Statistical inference and perfect simulation for point processes observed with noise

Jens Lund
The Royal Veterinary & Agricultural University
Copenhagen, Denmark

Ph.D. defence
February 25, 2000

E-mail `jlund@dina.kvl.dk`
Homepage `http://www.dina.kvl.dk/~jlund`

Overview of thesis

Introduction

Paper 1: Jens Lund and Mats Rudemo (1999)
Models for point processes observed with noise

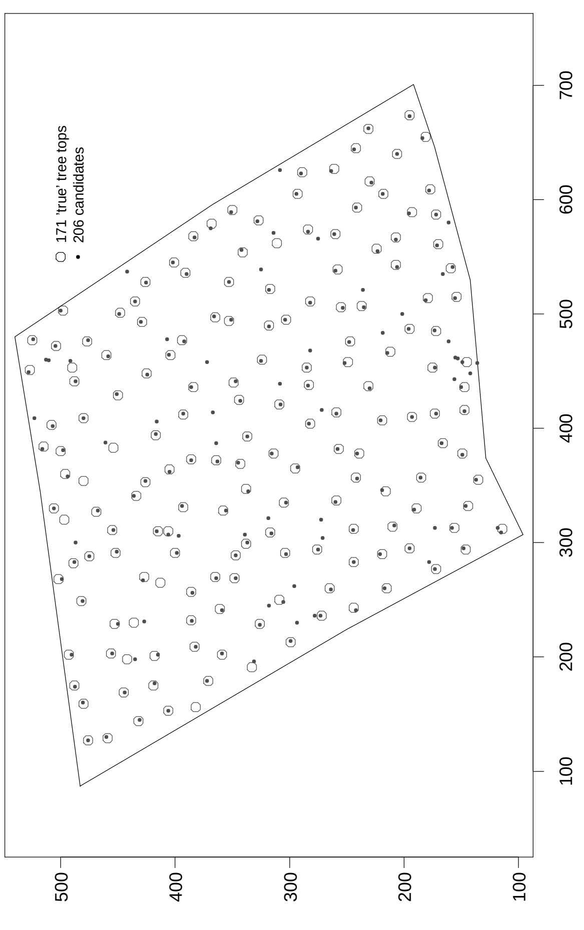
Paper 2: Jens Lund, Antti Penttinen, and Mats Rudemo (1999)
Bayesian analysis of spatial point patterns from noisy observations

Paper 3: Jens Lund and Elke Thönnes (1999)
Perfect simulation of point patterns from noisy observations
Now revised and improved!

Paper 4: Jens Lund (1998)
Survival of the Fattest? Self-thinning among Trees

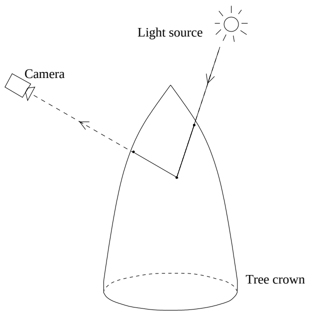
Preliminary Manuscript:
Jens Lund and Elke Thönnes (1999)
Perfect adaptive Metropolis-Hastings Simulation for Point Processes
Well, not that smart ...

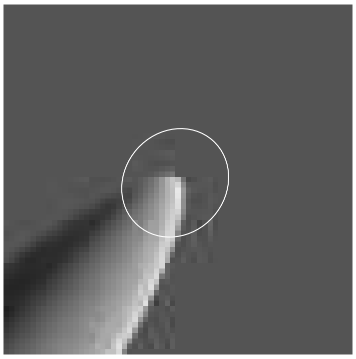
Data: two pointprocesses



The aerial photo

Tree model





(Larsen & Rudemo, 1998)

The data

True points $\mathbf{X}$

171 tree top positions in image as found from ground measurements

- $X = \{X_i, i \in M\}$, $M = \{1, \ldots, m\}$, $m = 171$
- Norway spruce
  Area $\approx 52 \times 52\text{m}^2$
  560m above ground
  Pixel size $15 \times 15\text{cm}^2$

Noisy points $\mathbf{Y}$

Tree top positions estimated from image

- $Y = \{Y_j, j \in N\}$, $N = \{1, \ldots, n\}$, $n = 206$
- Template model for one tree and light
  Best local match between template and image

(Dralle & Rudemo, 1996; Dralle & Rudemo, 1997)

Similar problems:
Indirect observations of positions
Image analysis

Model: from $\mathbf{X}$ to $\mathbf{Y}$

A area of observation

Thinning
Keep X_i, $i \in M$, with propability p
Remove with propability $1 - p$

Displacement
Remaining X_i
Generate Y-point according to density $f(\,\cdot\,|X_i)$,
$N_2(X_i + \mu, \Sigma)$

Censoring
Displaced points outside A disappear
Censored with probability $\int_{A^c} f(y|X_i)\,dy$ (≈ 0)

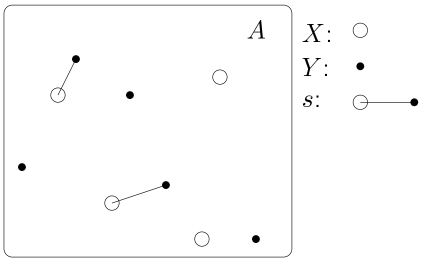
Ghost points
Poisson process, intensity $g(\,\cdot\,|X)$
(Can contain information on X)
Here homogeneous, $g(\,\cdot\,|X) = \lambda$

Parameter vector $\theta = (p, \lambda, \mu_1, \mu_2, \sigma_1, \sigma_2, \rho)$
(Assume suitable independence relations ...)

Three components

- X points, true points
- Y points, observed noisy version of X
- s, a matching, the connection between X and Y

A observation area



Goals

- Make model for transformation $X \to Y$
  Both X and Y observed,
  X fixed,
  Y noisy observation

- Maximum Likelihood Estimation of parameter θ,
  tests, etc.
  Both X and Y observed
  Training data
  Paper 1: Lund & Rudemo (1999)

- Reconstruct X from observation(s) of Y, i.e.
  Reconstruct ◯ from observation of •
  Bayes setup
  Y observed,
  θ *known* from training data
  Paper 2: Lund et al. (1999)

Notation

- True points
  $X = \{X_i, i \in M\}$, $M = \{1, \ldots, m\}$,
  $m = 171$ in general unknown

- Noisy version
  $Y = \{Y_j, j \in N\}$, $N = \{1, \ldots, n\}$, $n = 206$

- Parametervector $\theta = (p, \lambda, \mu_1, \mu_2, \sigma_1, \sigma_2, \rho)$

- Space of Mathcings S
  $s = (M_1, N_1, \pi) \in S$
  $Y_{\pi(i)} \sim X_i$ for $i \in M_1$.

- $M_1 \subseteq M = \{1, \ldots, m\}$,
  $N_1 \subseteq N = \{1, \ldots, n\}$,
  $|M_1| = |N_1|$

- $\mathcal{P}(M_1, N_1)$: bijections π from M_1 to N_1

Conditional likelihood

Theorem The conditional likelihood of Y given X is

$$L(Y|X,\theta) = \sum_{s \in S} T(X, Y, \theta, s)$$

where

$$T(X, Y, \theta, s) = L_1\, L_2\, L_3\, L_4$$

with

$$L_1 = p^{|M_1|} \prod_{i \in M_1} f(Y_{\pi(i)}|X_i, \mu, \Sigma)$$
$$L_2 = \prod_{i \in M \setminus M_1} \left(p \int_{A^c} f(y|X_i, \mu, \Sigma)\, dy + 1 - p \right)$$
$$L_3 = \lambda^{|N \setminus N_1|} \exp((1-\lambda)|A|)$$
$$L_4 = I\left[s \in S_{|X|,|Y|}\right]$$

Approximation: $\int_{A^c} f(y|X_i, \mu, \Sigma)\, dy = 0$
Paper 1: Lund & Rudemo (1999)

MLE estimation of θ

- X, Y observed

- Estimate θ

Problem: HUGE number of terms, $\gg 2^{171} \approx 3 \cdot 10^{51}$
Approximation

$$L(Y|X,\theta) \approx \sum_{s \in \tilde{S}} T(X, Y, \theta, s)$$

Find suitable $\tilde{S}$ with $|\tilde{S}|$ small, i.e.
Find a few large terms in the likelihood function
Iterative algorithm

Estimation of θ, tests, etc.

Joint distribution (Y, s): $L(Y, s|X, \theta) = T(X, Y, \theta, s)$
The matching s is missing data

Reconstruction of $\mathbf{X}$

- Y observed, θ known from training data
- Reconstruct X
  Estimation of derived statistics:
  K-, L-, F-, G-function etc.

