Plant individual-based modelling: More than meets the eye

Oscar García

UNBC UNIVERSITY OF NORTHERN BRITISH COLUMBIA

forestgrowth.unbc.ca

2008 World Conference on Natural Resource Modeling

Warsaw, Poland, 15-18 June 2008

http://rma2008.wl.sggw.pl

Outline

- Individual-based models (IBMs)
- Interactions and spatial structure
- From individuals to distributions
- Why size matters
- Conclusions

Individual-based models (IBMs)

Conceptually simple. Easier to understand individuals than more abstract aggregates.

Arithmetic replaces math.

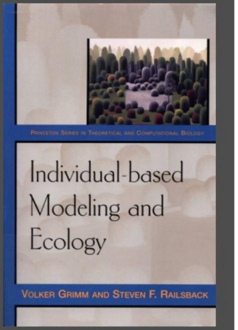
Computer intensive (often an advantage).

Popular, productive (publication-wise). As in Neural Networks or optimization heuristics, anything goes.

Useful, but some limitations. Apparent simplicity can be misleading.

IBMs in Ecology

- Grimm & Railsback (2005)
- Started around 1990
- Agent-based models in Economics, etc.
- "... less productive than expected..." "... may be ripe for a backlash ..."
- Plants: local interactions, size



Individual-based Modeling and Ecology

Volker Grimm and Steven F. Railsback

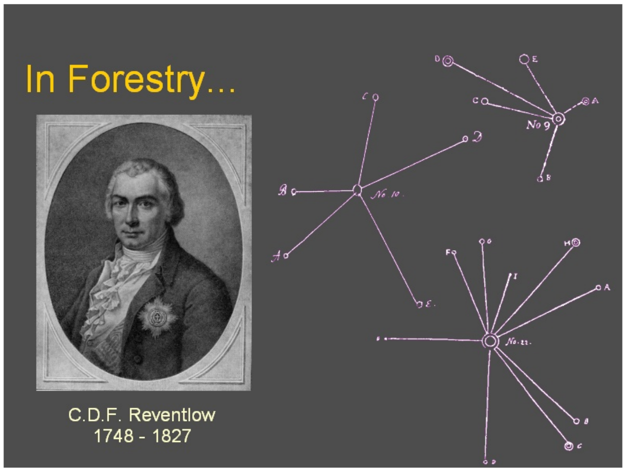
Comprehensive coverage, no need to add much here.

Book: "the last decade", or the "last 1 ½ decades". Grimm (1999) "Ten years of individual-based modelling in Ecology..."

In other fields at about the same time, jargon varies. References to forestry are rare.

Initial wild enthusiasm, although some disenchantment lately (quotes from Grimm & Railsback's Preface).

Unlike in other areas such as population ecology, for plants, and especially trees, numbers are not sufficient, size is important. Also, "mixing" assumptions are questionable.



In Forestry...

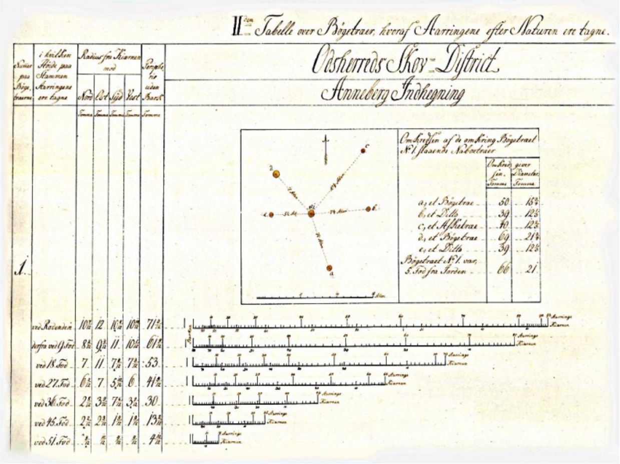
C.D.F. Reventlow
1748 - 1827

Models have always been important in forestry, due to obvious difficulties in learning from experience or experimentation.

Count Reventlow, Lord Chancellor of Denmark around 1800, keen interest in forestry.

Mapping of locations and sizes of competing trees.

