When the wrong model is best: Confounding and confusion in individual-tree models

Oscar Garcia

UNBC UNIVERSITY OF NORTHERN BRITISH COLUMBIA

forestgrowth.unbc.ca

2014 Western Mensurationists Meeting
June 22-24, Monterey, CA

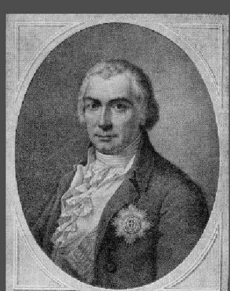
Will talk about one of the things that has always bother me about individual-tree growth models, having to do with statistical confounding between size and growth rate.
Does not seem to worry anybody else, so most likely the confusion is only in my own mind. Still, would be interested in your thoughts.

Outline

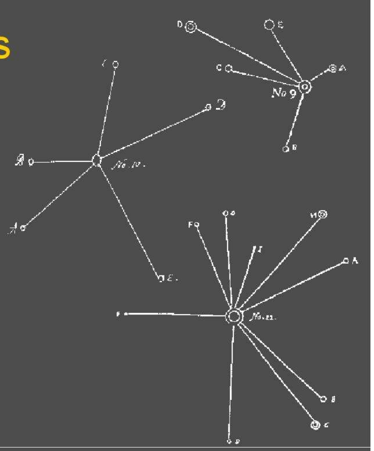
- Individual-tree growth models
  - Pairwise interactions
  - Fully spatial
  - Aspatial, aka distance-independent
- Models comparison: an example
- Growth and size
- Implications, alternatives?
- Comments?

First, a brief review, to be on the same page. Followed by a model comparison study done with PhD student Min Jun Lee. And then the confounding/confusion issues.

Spatial models

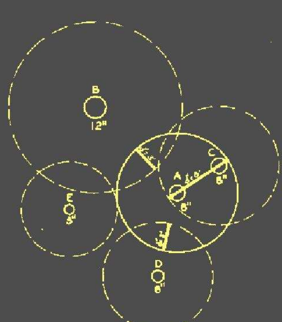


C.D.F. Reventlow
1748 - 1827



Count Reventlow dabbled in forestry, in-between managing the Danish Crown finances and distributing land to the peasants in the agrarian reform.
Aroud 1800 he used spatial individual-tree modelsto somehow produce his yied tables.

Staebler 1951 – ZOI overlaps



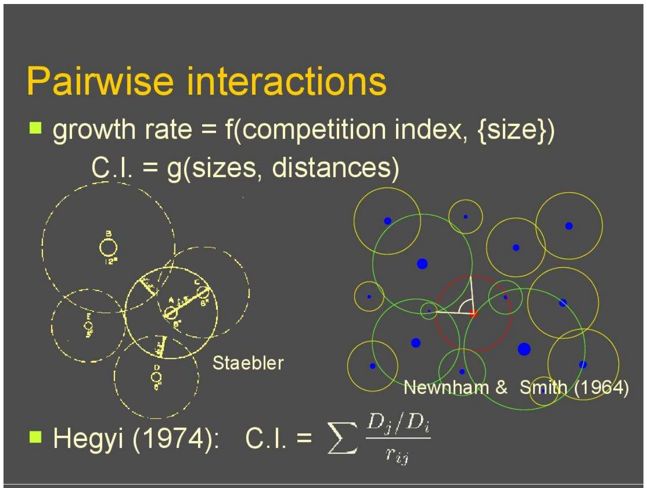
Growth and Spacing
in an
Even-aged Stand of Douglas-Fir

George R. Staebler

Thesis submitted in partial fulfillment of the requirements for the degree of Master of Forestry, School of Natural Resources, University of Michigan.

May 1951

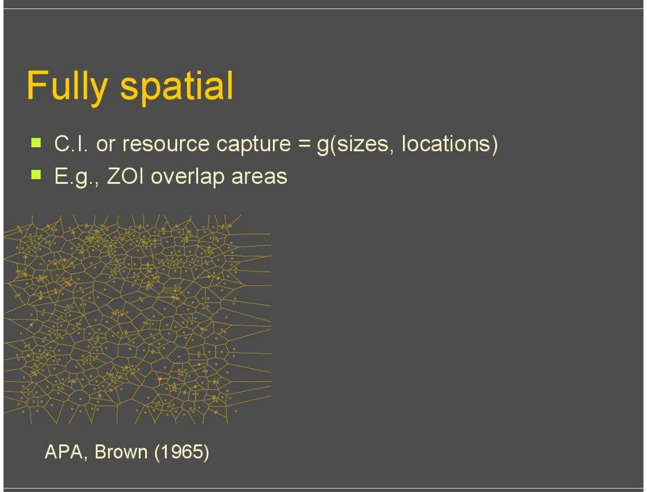
Fast-forward to 1951. Staebler introduces the idea of overlapping zones of influence (ZOI), used in many of the later models.
Using paper and pencil, so could not get very far. The subject exploded in the 1960's and 70's with the availability of digital computers.