Bayes setup:

- Likelihood function: $L(Y, s|X, \theta) = T(X, Y, \theta, s)$
- Prior distribution on X: $L(X)$
  Penalizes irregular point patterns
- Posterior distribution:
  $L(X, s|Y, \theta) \propto L(Y, s|X, \theta)L(X) \;=\; T(X, Y, \theta, s)L(X)$
- Make MCMC samples from (X, s)
  Marginal X: $(X, s) \to X$

Point process priors

Prior:

- Poisson
- Gibbs $f(x) \propto \alpha^{|X|} \prod_{i_1 \sim i_2} H(\|X_{i_1} - X_{i_2}\|)$
  - Strauss $H(r) = \beta^{I[r \leq R]}$
  - Logistic

$$H(r) = \frac{1}{1 + \exp(-\beta_1(r - R))}$$

No obvious connection between α, β, and $\mathrm{E}|X|$
Pilot studies = trial and error, $\mathrm{E}|X| = 171$

$L(X) = L(X|\,|X| = n)L(n)$ doesn't work
Unknown normalizing constants

$$r_{\text{posterior}} = \frac{L(X'_{t+1}, s'_{t+1}|Y, \theta)}{L(X_t, s_t|Y, \theta)} = \frac{L(X'_{t+1}|n+1)L(n+1)}{L(X_t|n)L(n)} \cdots$$

(Geyer & Thompson, 1995)

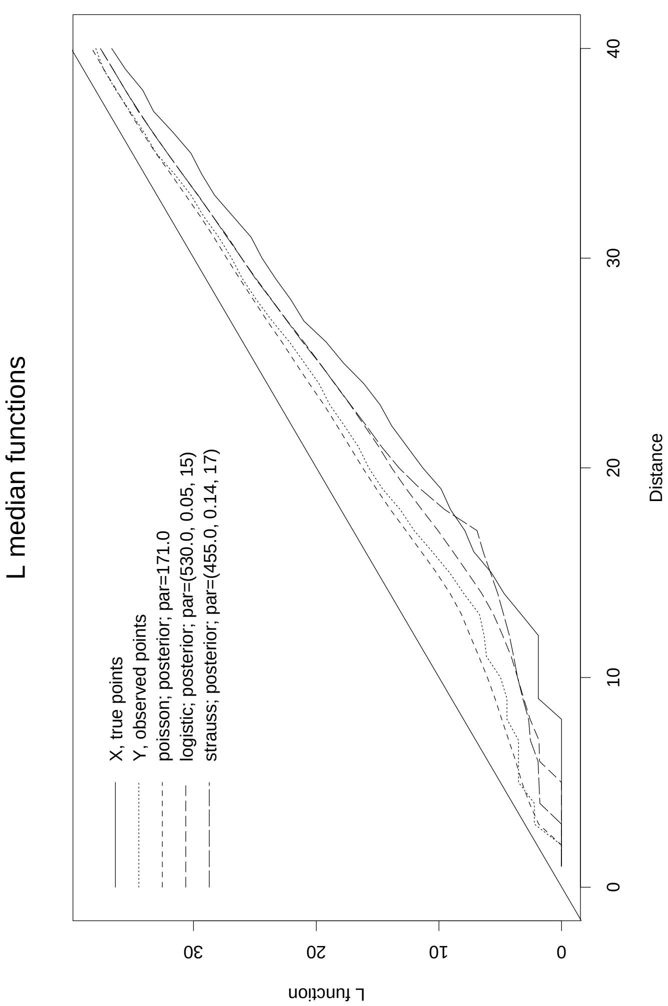
Hyperprior on α and β not possible

$(\mathbf{X}, \mathbf{s})$ MCMC

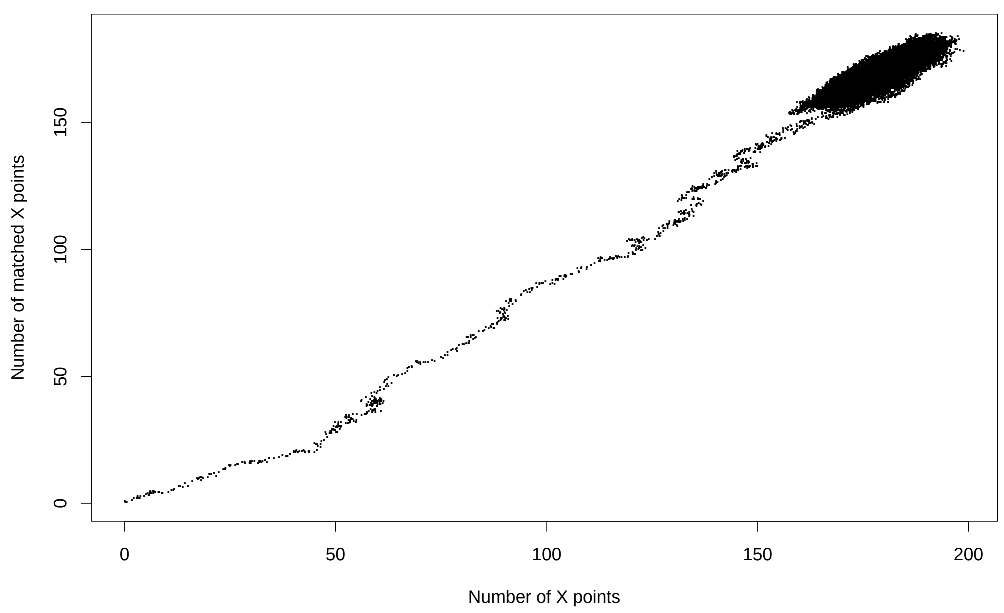
- Construct Markov Chain: (X_t, s_t)
  Stationary distribution $L(X, s|Y, \theta)$
- Standard way to do this!
  Metropolis-Hastings MCMC
  (Gilks et al., 1996; Geyer & Møller, 1994)
- Make proposals (X'_{t+1}, s'_{t+1})
  $Q(X'_{t+1}, s'_{t+1}|Y, \theta, X_t, s_t)$
  Choose proposal distribution in a clever way :-)
- Accept the proposal with probability
  $\alpha = \min(1, r_{\text{posterior}} \times r_{\text{proposition}} \times |J|)$
  Accept: $(X_{t+1}, s_{t+1}) = (X'_{t+1}, s'_{t+1})$
  Reject: $(X_{t+1}, s_{t+1}) = (X_t, s_t)$
  Ensures the correct stationary distribution
- Hastings-Green ratio

$$r_{\text{posterior}} = \frac{L(X'_{t+1}, s'_{t+1}|Y, \theta)}{L(X_t, s_t|Y, \theta)}$$

$$r_{\text{proposition}} = \frac{Q(X_t, s_t|Y, \theta, X'_{t+1}, s'_{t+1})}{Q(X'_{t+1}, s'_{t+1}|Y, \theta, X_t, s_t)}$$



102000 simulations



Perfect simulation of $\mathbf{L}(\mathbf{X}, \mathrm{s} | \mathbf{Y}, \theta)$

Crash course on perfect simulation:

① ② ③ ④ ... H H H T H T T H

Coupling from the past
Sandwiching
Dominated coupling from the past

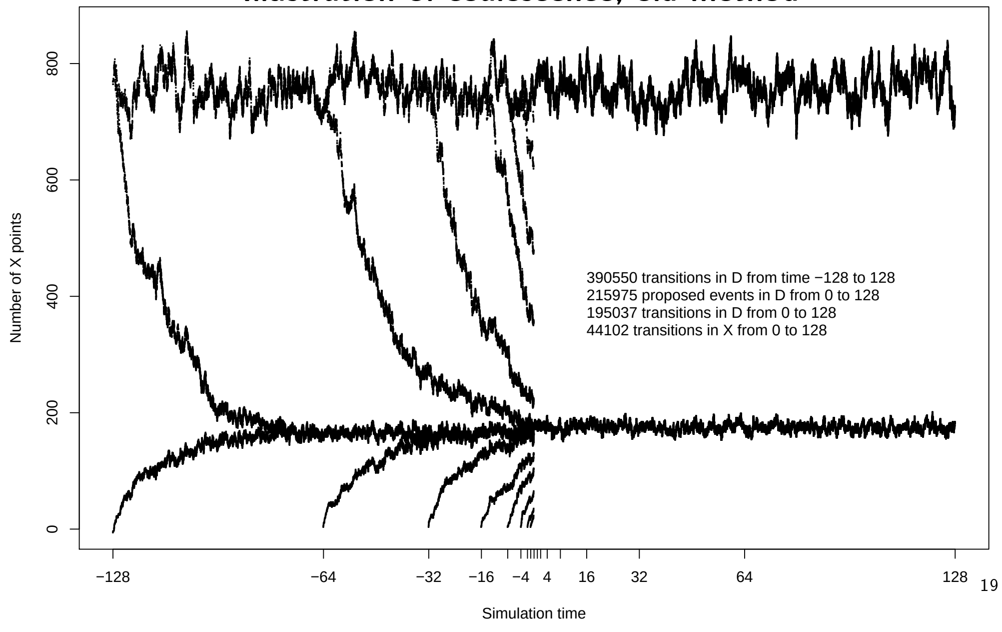
Spatial birth and death processes
Continuous time Markov processes
Free and matched X-points separately

Point processes: $\leq$ is the subset relation

Logistic; posterior; par=(175.0, 0.8, -3)
Logistic; posterior; par=(225.0, 0.1, 7)

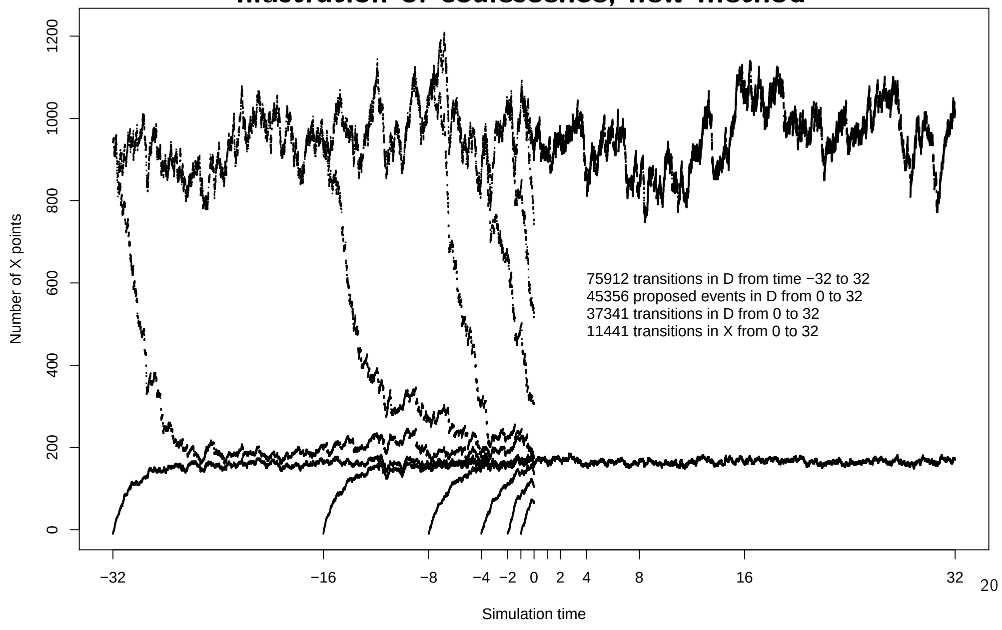
CFTP

Illustration of coalescence; old method



CFTP

Illustration of coalescence; new method



Conclusions

- New model for point processes observed with noise
- Estimation of parameters in this model
- Improved inference for the noisy observation
- Perfect simulation can (almost) be used on real data, in real models

