Comparing four European stand growth models for Douglas fir

Thoughts on some methodological challenges

Minna Pulkkinen, Holger Wernsdörfer & François Ningre

AgroParisTech-INRA, Laboratoire d'Etude des Ressources Forêt-Bois (LERFoB)
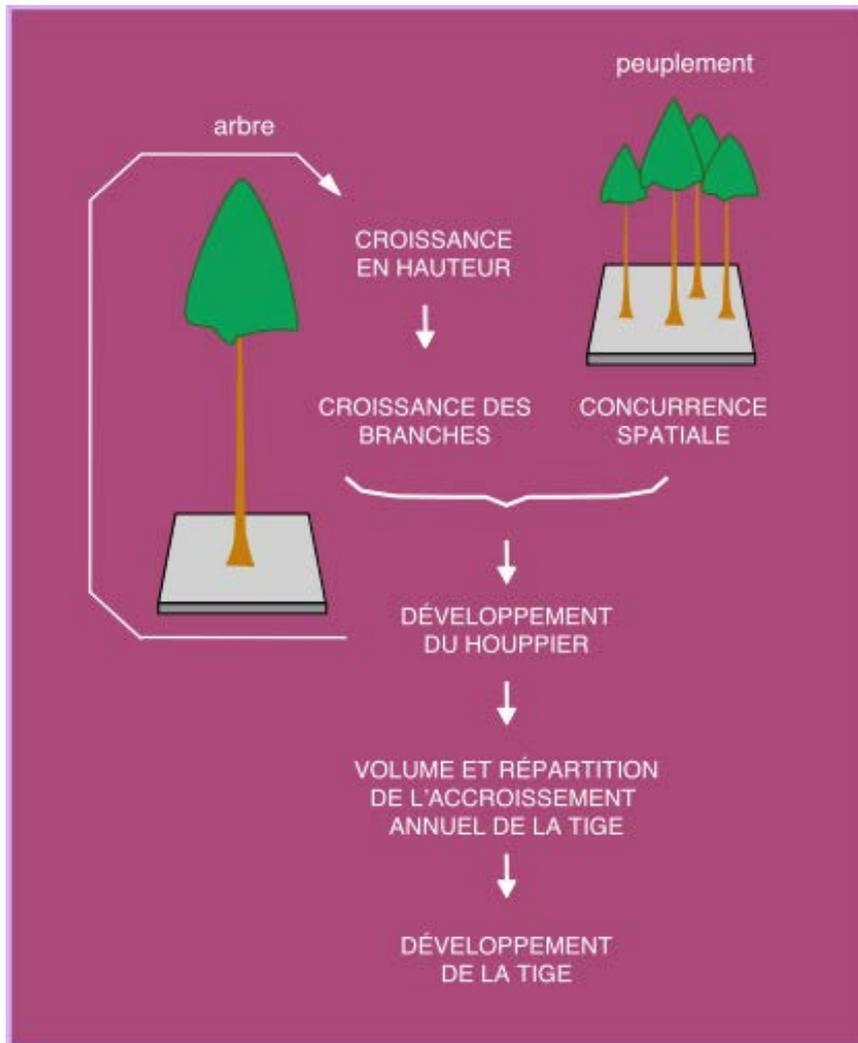
Journées CAQSIS, Gembloux, 8–10/4/2014

1

Background

- Post-doc project "Assessment of new forest production systems for Douglas fir – towards a simulation tool linking research, development and teaching"

- Aims of the project
  1. Assessment of existing growth models for Douglas fir in view of realistic simulation of new silvicultural management scenarios in France
  2. Simulation and comparison of new and conventional scenarios in terms of timber (quantity & quality) and energy biomass production
  3. Development and implementation of a teaching tool in Capsis

- *Key question*: how well are growth models able to simulate new silvicultural management scenarios?

- *Model assessment against observed data*: if a model is able to reproduce realised contrasting silvicultural management scenarios, it is likely to be able to simulate new scenarios as well

- *Data for model assessment*: field experiments representing contrasting silvicultural management scenarios

- *Sensitivity analysis*: identifying relations and phases of development critical for model performance
  $\rightarrow$ model improvement

- Steps of model assessment in the project
  1. Compilation of data from field experiments with (i) highly contrasting initial densities and (ii) varied timings and intensities of thinnings
  2. Comparison of simulation results of 4 European growth models for Douglas fir with field observations
     - SimCoP (LERFoB)
     - BWinPro (FVA Freiburg)
     - Gymnos (Université de Liège Gembloux)
     - Oasis (FCBA)
  3. Sensitivity analysis on the most promising model(s) in view of simulation of new scenarios
  4. Improvement the most critical relations revealed by the sensitivity analysis (beyond the project)