Perhaps not quite an IBM, although the ideas were used in the development of his yield tables.



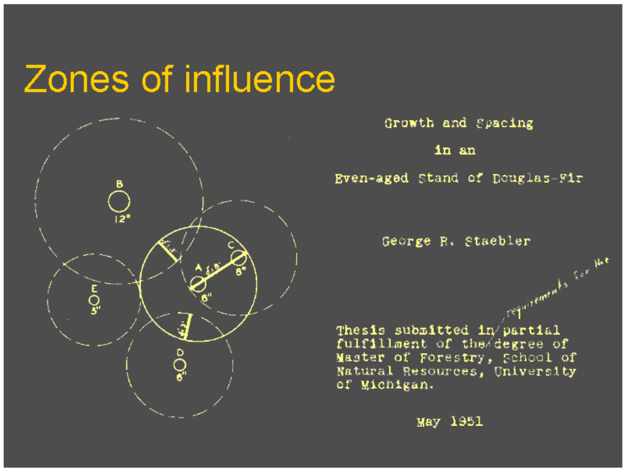
Reventlow, ca. 1800 (courtesy of JP Skovsgaard).

Sizes and distances to neighbours, radial growth ring measurements in 4 directions and at various heights.

He also reconstructed height growth through full stem analysis, often thought as a modern development.

Among other things, contributed also to financial analysis (compound interest, discounting), similar to later work by Faustmann (1848). Optimizing wood production, in particular for ship building.

The oaks he planted were ready to be used by the Navy around the time of the 2nd World War. Yes, foresters are funny people!



Zones of influence

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.

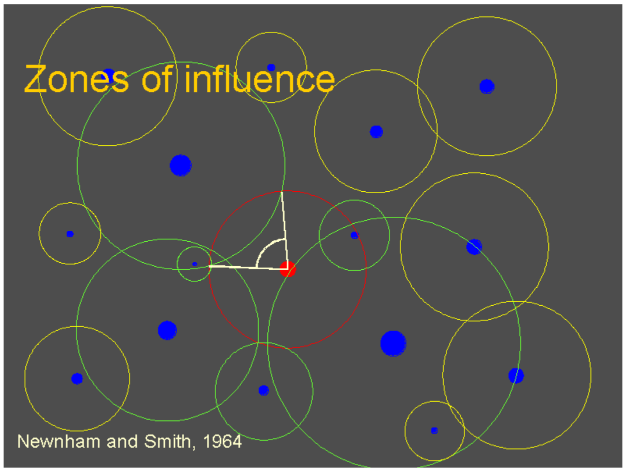
May 1951

Staebler, 1951, in North America, which is what really counts ;-)

Zone of influence concept, used ever since.

Formula for overlap area is a bit messy, so he used overlap width.

No computers, hand calculations did not get him very far.

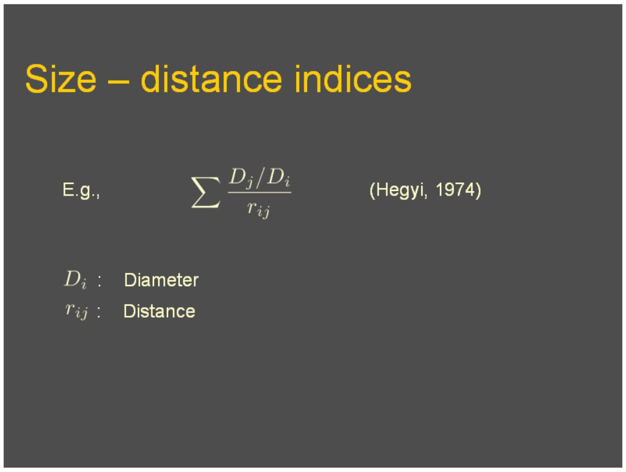


Zones of influence

Newnham and Smith, 1964

Now with computers. Prof. JHG Smith and students, at the University of British Columbia.

Competition index: sum of overlap angles.