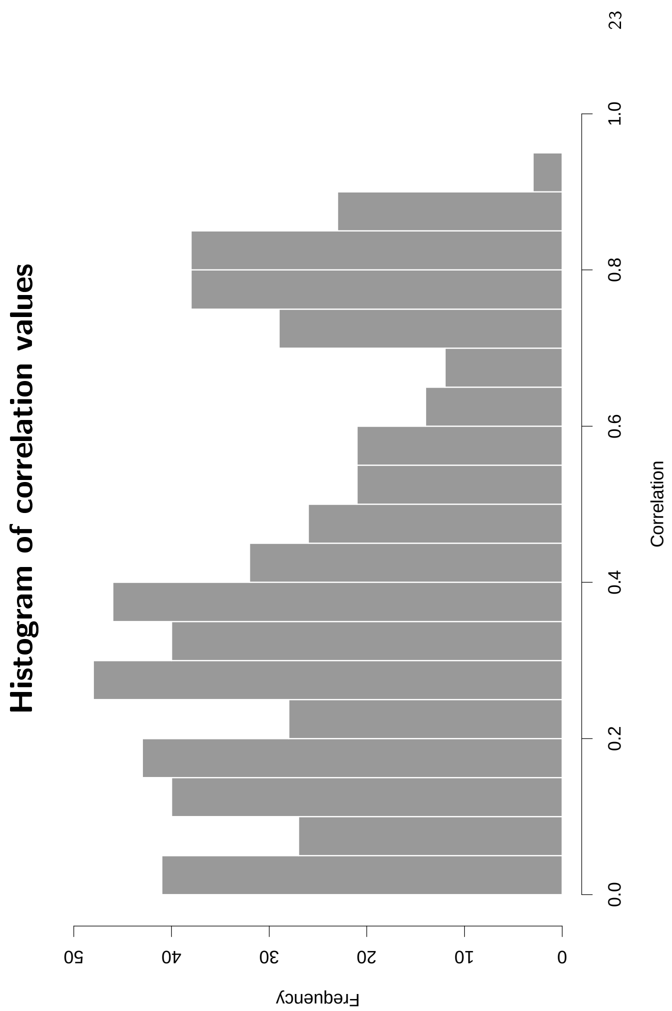
- The forest people were helped by a model for self thinning, perhaps not for positioning (Paper 4)

Further developments:

- Improve positioning from images:
  Couple image analysis and point process analysis closer together
- Use several images and do 3-dimensional positioning
- Perfect simulation

Additional slides

Both as a supplement to the previous slides, and to describe the two papers not covered so far



Neighbours of state s

Remember: $s = (M_1, N_1, \pi)$, M_1: the matched X-points, N_1: the matched Y-points, $Y_{\pi(i)} \sim X_i$

s itself

Add pair unpaired X_i, closest unpaired Y_j

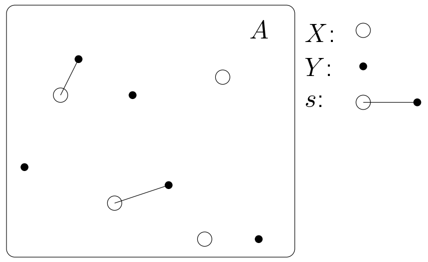
Remove pair

Shift X point with closest unpaired X_i

Shift Y point with closest unpaired Y_j

Exchange Y-points among two pairs
Only between close X-points

Typical $\approx |M| + 3|M_1|$ neighbours



Iterative algorithm

Given S_t with r matchings s, estimate θ_t

1. S_{t+1}
   r matchings among neighbours of $s \in S_t$
   with largest value of
   $T(X, Y, \theta_t, s')$, s' a neighbour

2. θ_{t+1} maximizes $L(S_{t+1}, \theta)$

3. Stop if $S_t = S_{t+1}$

Approximation

$$L(S_{t+1}, \theta) = \sum_{s \in S_{t+1}} T(X, Y, \theta, s)$$

Estimate θ_{t+1}

Starting set S_0

$S_0 =$ the r states among s_k, $k = 0, \ldots, \min(m, n)$ with highest value $\sup_\theta T(s_k, \theta)$

Likelihood approximation with one term (fast)

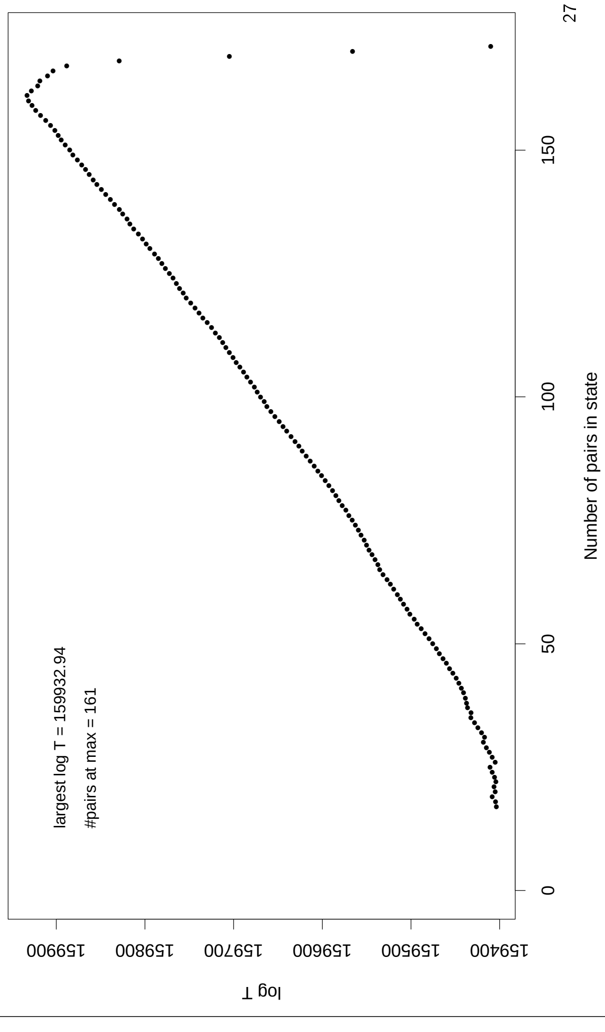
The states s_k, $k = 0, \ldots, \min(m, n)$:

1. s_k consists of k pairs of X- and Y-points

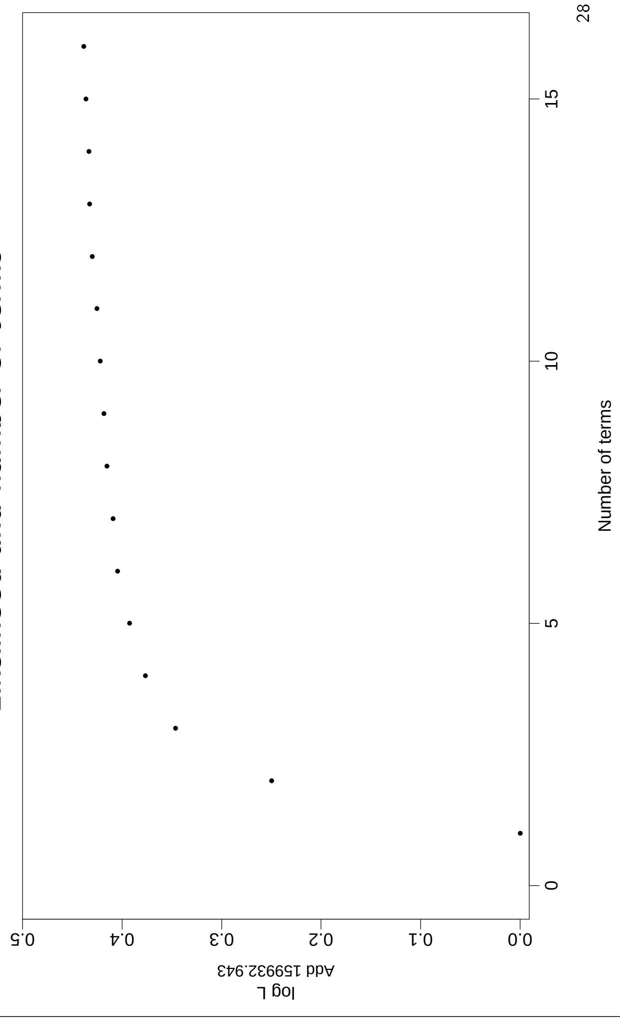
2. From each X_i: let a circle grow until it hits an unpaired Y_j-point

