undetected. The consequent detectability of a tree on the ALS image is a function of its size.
- We present and evaluate model-based estimators for stand density and detectability in this context.

# The Boolean model in $R^2$

Consider a marked point process

$$\left\{\mathbb{Z}(s_i) : s_i \in \mathbb{R}^2\right\},$$

where

- the locations (germs) $s_i$ are generated by the homogeneous (=stationary, isotropic) Poisson process, with intensity $\lambda$.
  - $\rightarrow$ For any region $B$, the number of points follows the *Poisson*$(\lambda\,|B|)$ distribution and $s_i$'s are scattered independently and unformly.
- The marks (primary grains) $\mathbb{Z}(s_i)$ are independent realizations of random compact sets (RACS) in $\mathbb{R}^2$, with areas $Z$ with pdf $f(z)$.
- The union of the primary grains $\bigcup_{i=1}^{\infty} \mathbb{Z}_i$ is the Boolean model, also known as Poisson germ-grain model.

# The Boolean model in $R^2$

Consider a marked point process

$$\left\{ \mathbb{Z}(s_i) : s_i \in \mathbb{R}^2 \right\} ,$$

where

- the locations (germs) $s_i$ are generated by the homogeneous (=stationary, isotropic) Poisson process, with intensity $\lambda$.
  - $\rightarrow$ For any region $B$, the number of points follows the *Poisson*$(\lambda |B|)$ distribution and $s_i$'s are scattered independently and unformly.
- The marks (primary grains) $\mathbb{Z}(s_i)$ are independent realizations of random compact sets (RACS) in $\mathbb{R}^2$, with areas $Z$ with pdf $f(z)$.
- The union of the primary grains $\bigcup_{i=1}^{\infty} \mathbb{Z}_i$ is the Boolean model, also known as Poisson germ-grain model.

# The Boolean model in $R^2$

Consider a marked point process

$$\left\{\mathbb{Z}(s_i) : s_i \in \mathbb{R}^2\right\},$$

where

- the locations (germs) $s_i$ are generated by the homogeneous (=stationary, isotropic) Poisson process, with intensity $\lambda$.
  - $\rightarrow$ For any region $B$, the number of points follows the *Poisson*$(\lambda\,|B|)$ distribution and $s_i$'s are scattered independently and unformly.
- The marks (primary grains) $\mathbb{Z}(s_i)$ are independent realizations of random compact sets (RACS) in $\mathbb{R}^2$, with areas $Z$ with pdf $f(z)$.
- The union of the primary grains $\bigcup_{i=1}^{\infty} \mathbb{Z}_i$ is the Boolean model, also known as Poisson germ-grain model.

# An example

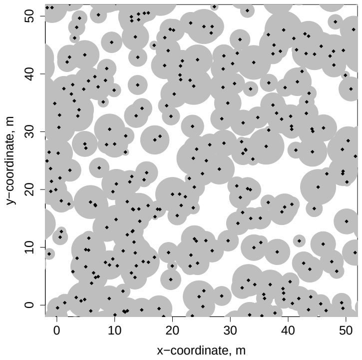



Figure : A realization of the Boolean model with circular primary grains.

# Estimator for stand density

- The area fraction
$$p = 1 - e^{-\lambda \mathrm{E}(Z)}$$
gives the mean fraction occupied by the model in a region of unit area (also called hitting probability or vacancy)
- In our application, the canopy closure provides a measurement of the volume fraction. Setting $cc = p$ yields an estimator for stand density
$$\hat{\lambda} = -\frac{\ln(1 - cc)}{\mathrm{E}(Z)} \quad (1)$$
where $\mathrm{E}(Z)$ is the mean crown area (over all trees, not only over the observed ones) and $cc \in [0, 1]$ is the relative canopy closure.
- To compute $\mathrm{E}(Z)$, we need an expression of detectability as a function of crown area, $\pi(z)$.

# Estimator for stand density

- The area fraction

$$p = 1 - e^{-\lambda \mathrm{E}(Z)}$$

gives the mean fraction occupied by the model in a region of unit area (also called hitting probability or vacancy)

- In our application, the canopy closure provides a measurement of the volume fraction. Setting $cc = p$ yields an estimator for stand density

$$\hat{\lambda} = -\frac{\ln(1 - cc)}{\mathrm{E}(Z)} \quad (1)$$

where $\mathrm{E}(Z)$ is the mean crown area (over all trees, not only over the observed ones) and $cc \in [0, 1]$ is the relative canopy closure.

- To compute $\mathrm{E}(Z)$, we need an expression of detectability as a function of crown area, $\pi(z)$.