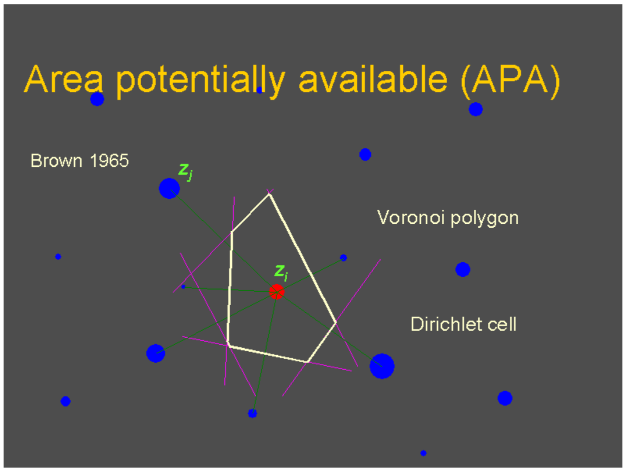


Size – distance indices

E.g., $\sum \frac{D_j/D_i}{r_{ij}}$ (Hegyi, 1974)

D_i : Diameter
r_{ij} : Distance

Another class of competition indices, combinations of sizes and distances



Area potentially available (APA)

Brown 1965

Voronoi polygon

Dirichlet cell

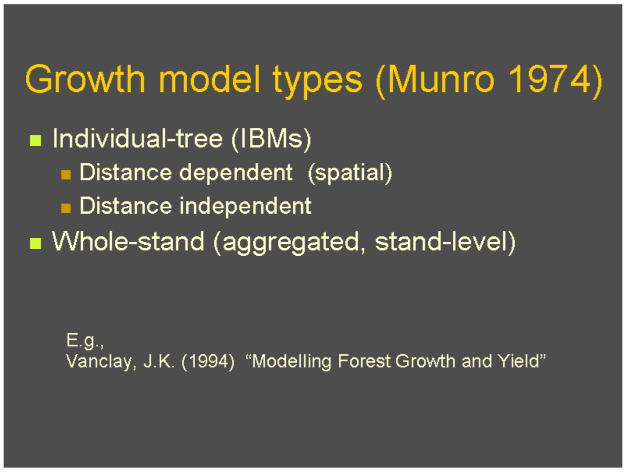
And another one. Brown (New Zealand) 1965, Mead (UK, annual plants) 1966.

Jack (1967) recognized tree APAs as the Dirichlet cells or Voronoi polygons in math and computer science.

APA

Jack Pine
Manitoba

Dirichlet
Voronoi
Thiessen



Growth model types (Munro 1974)

- Individual-tree (IBMs)
  - Distance dependent (spatial)
  - Distance independent
- Whole-stand (aggregated, stand-level)

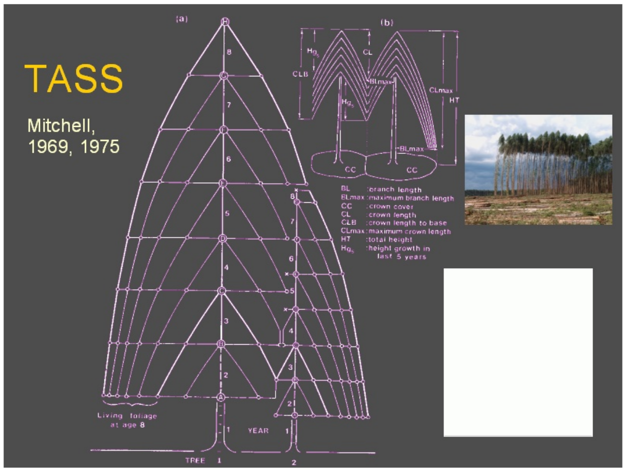
E.g.,
Vanclay, J.K. (1994) "Modelling Forest Growth and Yield"

Munro's classification. Derived in part from Goulding's 1968 UBC thesis (probably the first distance-independent model).

For more on forest growth models see Vanclay's book. Electronic copies available at http://jkv.5megs.com or http://web.unbc.ca/~garcia/FSTY405/book.html

Spatial models were all the rage in the 1970's. Ideal thesis topic: guaranteed some result at the end, and nobody could prove you wrong!