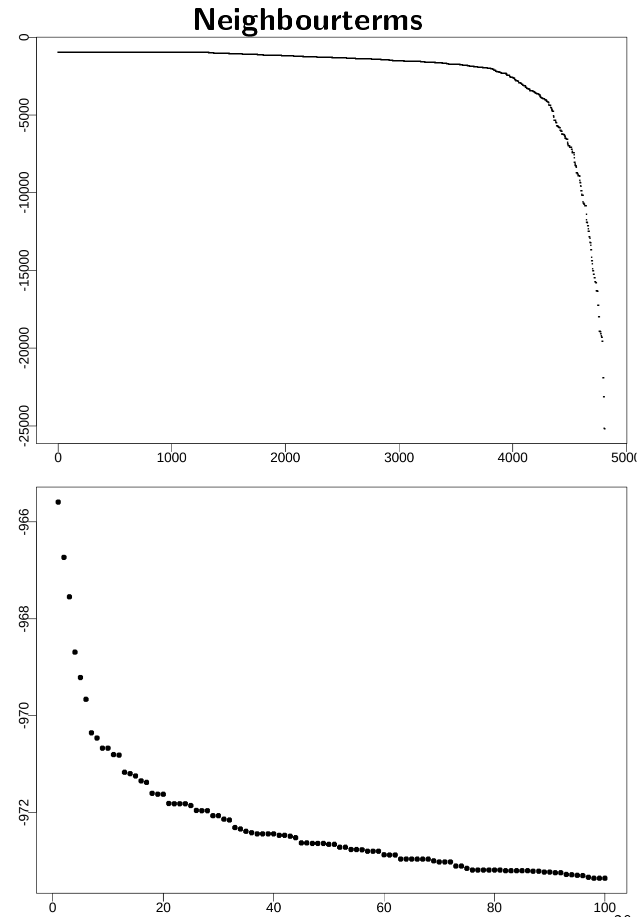
3. s_1 with 1 hit, s_2 with 2 hits, ...

Starting terms



Likelihood and number of terms





Area = 160898.5, number of neighbours = 4812,
term(500) = -975.89, term(1000) = -977.91, term(2000) = -1182.55

The 8 pairs in the approximation

- $\log T(s, \hat{\theta})$ for each of the 8 terms
- $\hat{\theta}$ maximum likelihood estimate
- number of pairs in each of the states

term	1	2	3	4	5	6	7	8
$\log T(s, \hat{\theta})$	159932.9	159931.8	159931.0	159929.8	159929.3	159928.8	159928.1	159928.0
# pairs	161	161	160	160	162	160	162	160

$$\exp(159928.0 - 159932.9) = \exp(-4.9) = 0.0074$$

Parameter estimates

Based on the 8 terms

parameter	p	λ	μ_1	μ_2	σ_1^2	σ_2^2	$\text{cov}_{1,2}$
estimate	0.941	0.00028	-0.342	0.082	1.004	2.028	-0.049
std. dev.	0.018	0.00004	0.080	0.113	0.115	0.235	0.115

Standard deviations: Numerical differentiation of $L(S, \theta)$

Tests

Likelihood ratio tests based on $L(S, \theta)$

Test of the hypotheses

- $H_0 : (\mu_1, \mu_2) = (0, 0)$
- $H_1 : \mu_2 = 0$
- H_2 : full model

Model	$-2 \log Q$	df	p-value
H_1 (against H_2)	0.523	1	0.47
H_0 (against H_1)	17.32	1	0.00003
H_0 (against H_2)	17.85	2	0.00013

$(\mathbf{X}, \mathrm{s})$ proposals

Remember: $s = (M_1, N_1, \pi)$
Current value: (X_t, s_t)

Proposal (X'_{t+1}, s'_{t+1}):

Add a point b

A matched point h
Choose unmatched free Y-point, say Y_j
$\xi \sim N(Y_j - \mu, \Sigma)$, $X'_{t+1} = X_t \cup \{\xi\}$
s'_{t+1} from s_t by adding ξ and connection between Y_j and ξ

An unmatched point $1 - h$
$(X'_{t+1}, s'_{t+1}) = (X_t \cup \{\xi\}, s_t)$
ξ uniformly distributed in A

Delete a point $1 - b$

A matched point $1 - h$
X_i uniformly among $M_{1,t}$ paired X-points
$X'_{t+1} = X_t \setminus \{X_i\}$
Delete connection X_i, $Y_{\pi(i)}$ to obtain s'_{t+1}

An unmatched point h
X_i uniformly among $M_t \setminus M_{1,t}$ unpaired X-points
$(X'_{t+1}, s'_{t+1}) = (X_t \setminus \{X_i\}, s_t)$

Accept probability α

$|J| \equiv 1$
No changes to points, just add and delete

Add point case

Matched points

$$r_{\text{posterior}} = \frac{L(X'_{t+1}, s'_{t+1}|Y, \theta)}{L(X_t, s_t|Y, \theta)} = \frac{L(X'_{t+1})}{L(X_t)} \frac{p}{\lambda} f(Y_j|\xi, \mu, \Sigma)$$

$$r_{\text{proposal}} = \frac{1-b}{b} \frac{1-h}{h} \frac{1/|M'_{1,t+1}|}{f(Y_j|\xi, \mu, \Sigma)/|N_t \setminus N_{1,t}|}$$

Unmatched points

$$r_{\text{posterior}} = \frac{L(X'_{t+1}, s'_{t+1}|Y, \theta)}{L(X_t, s_t|Y, \theta)} = \frac{L(X'_{t+1})}{L(X_t)}(1-p)$$

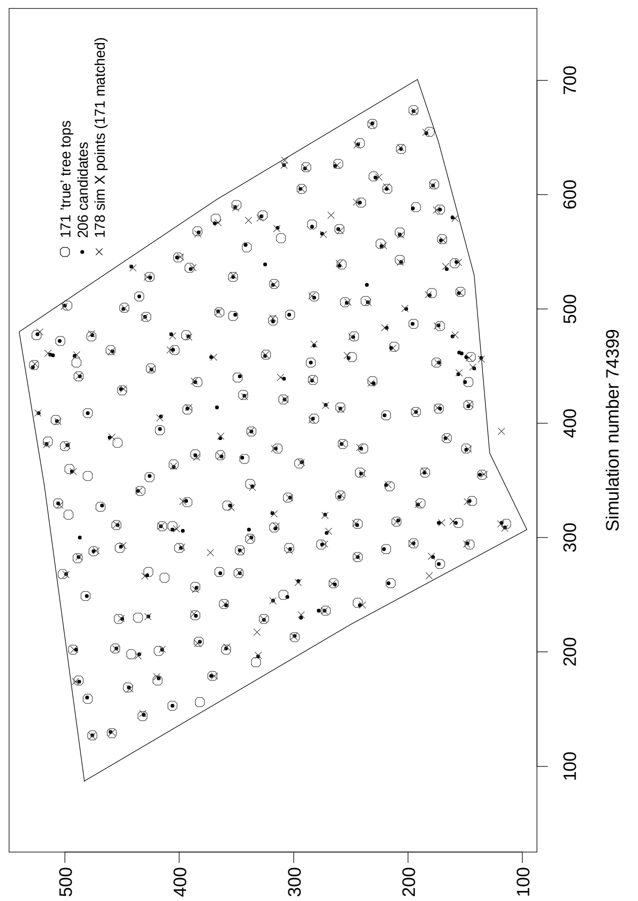
$$r_{\text{proposal}} = \frac{1-b}{b} \frac{h}{1-h} \frac{1/|M'_{t+1} \setminus M'_{1,t+1}|}{1/|A|}$$

$b = h = \frac{1}{2}$

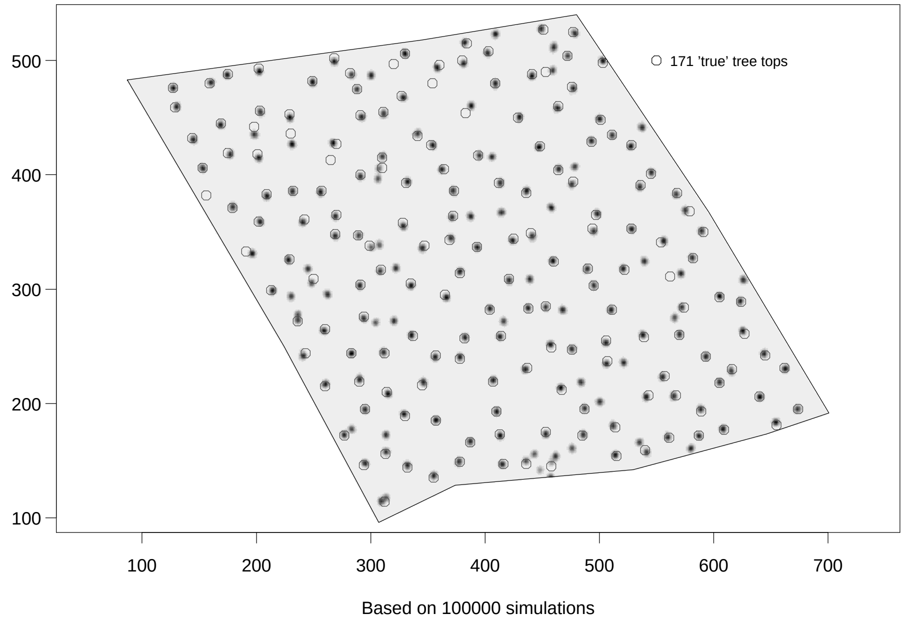
Specific priors

Distribution	Parameters
Poisson	$\alpha\|A\| = 171$
Strauss	$(\alpha\|A\|, \beta, R) =$ (455, 0.14, 17)
Logistic	$(\alpha\|A\|, H(0), R) =$ (530, 0.05, 15)