# Estimator for stand density

- The area fraction
$$p = 1 - e^{-\lambda \mathrm{E}(Z)}$$
gives the mean fraction occupied by the model in a region of unit area (also called hitting probability or vacancy)
- In our application, the canopy closure provides a measurement of the volume fraction. Setting $cc = p$ yields an estimator for stand density
$$\hat{\lambda} = -\frac{\ln(1 - cc)}{\mathrm{E}(Z)} \quad (1)$$
where $\mathrm{E}(Z)$ is the mean crown area (over all trees, not only over the observed ones) and $cc \in [0, 1]$ is the relative canopy closure.
- To compute $\mathrm{E}(Z)$, we need an expression of detectability as a function of crown area, $\pi(z)$.

# Detectability $\pi(z)$

- Consider trees with crown area above a fixed $z$. Due to independence of $Z$, the locations of trees with $Z > z$ is realization of a homogeneous Poisson process with intensity

$$\lambda_{Z>z} = \lambda P(Z > z).$$

- The mean crown area for these trees is:

$$\mathrm{E}_{Z>z}(Z) = \frac{1}{P(Z > z)} \int_z^\infty t f(t) \mathrm{d}t.$$

- The trees with $Z > z$ form also a Boolean model, with area fraction

$$p_{Z>z} = 1 - e^{-\lambda_{Z>z} \mathrm{E}_{Z>z}(Z)} = 1 - e^{-\lambda \int_z^\infty t f(t) \mathrm{d}t}$$

- We assume that a tree remains undetected if the center is within the crown of a bigger tree. Then the detectability is directly

$$\pi(z; \lambda) = 1 - p_{Z>z} = e^{-\lambda \int_z^\infty t f(t) \mathrm{d}t}$$

# Detectability $\pi(z)$

- Consider trees with crown area above a fixed $z$. Due to independence of $Z$, the locations of trees with $Z > z$ is realization of a homogeneous Poisson process with intensity

$$\lambda_{Z>z} = \lambda P(Z > z).$$

- The mean crown area for these trees is:

$$\mathrm{E}_{Z>z}(Z) = \frac{1}{P(Z > z)} \int_z^{\infty} tf(t)\mathrm{d}t.$$

- The trees with $Z > z$ form also a Boolean model, with area fraction

$$p_{Z>z} = 1 - e^{-\lambda_{Z>z}\mathrm{E}_{Z>z}(Z)} = 1 - e^{-\lambda \int_z^{\infty} tf(t)\mathrm{d}t}$$

- We assume that a tree remains undetected if the center is within the crown of a bigger tree. Then the detectability is directly

$$\pi(z;\lambda) = 1 - p_{Z>z} = e^{-\lambda \int_z^{\infty} tf(t)\mathrm{d}t}$$

# Detectability $\pi(z)$

- Consider trees with crown area above a fixed $z$. Due to independence of $Z$, the locations of trees with $Z > z$ is realization of a homogeneous Poisson process with intensity

$$\lambda_{Z>z} = \lambda P(Z > z).$$

- The mean crown area for these trees is:

$$\mathrm{E}_{Z>z}(Z) = \frac{1}{P(Z > z)} \int_z^\infty tf(t)\mathrm{d}t.$$

- The trees with $Z > z$ form also a Boolean model, with area fraction

$$p_{Z>z} = 1 - e^{-\lambda_{Z>z}\mathrm{E}_{Z>z}(Z)} = 1 - e^{-\lambda \int_z^\infty tf(t)\mathrm{d}t}$$

- We assume that a tree remains undetected if the center is within the crown of a bigger tree. Then the detectability is directly

$$\pi(z;\lambda) = 1 - p_{Z>z} = e^{-\lambda \int_z^\infty tf(t)\mathrm{d}t}$$

# Detectability $\pi(z)$

- Consider trees with crown area above a fixed $z$. Due to independence of $Z$, the locations of trees with $Z > z$ is realization of a homogeneous Poisson process with intensity

$$\lambda_{Z>z} = \lambda P(Z > z).$$

- The mean crown area for these trees is:

$$\mathrm{E}_{Z>z}(Z) = \frac{1}{P(Z > z)} \int_z^\infty tf(t)\mathrm{d}t.$$

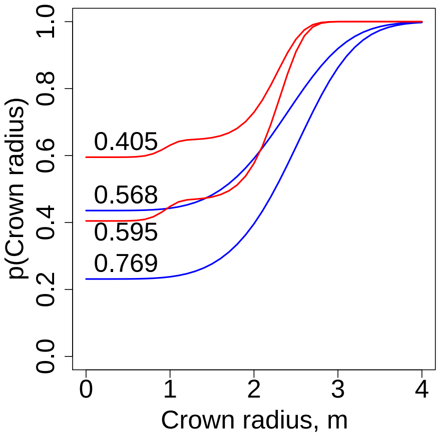
- The trees with $Z > z$ form also a Boolean model, with area fraction

$$p_{Z>z} = 1 - e^{-\lambda_{Z>z}\mathrm{E}_{Z>z}(Z)} = 1 - e^{-\lambda \int_z^\infty tf(t)\mathrm{d}t}$$

- We assume that a tree remains undetected if the center is within the crown of a bigger tree. Then the detectability is directly

$$\pi(z;\lambda) = 1 - p_{Z>z} = e^{-\lambda \int_z^\infty tf(t)\mathrm{d}t}$$

# Examples on detectability




Detectability in hypothetical stands with unimodal (blue) and bimodal (red) size distribution of trees and two stand densities. The values show the canopy closure.