Largely abandoned in favor of distance independent models in the 1980's. Spatial indices often found to have poor predictive power (possible reasons later).



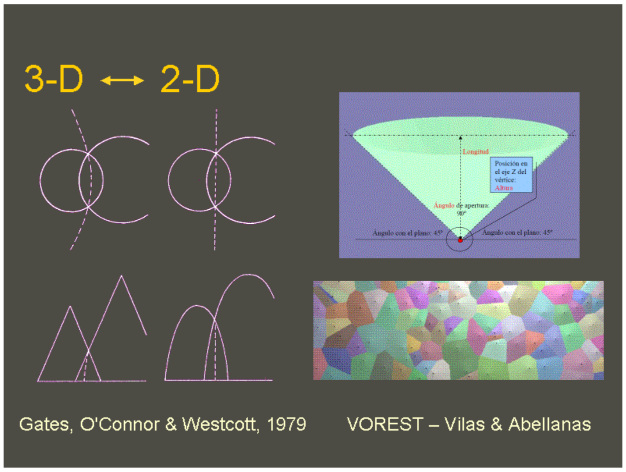
TASS

Mitchell,
1969, 1975

One of the few (only?) spatial models still being used for practical forest management. Also one of the most complex (3-dimensional). British Columbia.

Atypical in being driven by height and crown development, rather than by stem diameter. Wood accumulation is a “side effect”, as it should.

Note in the top figure that on horizontal projection it resembles ordinary 2-D models.



3-D ⟷ 2-D

Gates, O'Connor & Westcott, 1979

VOREST – Vilas & Abellanas

Left: 3-D to 2-D relationships investigated by Gates et al.

Right: Similar relationships in generalized Voronoi polygons, resulting from the intersection of cones or other figures (see http://www.sciencedaily.com/releases/2007/11/071117111207.htm).

Of course, to mathematicians upside-down trees make no difference!

General

z_i : size

p_i : location (coordinates)

$$\Delta z_i = f(z_1, p_1, z_2, p_2, \dots)$$

$$\Delta z_i = f(z_i, c_i)$$

increment

$$c_i = g(z_i, p_i, z_1, p_1, z_2, p_2, \dots)$$

competition index

- “Size”
  - Stem diameter a causal variable?
  - Plasticity

For convenience, competitor effects are collected in a “competition index”. No coordinates in distance-independent models.

A single variable (diameter or biomass) used as both cause and consequence.

Size is usually stem diameter. But, how does accumulated mostly dead xylem affect growth?

Unlike fish. mammals, and some plants, the shape of trees can vary tremendously. Many combinations of diameter, height, and crown dimensions are possible. Under certain natural stand conditions these dimensions can be well correlated, but this is not so under management alternatives. E.g., thinning aims to produce different diameters for a same height. A single scalar measure of size may be inadequate, vector generalizations may be needed.

Interactions

Sources of spatial correlation

- Competition
- Facilitation
- Micro-site

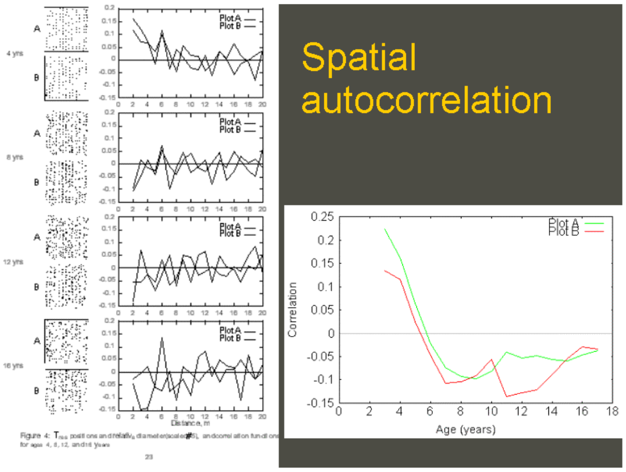
With few exceptions, modelling has focused exclusively on competition.

Competition induces negative spatial size correlations: the neighbor of a large tree would tend to be smaller than average. And vice-versa.