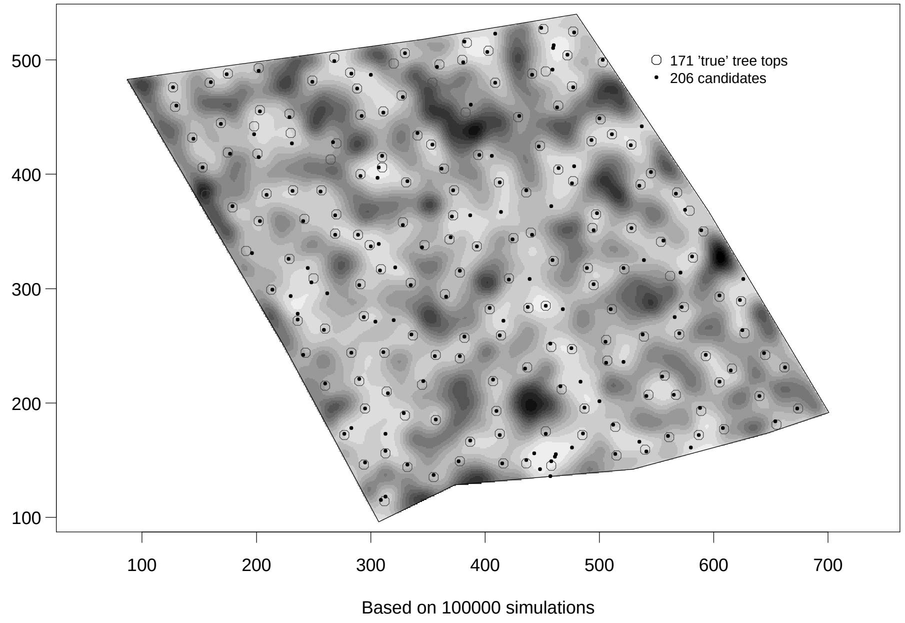
Logistic prior: main example in the following



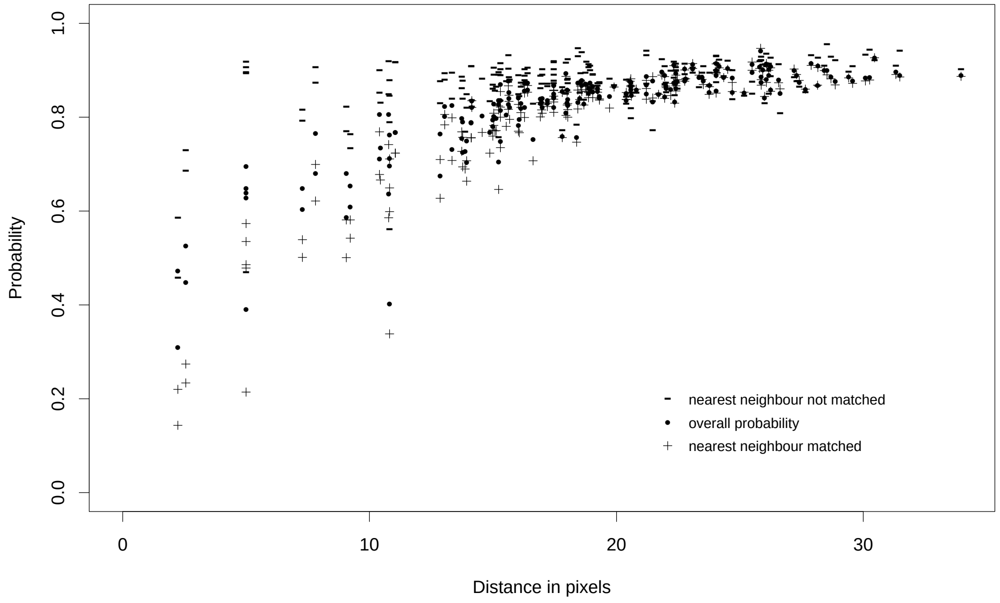
Intensity surface for all simulated X points



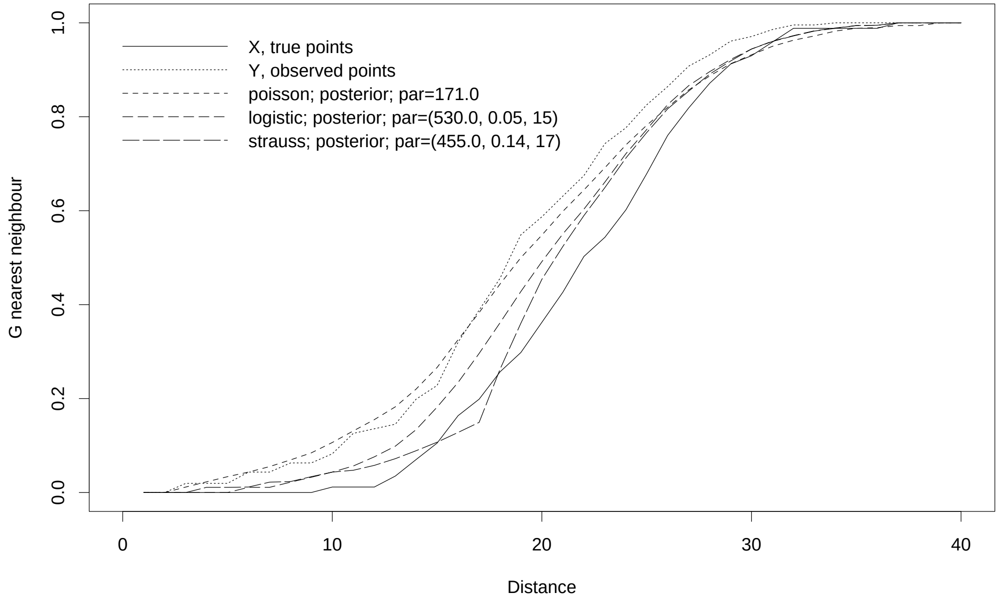
Intensity surface for free simulated X points