Effect of neighbors usually encapsulated in a competition index. Or a resource capture or assimilation index.

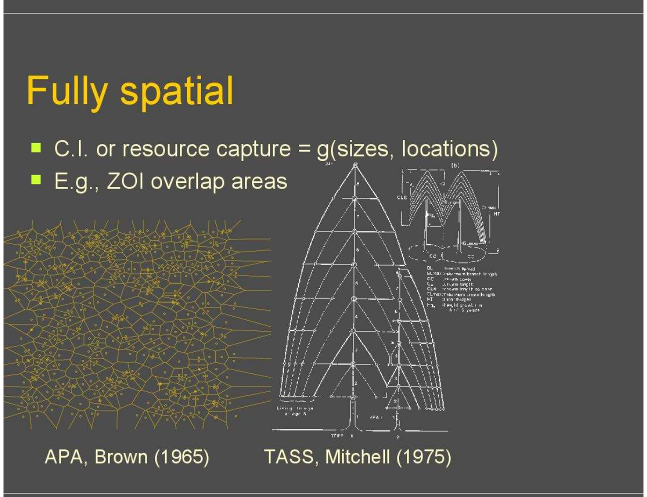
"Semi-spatial" models simplify by considering only interactions among pairs of trees. Only distances; azimuth and three-way interactions are ignored.

E.g., Staebler used the width of the ZOI overlaps, Newnham the angles, others functions of distances and sizes.

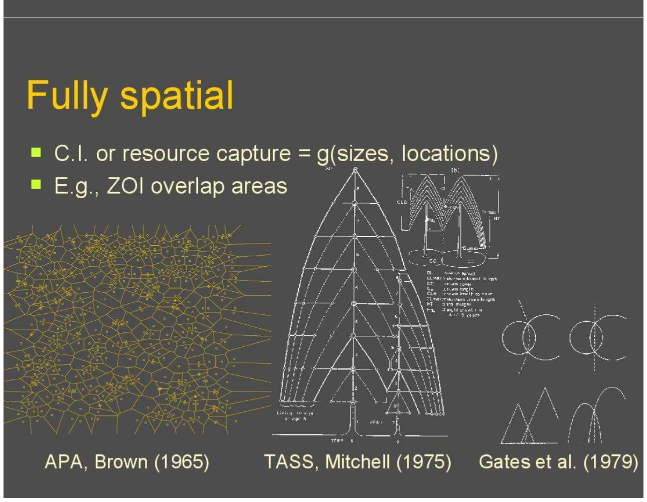


"Fully spatial" models depend on the configuration of neighbor locations.

E.g., functions of the overlap areas (including 3-at-a-time), and Brown's area potentially available (APA).



More complicated, Mitchell's TASS model. Crown surface moves upward through branch growth, preserving shape. Crown width expansion stops on contact with neighbors. Essentially, this induces a horizontal space partition similar to the APA (see small top-right figure, from Mitchell 1975).



Gates et al looked at relationships between 3-D and 2-D intersections.

Distance-independent models

- Usually locations and indices contributed little to prediction
- Aspatial growth = f(size) just as good
- Spatial abandoned for operational use in the 1980's
- Why?

Generally found that the spatial modelling does not help in growth and yield prediction. Distance-independent models now used for management; TASS is an (the?) exception.
Why? Biologically spatial models seem to make a lot of sense.

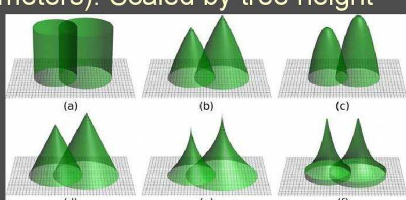
Lee & García (in prep.)

- Spruce–hardwoods mixtures in BC Central Interior
- Control and 2 brushing treatments, 50 x 50 m stem-mapped plots
- Estimate tree volume increment with 3 model classes:
  1. Fully spatial (several alternatives)
  2. Perfect plasticity approximations (PPA, aspatial)
  3. Empirical aspatial $\Delta v = f(v)$

A modelling and simulation study in mixwoods.
Model class explanations follow.

Spatial: `siplab` R package

- Influence functions (with species-dependent parameters). Scaled by tree height

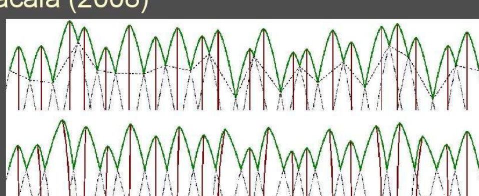


- Overlap partitioning rules
- Efficiency distance weighting

Fully spatial models implemented in siplab. Siplab unifies and generalizes modelling approaches, computing resource capture indices in three stages:
(1) An "influence function" represents crown extent as in TASS, or a more abstract shading potential or competitive pressure (above- and/or below-ground).
(2) In the overlaps the most dominant tree takes all, as in TASS and APA (one-sided competition), or the resource may be shared according to a less extreme asymmetry parameter.
(3) Resource contributions may diminish with distance, e.g., because of increased overheads from branch growth and maintenance.
Cylinders in (a) correspond to simple ZOI overlap models, (b) are the shapes used in TASS.