However, there are other sources of spatial structure that have generally been neglected. All the ones I can think of induce positive size correlations, counteracting the effect of competition.

One is facilitation, e.g., mutual sheltering from wind, or favorable alterations of the micro-climate. Perhaps only important at extreme stand densities.

Short-range variations of soil fertility, depth, moisture (micro-site), should cause positive size correlations: Neighboring trees tend to share more similar growing conditions than trees further apart.



Spatial autocorrelation

Figure 4: Tree positions and relative diameter (scaled #5), and correlation functions for ages 4, 8, 12, and 16 years

Two unthinned sample plots of planted radiata pine measured over a number of years. Left: diameter correlations over inter-tree distance, for various ages.

In a young 4-years-old stand competition has not yet started. Positive correlation decreasing with distance, as would be expected from micro-site alone.

At age 8, competition has started, and short-range correlations have become negative. The effect of micro-site is still present, but overridden by that of competition.

Later, mortality confuses things somewhat. The graph on the right shows the 2-meter correlation over time.

Sources of spatial correlation

- Competition
- Facilitation
- Micro-site

In natural stands:

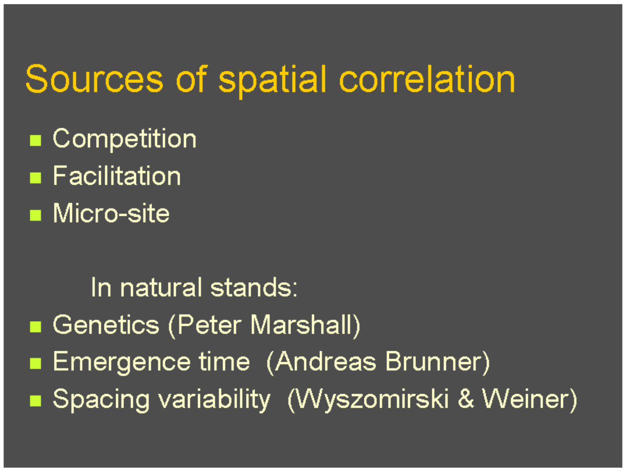
- Genetics (Peter Marshall)

Those effects occur in forest plantations. In natural stands there can be additional sources of spatial structure, that have been pointed out to me by various people (my past experience has been mostly with plantations, or at least that is my excuse).

In some species that regenerate vegetatively, such as aspen, neighbors tend to be genetically related, contributing to positive size correlations.

Patches belonging to a same clone can often be distinguished by slightly different Fall colors (photo by R. Kabzems).

A similar effect might occur in species with heavy seeds, with limited spatial dispersion, such as oaks.



Sources of spatial correlation

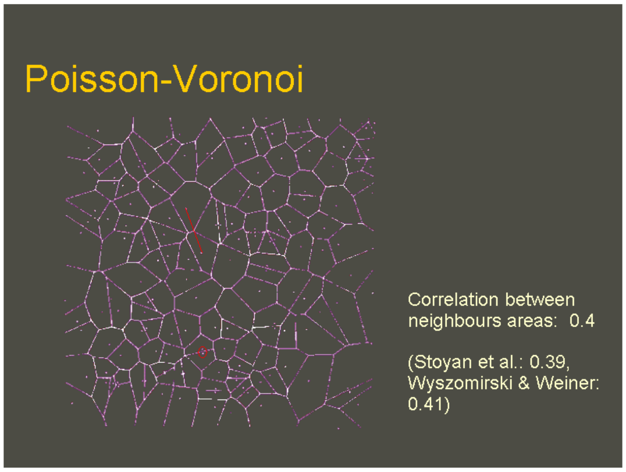
- Competition
- Facilitation
- Micro-site

In natural stands:

- Genetics (Peter Marshall)
- Emergence time (Andreas Brunner)
- Spacing variability (Wyszomirski & Weiner)

Germination and seedling development in natural stands does not occur simultaneously. If the timing is affected by spatially-correlated micro-site conditions, this will induce positive spatial size correlations. Related to micro-site, but affecting directly only the early growth.