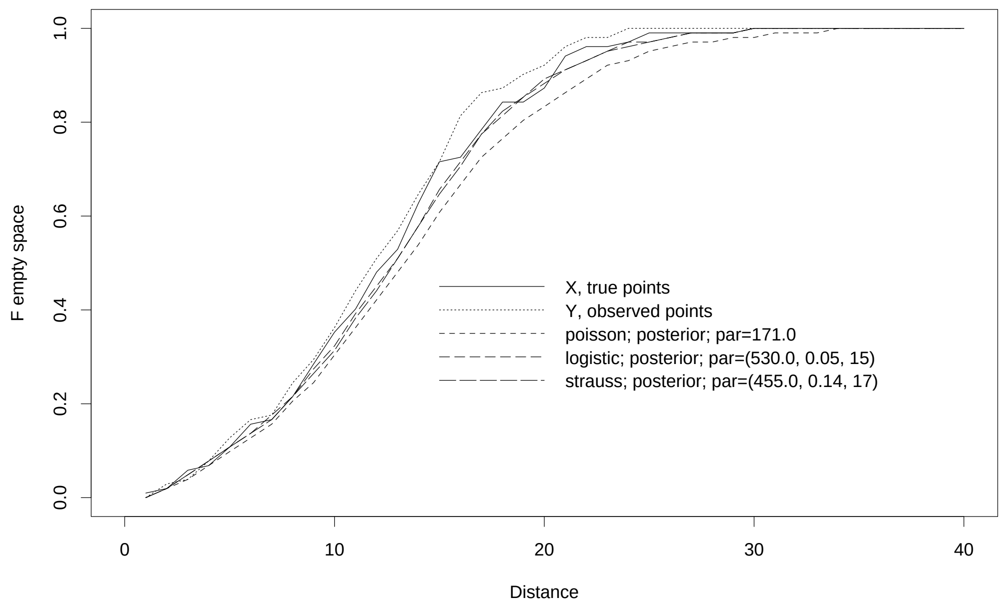
Probability of being matched



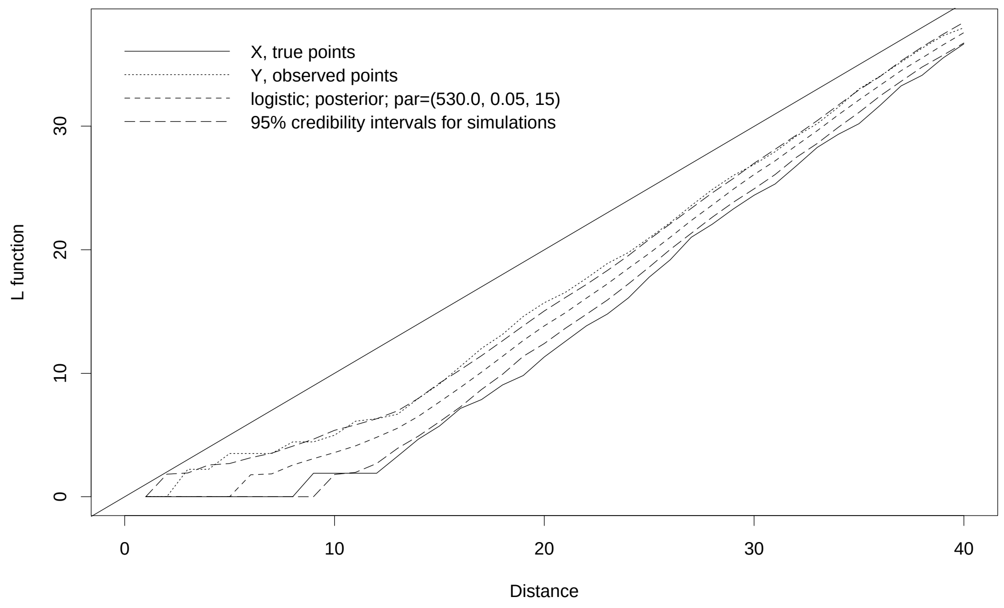
G median functions



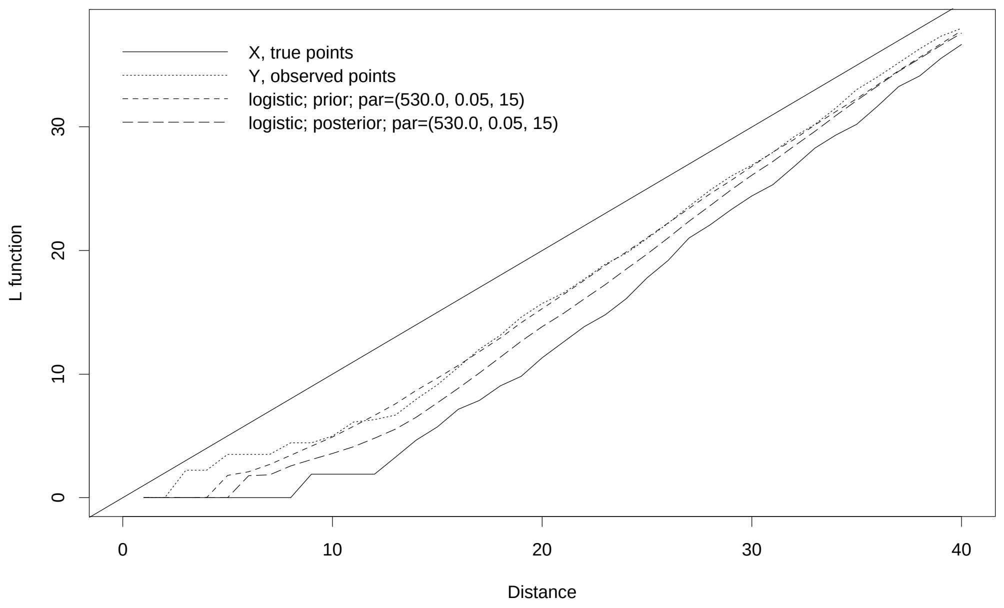
F median functions



L function credibility interval



L median functions

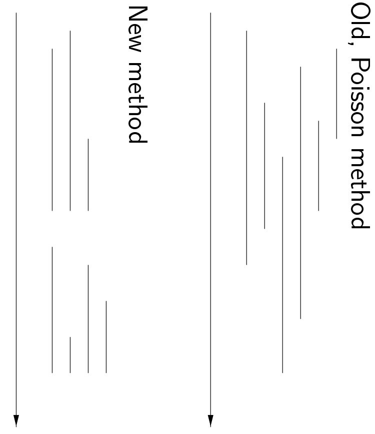


Improved version

N_t number of X-points matched to the Y_j-point in D_t

Old, Poisson method

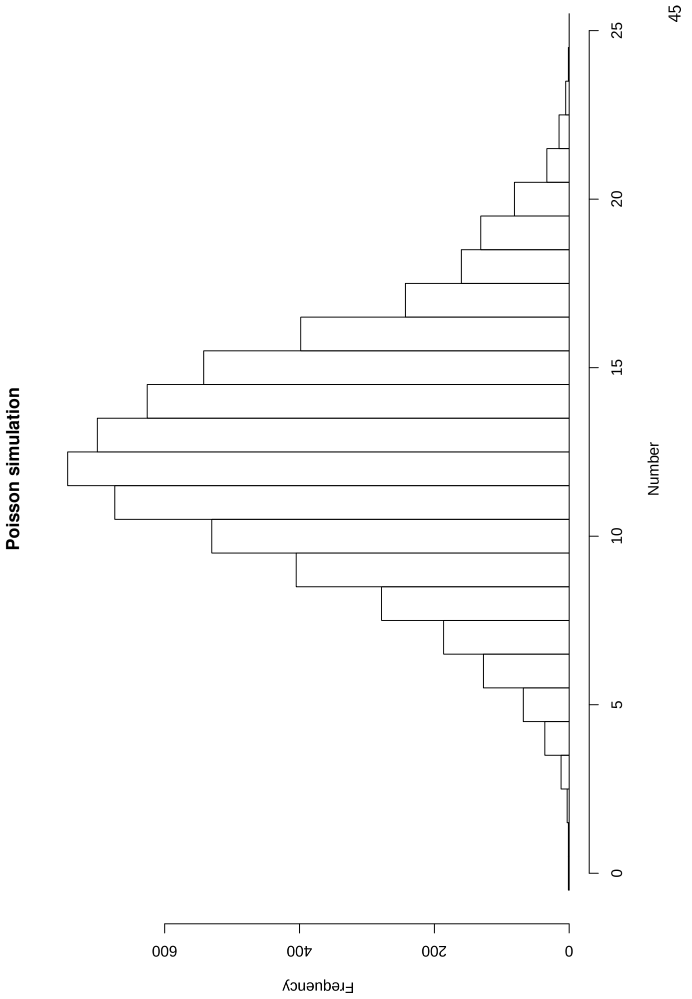
New method

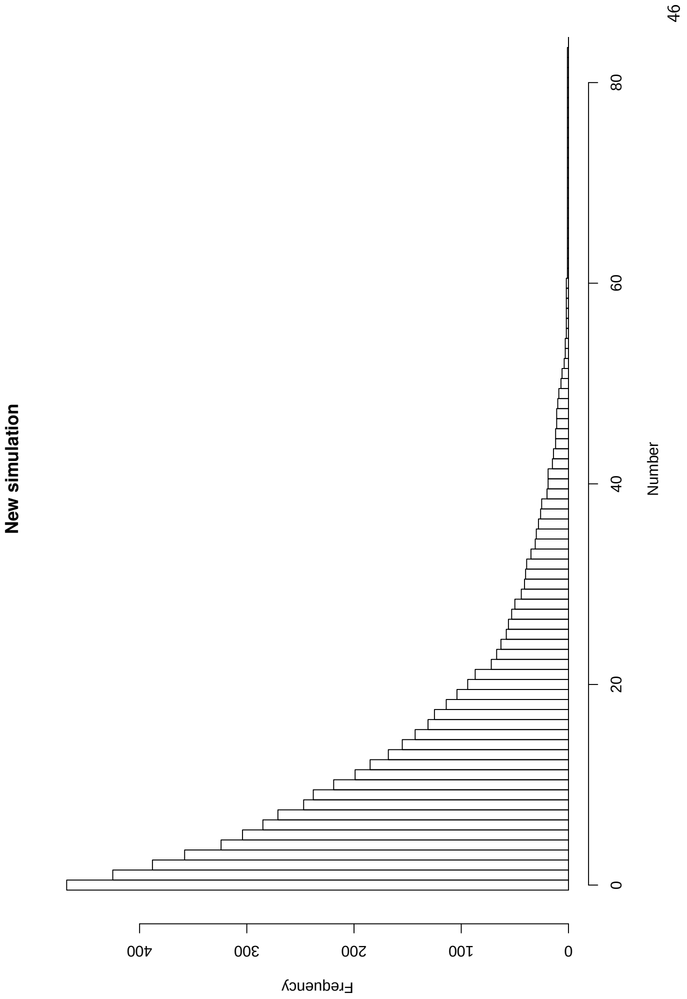


Sandwiching ⇒ several X-points per Y-point in dominating chain.

Poisson, “old method”:
Intensity up: λ
Intensity one down: N_t

New method:
Intensity up: λ
Intensity going to 0: 1



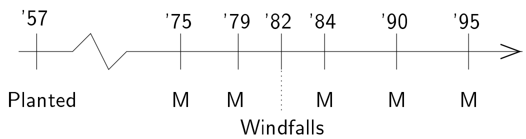


Survival of the fattest?

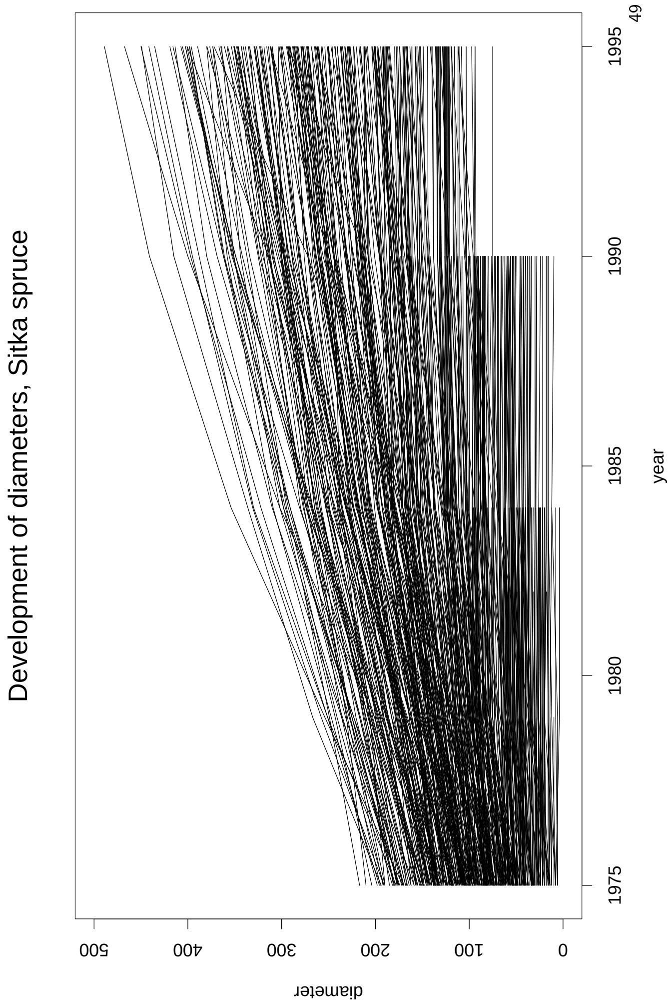
- Presentation of data
- Aspects of self-thinning
- Continuous life times, discrete observations
- Choice of covariates
- Time dependent covariates
- Spatial dependence, competition index
- Results
- Final remarks