Models

arbre
peuplement

CROISSANCE EN HAUTEUR

CROISSANCE DES BRANCHES

CONCURRENCE SPATIALE

DÉVELOPPEMENT DU HOUPPIER

VOLUME ET RÉPARTITION DE L'ACCROISSEMENT ANNUEL DE LA TIGE
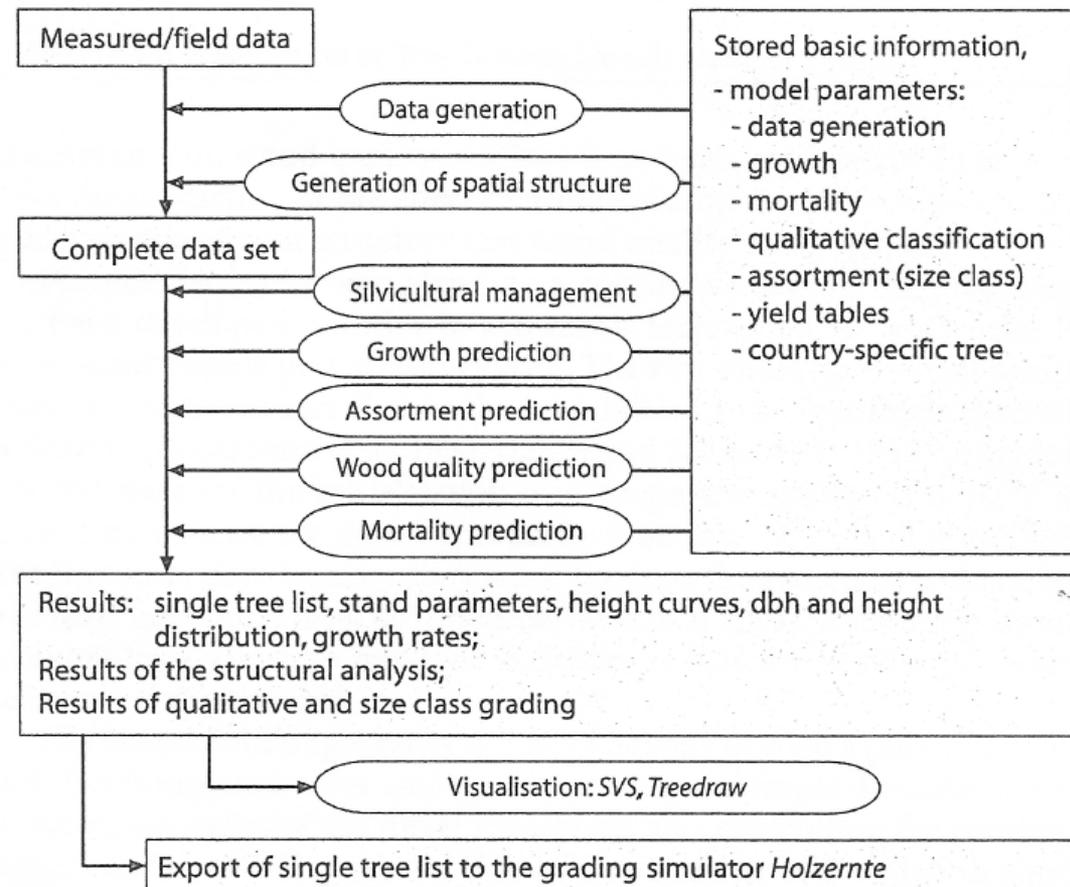
DÉVELOPPEMENT DE LA TIGE