The final one was discovered by Tomasz Wyszomirski, University of Warsaw; a paper with Jacob Weiner is under review. They talk about competition, but I see it as a purely geometric effect.



Poisson-Voronoi

Correlation between neighbours areas: 0.4

(Stoyan et al.: 0.39, Wyszomirski & Weiner: 0.41)

To examine the pure spacing variability effect in isolation, consider a completely random (Poisson) spatial distribution of tree seedlings. Two neighboring trees that happen to be closer than average (red circle), will both tend to have APAs smaller than average. Conversely, the growing space for two neighbors that are more distant than average (red line), will tend to be larger than average (for both). Therefore, the growing space of neighbors is positively correlated. Other things being equal, this will tend to make neighbor sizes to be positively correlated.

It does not seem easy to work out the correlations analytically, but Wyszormirski and Weiner found by simulation a correlation coefficient of 0.41 for neighboring areas in the Poisson-Voronoi diagram. Okabe et al (2002) in THE book on Voronoi polygons and related stuff mention a simulation by Stoyan et al (1989) giving 0.39.

Size distributions

Individuals → Distributions

- “Distribution modifying function”
- Distance-independent, deterministic:
  - Physics: Lagrangian to Eulerian view
  - Liouville + modifications for mortality
- Stochastic, no interactions:
  - Kolmogorov / Fokker-Plank

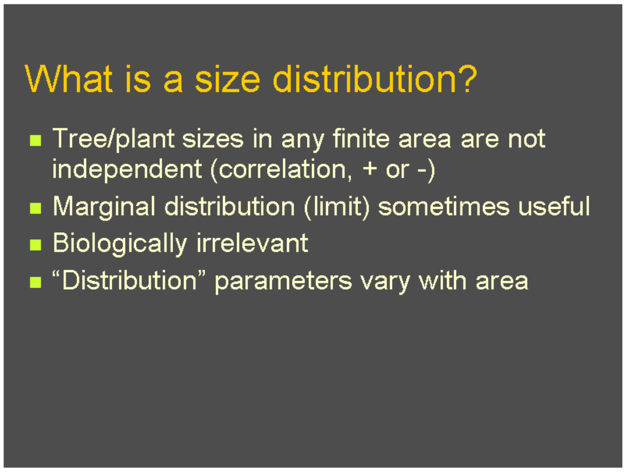
(E.g., Picard & Franc, 2004: www.FBMIS.info)

- Spatial??

The relationship between IBMs and the size distributions that they generate has for long been a source of speculation, mystery, and confusion. The term “distribution modifying function” has been coined for the individual growth rate function effect on the size distribution. There are several partially correct or completely wrong derivations both in the ecology and the forestry literatures (no names!)

Actually, there is a well-established mathematical theory for the distance-independent, deterministic case. And for the stochastic case, at least if interactions can be ignored. For pointers see, e.g., Picard and Franc.

With local interactions (spatial, distance-dependent models) the problem is much more difficult. Or impossible...



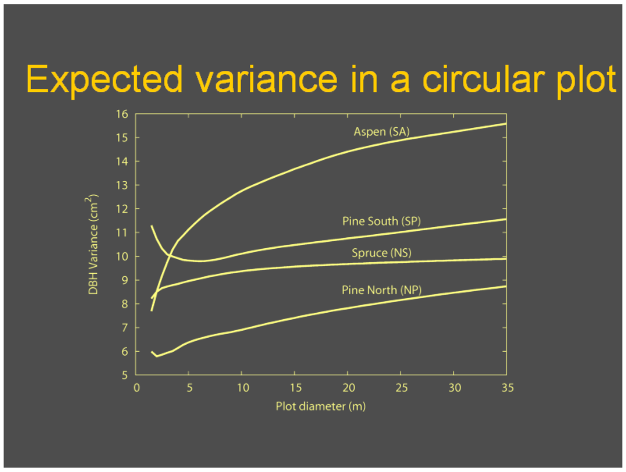
What is a size distribution?

- Tree/plant sizes in any finite area are not independent (correlation, + or -)
- Marginal distribution (limit) sometimes useful
- Biologically irrelevant
- “Distribution” parameters vary with area

Problem is, plant sizes are not independent, so that a size distribution is not well defined.