# Size distribution of detected trees

- The pdf of detectable trees is

$$f_D(z) = \frac{\pi(z;\lambda)f(z)}{\mathrm{E}(\pi(z;\lambda))}$$

  where $\lambda$ is replaced with $\hat{\lambda} = -\frac{\ln(1-cc)}{\mathrm{E}(Z)}$
- $f$ is a suitable continuous pdf (e.g Weibull or mixture of two Weibulls).
- Fitting $f_D(z)$ to the size distribution of detected trees (using ML) yields an estimate of $f(z)$, which is utilized to estimate $\mathrm{E}(Z)$ for estimation of stand density.

# Size distribution of detected trees

- The pdf of detectable trees is

$$f_D(z) = \frac{\pi(z;\lambda)f(z)}{\mathrm{E}(\pi(z;\lambda))}$$

  where $\lambda$ is replaced with $\hat{\lambda} = -\frac{\ln(1-cc)}{\mathrm{E}(Z)}$
- $f$ is a suitable continuous pdf (e.g Weibull or mixture of two Weibulls).
- Fitting $f_D(z)$ to the size distribution of detected trees (using ML) yields an estimate of $f(z)$, which is utilized to estimate $\mathrm{E}(Z)$ for estimation of stand density.

# Size distribution of detected trees

- The pdf of detectable trees is

$$f_D(z) = \frac{\pi(z;\lambda)f(z)}{\mathrm{E}(\pi(z;\lambda))}$$

  where $\lambda$ is replaced with $\hat{\lambda} = -\frac{\ln(1-cc)}{\mathrm{E}(Z)}$
- $f$ is a suitable continuous pdf (e.g Weibull or mixture of two Weibulls).
- Fitting $f_D(z)$ to the size distribution of detected trees (using ML) yields an estimate of $f(z)$, which is utilized to estimate $\mathrm{E}(Z)$ for estimation of stand density.

# Evaluation with empirical data

- 40 square 400 $m^2$ sample plots with 95% Scots pine in Kiihtelysvaara, North Carelia.
- Based on field measurements, the stand density varies between 400 ... 2250 trees per ha (mean 1223).
- The algorithm of Pitkänen (2004) was used for tree crown delineation from ALS data.
- Canopy closure measured as proportion of laser returns from the crown compared to all returns.

# Evaluation with empirical data

- 40 square 400 $m^2$ sample plots with 95% Scots pine in Kiihtelysvaara, North Carelia.
- Based on field measurements, the stand density varies between 400 . . . 2250 trees per ha (mean 1223).
- The algorithm of Pitkänen (2004) was used for tree crown delineation from ALS data.
- Canopy closure measured as proportion of laser returns from the crown compared to all returns.

# Evaluation with empirical data

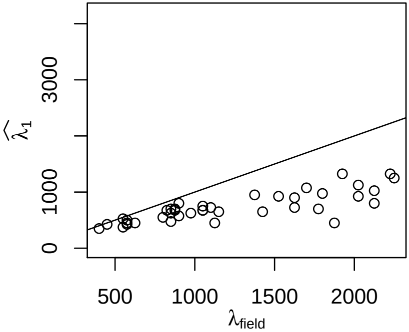
- 40 square 400 $m^2$ sample plots with 95% Scots pine in Kiihtelysvaara, North Carelia.
- Based on field measurements, the stand density varies between 400 . . . 2250 trees per ha (mean 1223).
- The algorithm of Pitkänen (2004) was used for tree crown delineation from ALS data.
- Canopy closure measured as proportion of laser returns from the crown compared to all returns.

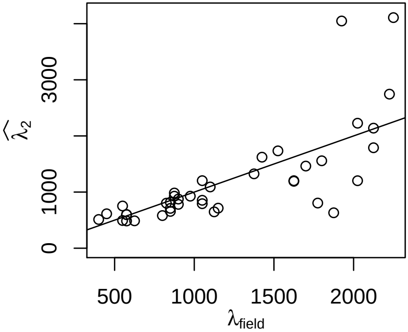
# Evaluation with empirical data

- 40 square 400 $m^2$ sample plots with 95% Scots pine in Kiihtelysvaara, North Carelia.
- Based on field measurements, the stand density varies between 400 ...2250 trees per ha (mean 1223).
- The algorithm of Pitkänen (2004) was used for tree crown delineation from ALS data.
- Canopy closure measured as proportion of laser returns from the crown compared to all returns.