Perfect plasticity

- Strigul, Pristinski, Purves, Dushoff and Pacala (2008)



- Explicit multi-cohort model, one-sided comp.
- Empirical polynomial in tree height

The second model class uses a "perfect plasticity approximation" (PPA). Contrary to the conventional assumption of rigid tree shapes, trees can lean and distort their crown (and root) development to occupy less contested spaces. The PPA ignores any limits to these adjustments, balancing competition intensity on all sides. The model becomes aspatial, analogous to a soap froth where bubbles move to reach equilibrium with their neighbors. Perfect plasticity may be a better assumption than no plasticity.
We used to variants: (a) A theoretical model that assumes one-sided competition and influence functions based on Gates et al (1979) (with species-dependent parameters). (b) A more general empirical function of tree height.

Results

Over-all R^2

- Best spatial: 0.62
- One-sided PPA:
- Polynomial PPA:
- $\Delta v = b_0 + b_1 v$:

I show an over-all tree volume increment r-squared for all species and all the 3 plots.
We tried two influence functions and two competition asymmetry alternatives.

Results

Over-all R^2

- Best spatial: 0.62
- One-sided PPA: 0.75
- Polynomial PPA: 0.75
- $\Delta v = b_0 + b_1 v$:

Both PPA variants were similar, and better than the spatial model. Support for the PPA.

Results

Over-all R^2

- Best spatial: 0.62
- One-sided PPA: 0.75
- Polynomial PPA: 0.75
- $\Delta v = b_0 + b_1 v$: 0.86

A simple linear regression over current tree volume was best.

Results

Over-all R^2

- Best spatial: 0.62
- One-sided PPA: 0.75
- Polynomial PPA: 0.75
- $\Delta v = b_0 + b_1 v$: 0.86

To be expected?

Models typically include the same variable on both sides, maybe embedded in a C.I.

Surprising? Actually, most studies have shown little or no improvements of more complex models over simply using tree size as a predictor.
Most models use dbh or tree basal area instead of volume (which introduces other issues), but the basics are the same.

Correlation is not causation

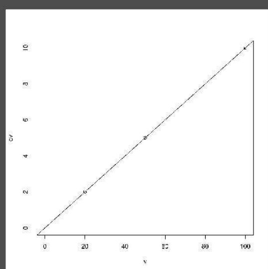
- Assume a true model $\Delta v_i = \beta_i$

That is, growth is unrelated to size, but varies among trees due to genetics, microsite, competition, etc.

- Take 3 trees, with $\Delta v_i = \beta_i = 2, 5,$ and 10 dm^3/yr
- The volumes at age 10 are $v_i = 20, 50,$ and $100\ dm^3$
- Plot Δv_i over v_i :

A toy example may help.

... Correlation is not causation



- The "wrong" model $\Delta v = 0.1\ v$ (or $\Delta v = v / t$) produces perfect predictions

So what?

- Bigger trees may or may not grow faster. Faster-growing trees tend to be bigger. Confounding, how to disentangle?

We have it back to front,

So what?

- Bigger trees may or may not grow faster. Faster-growing trees tend to be bigger. Confounding, how to disentangle?
- Does it matter? It depends:
  - In unmanaged undisturbed stands the growth-size regression is a good (best?) predictor
  - Management or disturbances can break the correlation

Essentially an empirical extrapolation of past growth.

What to do?

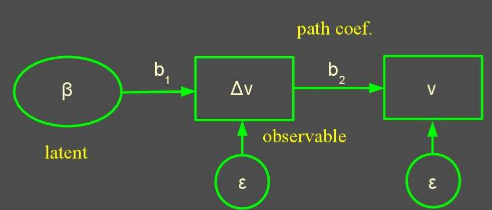
- Mixed effects? With multiple measurements one can try estimating the β_i

  Still not as “good” as the wrong model.

With additional information the β can be modelled as random effects. Some improvement over Δv = constant, but...

... What to do?

- Path Analysis (Sewall Wright, 1918, 21, 25)
- Structural Equation Modelling (SEM)



Might help for understanding what is going on, but perhaps not solve the fundamental problem.

... What to do?

- Whole-stand models

  For complex stands PPA approaches seem promising
- ...?

Avoid individual-tree modelling altogether?

PPA concepts may be used to produce good cohort (species and/or age-class) whole-stand models.

Other ideas?

forestgrowth.unbc.ca/siplab