More precisely, over any finite area there is a joint distribution, and the only scalar distribution that makes sense is the marginal (which is also the limit for an infinitely large stand). The marginal is close to what managers want for planning. But biologically, what matters is what happens in small neighborhoods.

If we insist in fitting a distribution (a popular pastime), the results depend on the plot size.



Expected variance in a circular plot

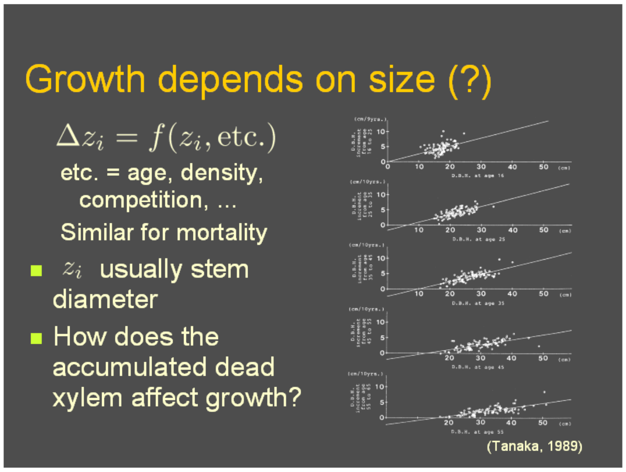
Examples of diameter variance for different plot sizes (Canadian J. Forest Research 36: 2983-2993, 2006).

Positive size correlations cause the variance in small plots to be lower than in larger plots.

Current distance-dependent IBMs predict negative correlations, distance-independent ones ignore them.

Why size matters

… in growth predictions.



Growth depends on size (?)

$$\Delta z_i = f(z_i, \text{etc.})$$

etc. = age, density, competition, …

Similar for mortality

- z_i usually stem diameter
- How does the accumulated dead xylem affect growth?

(Tanaka, 1989)

Graphs: 10-year diameter increment vs. current diameter, for various ages.

Size is usually the most significant increment predictor.

More generally…

- Do trees grow faster because they are larger?
- Or are they larger because they have been growing faster?
- $z_i \longrightarrow \Delta z_i$ or $\Delta z_i \longrightarrow z_i = \sum \Delta z_i$?

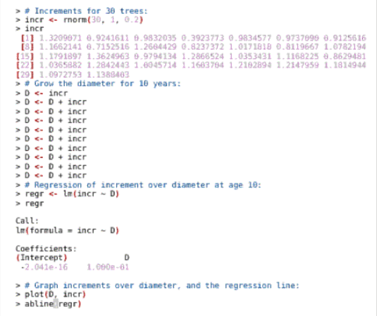
Is it possible that we could have some circularity problem?
Naaah! People have been doing this for ages!

Example

- Trees grow at 1 cm / year, on average
- Genetic / micro-site variation of 20%
- Ignore year-to-year, within-year, and measurement error noise
- I.e., $\Delta D_i = 1 + \epsilon_i, \qquad \epsilon_i \sim \text{Normal}(0, 0.2)$

Minimalist example, for transparency.
We could think it through, but we have computers!

Simulate in R

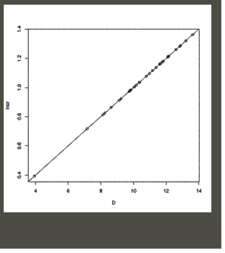


```
> # Increments for 30 trees:
> incr <- rnorm(30, 1, 0.2)
> incr
 [1] 1.3209071 0.9241611 0.9832035 0.3923773 0.9834577 0.9737090 0.9125616
 [8] 1.1662141 0.7152516 1.2604429 0.8237372 1.0171818 0.8119667 1.0782194
[15] 1.1791897 1.3624963 0.9794134 1.2866524 1.0353431 1.1168225 0.8629481
[22] 1.0365882 1.2842443 1.0045714 1.1603704 1.2102894 1.2147959 1.1814944
[29] 1.0972753 1.1388403
> # Grow the diameter for 10 years:
> D <- incr
> D <- D + incr
> D <- D + incr
> D <- D + incr
> D <- D + incr
> D <- D + incr
> D <- D + incr
> D <- D + incr
> D <- D + incr
> D <- D + incr
> # Regression of increment over diameter at age 10:
> regr <- lm(incr ~ D)
> regr

Call:
lm(formula = incr ~ D)

Coefficients:
(Intercept)            D
 -2.041e-16    1.000e-01

> # Graph increments over diameter, and the regression line:
> plot(D, incr)
> abline(regr)
```