# Results for estimation of $\lambda$

| Estimator | RMSE | bias |
|---|---|---|
| $\lambda_1 = \frac{N_{detected}}{A}$ | 629 | -500 |
| $\lambda_2 = -\frac{ln(1-cc)}{E(Z)}$ | 568 | -31 |



# Estimation errors vs. spatial pattern

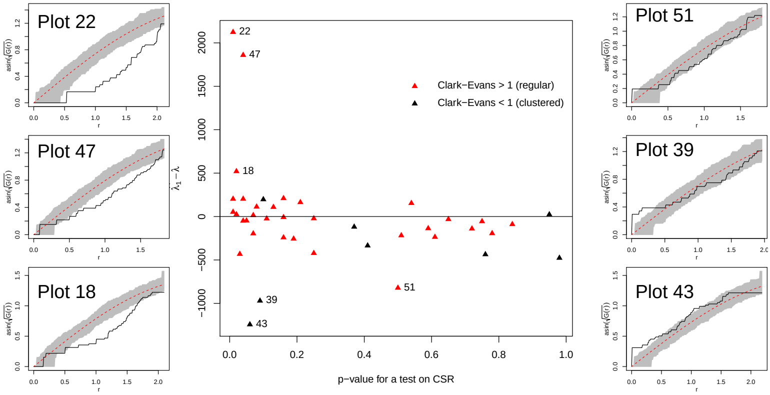



Figure : The estimation error of $\lambda$ versus the p-value of a test on CSR (global test on the $\widehat{G}$-function)

# Estimation errors vs. canopy cover

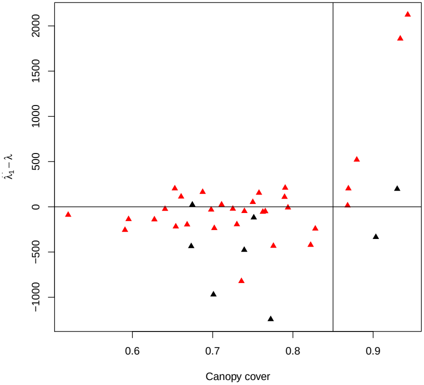



Figure : The estimation error of $\lambda$ versus the canopy cover.

# Discussion

- Even though the Boolean model has restrictive asumptions, it worked surprisingly well for most plots.
- The departure from random spatial pattern to the regular direction explained the most serious errors.
- Estimates were fairly accurate on plots with $cc < 0.85$.
- Extensions to other processes (Strauss, hard core) would be useful, but the equations get complicated.

# Discussion

- Even though the Boolean model has restrictive asumptions, it worked surprisingly well for most plots.
- The departure from random spatial pattern to the regular direction explained the most serious errors.
- Estimates were fairly accurate on plots with $cc < 0.85$.
- Extensions to other processes (Strauss, hard core) would be useful, but the equations get complicated.

# Discussion

- Even though the Boolean model has restrictive asumptions, it worked surprisingly well for most plots.
- The departure from random spatial pattern to the regular direction explained the most serious errors.
- Estimates were fairly accurate on plots with $cc < 0.85$.
- Extensions to other processes (Strauss, hard core) would be useful, but the equations get complicated.

# Discussion

- Even though the Boolean model has restrictive asumptions, it worked surprisingly well for most plots.
- The departure from random spatial pattern to the regular direction explained the most serious errors.
- Estimates were fairly accurate on plots with $cc < 0.85$.
- Extensions to other processes (Strauss, hard core) would be useful, but the equations get complicated.

# Bibliography

Maltamo, M., E. Næsset, J. Vauhkonen (ed.) (2014). *Forestry applications of airborne laser scanning- Concepts and Case Studies.* Springer..

Mehtätalo, L. (2006). Eliminating the effect of overlapping crows from aerial inventory estimates *Can. J. For. Sci. 36*, 1649–1660.

Pitkänen, J., Maltamo, M., Hyyppä, J. and Yu, X. (2004). Adaptive methods for individual tree detection on airborne laser based canopy height model. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXVI, part 8/W2: 187-191.

Stoyan, D., W.S. Kendall, J. Mecke (1995). *Stochastic geometry and its applications, $2^{nd}$ edition*. Chichester: John Wiley & Sons.

Vauhkonen, J., L. Mehtätalo (forthcoming). Matching remotely sensed and field measured tree size distributions. *Manuscript.*