Data

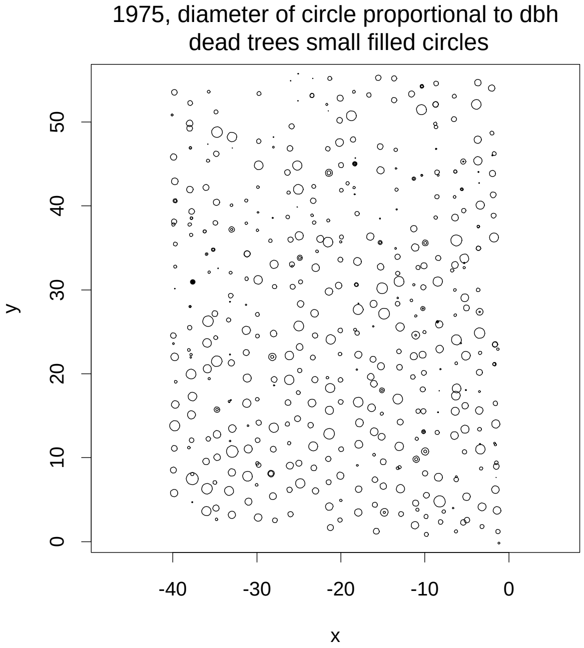
- Even-aged unthinned Sitka spruce, experiment MB-II
- 1975: 479 Sitka spruces, basal area 17.8 m^2/ha, 2228 trees/ha
- 1995: 243 Sitka spruces alive, basal area 61.7 m^2/ha, 1133 trees/ha
- Positions
- Diameters
- Some heights (will not be used)



M: Measurement time points

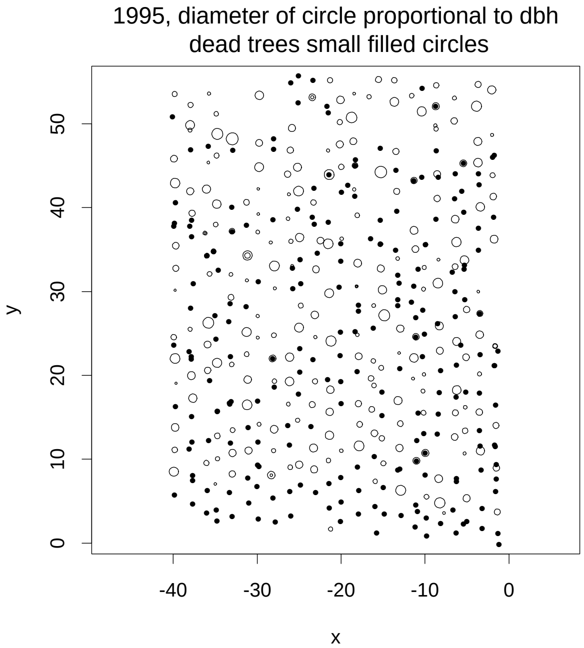


Map



478 trees alive

Map



243 trees alive

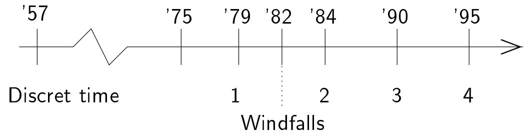
Aspects of Self-thinning

- Continuous life times, discrete observations
- Choice of covariates
- Time dependent covariates
- Spatial dependence, competition index

A spatio-temporal model

Continuous life times

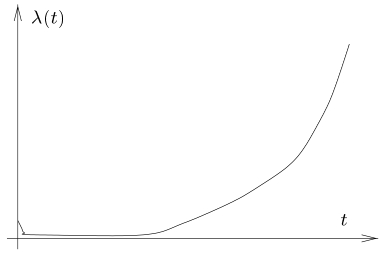
Life times of trees are continuous, but we observe only discrete life times, interval censoring:



Year	'75	'79	'84	'90	'95	Total
Dead	1	2	9	66	115	193
Censored	0	0	43	0	243	286
Total	1	2	52	66	358	479

Continuous life times

- Describe the life time by the hazard $\lambda(t)$
- Interpretation: $P(T \in [t, t+dt[\,|\, T \geq t) = \lambda(t)\,dt$



- Compare to interpretation of density $f(t)$: $P(T \in [t, t+dt) = f(t)\,dt$

Discrete Observations

Cox proportional hazards model

$$\lambda(t|x) = \lambda_0(t) e^{x'\beta}$$

Vector notation: $x'\beta = \beta_1 x_1 + \cdots + \beta_k x_k$
Discrete time hazard with cutpoints $\theta_0, \ldots, \theta_4$

$$\begin{aligned} \lambda(y|x) &= P(Y = y | Y \geq y) \\ &= 1 - \exp\left(-e^{x'\beta} \int_{\theta_{y-1}}^{\theta_y} \lambda_0(s)\,ds\right) \end{aligned}$$

This is in fact a "nice expression" since

$$\text{cloglog}(\lambda(y|x)) = \log(-\log(1 - \lambda(y|x))) = \tilde{\theta} + x'\beta$$

Likelihood function

$$P(Y = y) = \prod_{i=1}^{y-1} (1 - \lambda(i|x)) \lambda(y|x)$$

$$P(Y \geq y) = \prod_{i=1}^{y} (1 - \lambda(i|x))$$

- Easy estimation etc with software for Generalized Linear Models
- Markov structure in time

(Fahrmeir & Tutz, 1994)

Choice of Covariates

Two aspects:

1. Time dependent covariates
2. Spatial dependence, covariates depend on neighbour trees

One tree

- Diameter dbh
- Basal area
- Position x and y

Several trees

- Competition index
  Take spatial structure into account

(Rathbun & Cressie, 1994)

Time Dependent Covariates

Use covariates measured at previous time point:

	'57	'75	'79	'82	'84	'90	'95
				Windfalls			
Discret time		1			2	3	4
Covariates		'75			'79	'84	'90

Likelihood for one tree

$$P(Y=y)=\prod_{i=1}^{y-1}(1-\lambda(i|x_{i-1}))\lambda(y|x_{i-1})$$

Markov structure

Likelihood for one timepoint i, all trees

$$\prod_{\text{trees died}}(\lambda(i|x_{i-1}))\prod_{\text{trees survived}}(1-\lambda(i|x_{i-1}))$$

Spatial Dependence, Competition Index

A CI is a covariate that describes the competition among the trees,
the relative size of the trees

CI can be made in *many* ways, here Hegyi (1974) used (Johannsen, 1996; Tomé & Burkhart, 1989; Pukkala, 1989)

A_i: Basal area tree i

$$\mathsf{CI}_i=\sum_{j\neq i:\mathsf{dist}(i,j)\leq 4\mathsf{m}}\frac{A_j}{A_i\,\mathsf{dist}(i,j)}$$

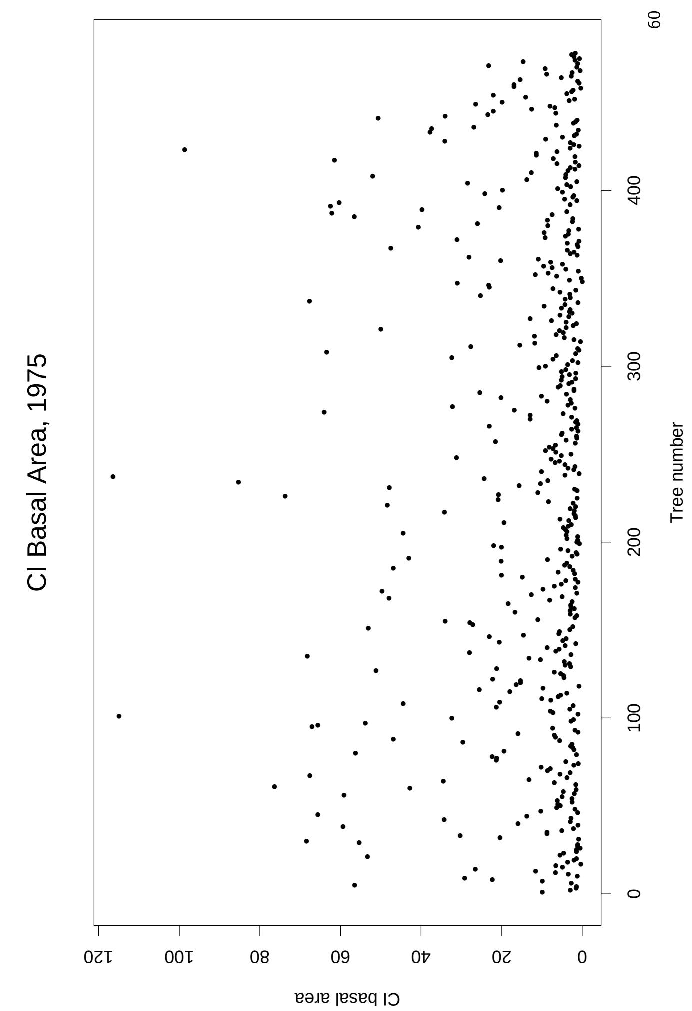
Based on living trees

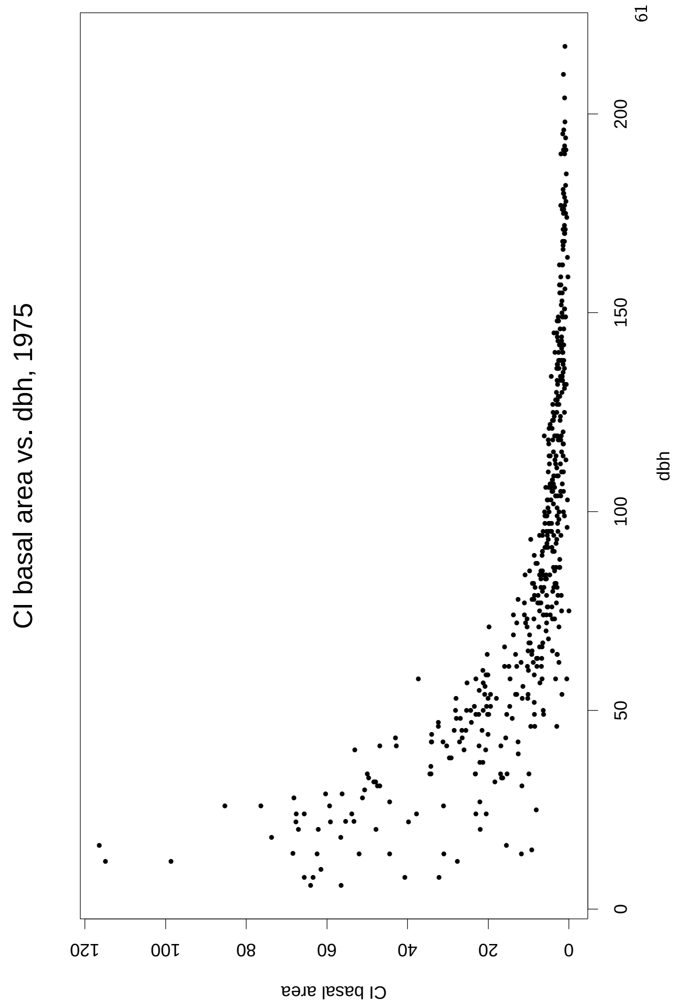
Very heavy-tailed distribution due to small trees and trees close together