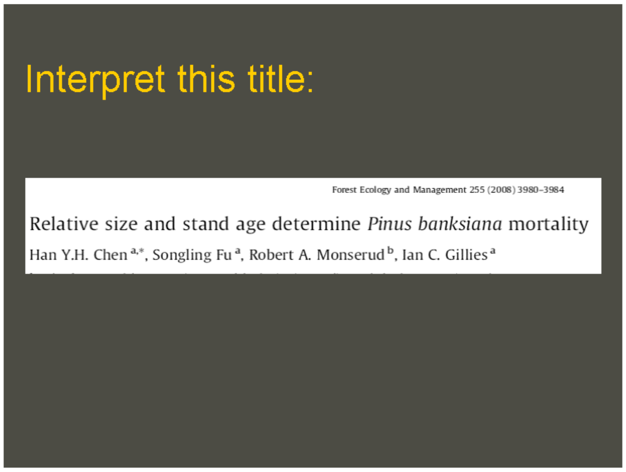
Steps:
- Increments for 30 trees, from a Normal with mean 1 and standard deviation 0.2
- Accumulate increments for 10 years, to get the diameters at age 10 (no shortcuts, the pedestrian way, as we would do it in a more complicated example)
- Regression of increment over diameter
- Plot the data and regression, increment vs. diameter

Not bad, R-squared 1.0!

So?

- The regression of increment on size was highly significant (perfect, in fact)
- Although in the true model they are (causally) independent
- We could recognize a mixed effects model, and give the "correct" prediction *incr* = *b*, with *b* ≈ 1 and a standard error of about 0.2
- But the "wrong" model *incr* = *b* * *D* / *Age* is much better for prediction!

In time-series forecasting competitions, often simplistic methods based on local extrapolation outperform sophisticated process-based models.



Interpret this title:

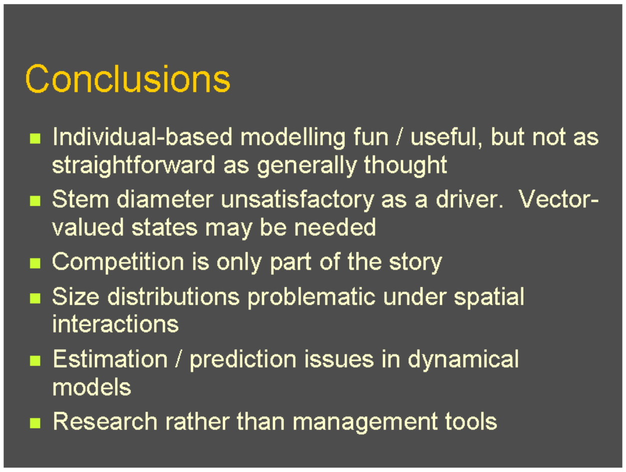
Forest Ecology and Management 255 (2008) 3980–3984

Relative size and stand age determine *Pinus banksiana* mortality

Han Y.H. Chen[a,*], Songling Fu[a], Robert A. Monserud[b], Ian C. Gillies[a]

What exactly does “determine” mean here?

Finally…



Conclusions

- Individual-based modelling fun / useful, but not as straightforward as generally thought
- Stem diameter unsatisfactory as a driver. Vector-valued states may be needed
- Competition is only part of the story
- Size distributions problematic under spatial interactions
- Estimation / prediction issues in dynamical models
- Research rather than management tools

In scientific research, models are pushed until they break. When that happens one should rejoice; knowledge is generated. Then the model is modified, and pushed until it breaks in some other way. On the other hand, managers tend to get upset when the model breaks.