$$\mathsf{CI}_i=\sum_{j\neq i:\mathsf{dist}(i,j)\leq 4\mathsf{m}}\frac{\min\left(\frac{A_j}{A_i},16\right)}{\max(\mathsf{dist}(i,j),0.5\mathsf{m})}$$

The CI has approximately the same distribution at different time points, and the CI is still heavy tailed

Define CI in the same way for dbh





Spatial inhomogeneity?

Include x and y coordinates for each tree as covariates

Also xy

Results

Covariates in final model

+ Diameter dbh

+ CI based on basal area

+ x and y position

− Basal area

− CI based on dbh

Interpretation of estimates in the continuous time hazards model

$$\lambda(t|x) = \lambda_0(t)e^{x'\beta}$$

Factor to multiply on the baseline hazard in the Cox proportional hazards model

Variable	Increase	Factor	Confidence interval
dbh	5cm	0.32	$[0.24; 0.42]$
ciba	10	1.28	$[1.13; 1.43]$
x	10m	1.30	$[1.12; 1.51]$
y	10m	0.85	$[0.75; 0.97]$

The time parameters are not so interesting per se

Open Problems

- Does the model describe the data in a satisfactory way? Results seem resonable (!?)
- Further inhomogeneity or over-dispersion? Possible solution: a Frailty model from survival analysis (Scheike & Jensen, 1995)
- Are the trees independent in a single time period given previous measurements?

Somewhat arbitrary choices e.g.

- Interaction radius? (Here 4m)
- The basic form of the CI

Description of diameter development

Final Remarks

Differences in tradition between forest research and statistics

Statistics

- Define

$$\text{CI}_i = \sum_{j \neq i : \text{dist}(i,j) \leq r} \frac{\text{dbh}_j^\alpha}{\text{dbh}_i^\alpha \, \text{dist}(i,j)^\beta}$$

  estimate r, α, β (I didn't do that :-)
- Take only measurable covariates into account

Forest research

- Compare the results for different fixed values of r, α, β (e.g. Tomé & Burkhart, 1989)
- A lot of verbal describtions of differences between "subjects" (e.g. Henriksen, 1981; Skovsgaard, 1997)

A counter example

State space $\Omega = \{0, 1\}$

Dominating chain $D_t \sim \pi = (0.5, 0.5)$

Target chain $X_t \sim (1-p, p)$

Transition matrix for (X_t):

$$P = \begin{pmatrix} 1-q & q \\ 1-q' & q' \end{pmatrix}$$

with $q = p/(1-p)(1-q')$

Want $X_t \leq D_t$

Simultaneous distribution (X, D)

X\D	0	1	Total
0	0.5	0.5-p	1-p
1	0	p	p
Total	0.5	0.5	1

$p < 0.5$, $0 < p/(1-p) < 1$, $q' \in]0, 1[$

The coupling

Randomness (U_t) iid uniform $[0, 1]$

Furthermore:
$q' < 0.5$ and
$q' > 1 - 0.5(1-p)/p$ (that is $q < 0.5$)

$$D_{t+1} = \begin{cases} 0 & \text{if } U_t \leq 0.5 \\ 1 & \text{if } U_t > 0.5 \end{cases}$$

If $X_t = 0$:

$$X_{t+1} = \begin{cases} 0 & \text{if } U_t \leq 1-q \\ 1 & \text{if } U_t > 1-q \end{cases}$$

If $X_t = 1$:

$$X_{t+1} = \begin{cases} 0 & \text{if } U_t \leq 1-q' \\ 1 & \text{if } U_t > 1-q' \end{cases}$$

This ensures $X_t \leq D_t$ for all t and $X_t \sim (1-p, p)$

The problem

Choose between two transition kernels P_0 and P_1 for X_t

Choose between two values of q', q'_0 and q'_1

Make the choice dependent on the value of D_t:

$$q' = \begin{cases} q'_0 & \text{if } D_t = 0 \\ q'_1 & \text{if } D_t = 1 \end{cases}$$

Let $s \equiv P(D_t = 0 | X_t = 0)$

$X_t = 1 \Rightarrow D_t = 1 \Rightarrow q' = q'_1$
$X_t = 0 \Rightarrow P = sP_0 + (1-s)P_1$

That is X_t evolves as

$$\tilde{P} = \begin{pmatrix} s(1-q_0) + (1-s)(1-q_1) & sq_0 + (1-s)q_1 \\ 1-q'_1 & q'_1 \end{pmatrix}$$

$\tilde{P}$ does *not* have $(1-p, p)$ as stationary distribution

References

Dralle, K. & Rudemo, M. (1996), 'Stem number estimation by kernel smoothing of aerial photos', *Canadian Journal of Forest Research* **26**, 1228–1236.

Dralle, K. & Rudemo, M. (1997), 'Automatic estimation of individual tree positions from aerial photos', *Canadian Journal of Forest Research* **27**, 1728–1736.

Fahrmeir, L. & Tutz, G. (1994), *Multivariate Statistical Modelling Based on Generalized Linear Models*, Springer Series in Statistics, Springer-Verlag.

Geyer, C. J. & Møller, J. (1994), 'Simulation procedures and likelihood inference for spatial point processes', *Scandinavian Journal of Statistics* **21**(4), 359–373.

Geyer, C. J. & Thompson, E. A. (1995), 'Annealing Markov chain Monte Carlo with applications to ancestral inference', *Journal of the American Statistical Association* **90**(431), 909–920.

Gilks, W. R., Richardson, S. & Spiegelhalter, D. J. (1996), *Markov Chain Monte Carlo in Practice*, Chapman & Hall.

Hegyi, F. (1974), A simulation model for managing Jack-pine stands, *in* J. Fries, ed., 'Growth Models for Tree and Stand Simulation', number 30 *in* 'Royal College of Forestry, Sweden, Research Notes', pp. 74–90. (Skogshögskolan, Rapporter och Uppsatser).

Henriksen, H. A. (1981), *Træmåling*. Teaching notes.

Johannsen, V. K. (1996), Statistical aspects of a spatial single-tree growth model for even-aged oak stands in Denmark, Dynamic Growth Models for Danish Forest Tree Species Working Paper No. 10, Danish Forest and Landscape Research Institute, Hørsholm.

Larsen, M. & Rudemo, M. (1998), 'Optimizing templates for finding trees in aerial photographs', *Pattern Recognition Letters* **19**, 1153–1162.

Lund, J. (1998), Survival of the fattest? Self-thinning among trees, Report 3, The Royal Veterinary and Agricultural University, Department of Mathematics and Physics. Course report from PhD course in forest biometrics.

Lund, J., Penttinen, A. & Rudemo, M. (1999), Bayesian analysis of spatial point patters from noisy

observations, Preprint, Department of Mathematics and Physics, The Royal Veterinary and Agricultural University. Submitted. Available from `ftp://ftp.dina.kvl.dk/pub/Staff/Jens.Lund/recon`

Lund, J. & Rudemo, M. (1999), 'Models for point processes observed with noise', *Biometrika* . To appear.

Lund, J. & Thönnes, E. (1999*a*), Perfect adaptive Metropolis-Hastings simulation for point processes, preliminary manuscript.

Lund, J. & Thönnes, E. (1999*b*), Perfect sampling from point patterns observed with noise, manuscript.

Lund, J. & Thönnes, E. (1999*c*), Perfect simulation of point processes given noisy observations, manuscript. Submitted.

Pukkala, T. (1989), 'Methods to describe the competition process in a tree stand', *Scandinavian Journal of Forest Research* **4**, 187–202.

Rathbun, S. L. & Cressie, N. (1994), 'A space-time survival point process for a longleaf pine forest in Southern Georgia', *Journal of the American Statistical Association* **89**(428), 1164–1174.

Scheike, T. H. & Jensen, T. K. (1995), A discrete survival model with random effects: An application to time to pregnancy, Research Report 95/11, Department of Biostatistics, University of Copenhagen.

Skovsgaard, J. P. (1997), Tyndingsfri drift af sitkagran. En analyse af bevoksningsstruktur og vedmasseproduktion i utyndede bevoksninger af sitkagran (*Picea sitchensis* (Bong.) Carr.) i Danmark, Forskningsserien nr. 19, Forskningscentret for Skov & Landskab, Hørsholm. 525 pp.

Tomé, M. & Burkhart, H. E. (1989), 'Distance-dependent competition measures for predicting growth of individual trees', *Forest Science* **35**(3), 816–831.