SimCop

- Ottorini (1991, 1995); based on work by Mitchell (1975)
- For even-aged pure Douglas fir stands
- Empirical, spatially explicit
- Basic idea: height growth
  $\rightarrow$ crown growth (restricted by neighbouring crowns)
  $\rightarrow$ "effective" foliar volume
  $\rightarrow$ stem growth
- Mortality with (i) suppression by neighbouring trees and (ii) Reineke's self-thinning rule
- Initial state: plot size, spacing & vigours ($\rightarrow$ heights) of trees; H_{dom} vs. age
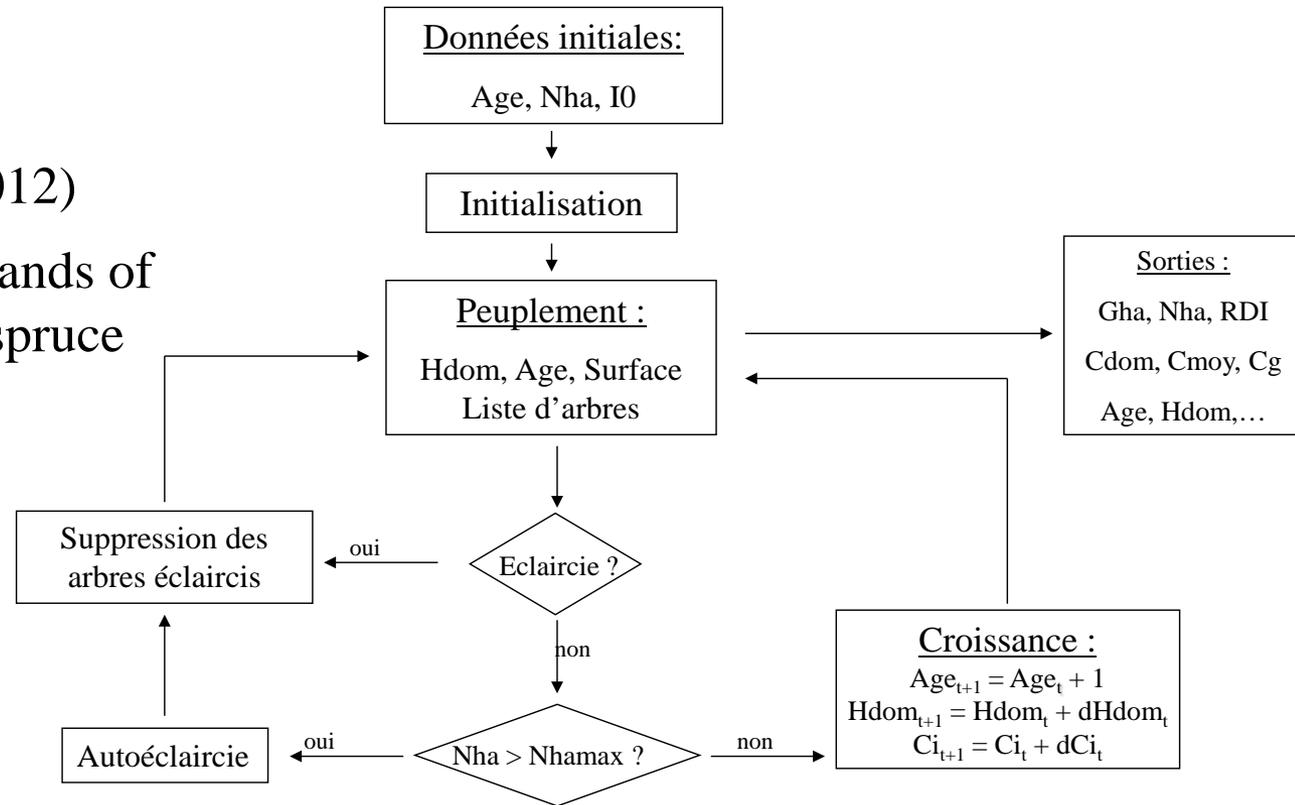- Implemented in Capsis (interactive and batch mode)

BWinPro

- Nagel *et al.* (2002), Albrecht *et al.* (2012)

- For uneven-aged mixed stands, parameterised for several tree species

- Empirical, spatially "semi-explicit"

- Basic idea: diameter
  → height
  → crown dimensions
  → growth of basal area
  and height

- Mortality with (i) maximum age and (ii) critical degree of crown closure

- Initial state: plot size, ages & diameters of trees

- Open-source implementation in Java (interactive and batch mode)

(Nagel & Schmidt 2006)

Measured/field data

Data generation

Generation of spatial structure

Complete data set

Silvicultural management

Growth prediction

Assortment prediction

Wood quality prediction

Mortality prediction

Stored basic information,
- model parameters:
  - data generation
  - growth
  - mortality
  - qualitative classification
  - assortment (size class)
  - yield tables
  - country-specific tree

Results: single tree list, stand parameters, height curves, dbh and height distribution, growth rates;
Results of the structural analysis;
Results of qualitative and size class grading

Visualisation: *SVS, Treedraw*

Export of single tree list to the grading simulator *Holzernte*

Gymnos

- Perin, Ligot *et al.* (2012)

- For even-aged pure stands of Douglas fir, Norway spruce and two larch species

- Empirical, spatially inexplicit

- Fundamental relationships: H_{dom} vs. age, growth of tree basal area

- Mortality: maximum basal area w.r.t. mean circumference of stand

- Initial state: (i) plot size, density & age of stand ($\rightarrow$ virtual stand) or (ii) inventory file

- Implemented in Capsis

Données initiales:
Age, Nha, I0

Initialisation

Peuplement :
Hdom, Age, Surface
Liste d'arbres

Sorties :
Gha, Nha, RDI
Cdom, Cmoy, Cg
Age, Hdom,…

Suppression des arbres éclaircis

oui — Eclaircie ?

non

Croissance :
$Age_{t+1} = Age_t + 1$
$Hdom_{t+1} = Hdom_t + dHdom_t$
$Ci_{t+1} = Ci_t + dCi_t$

Autoéclaircie — oui — Nha > Nhamax ? — non

(Perin 2013)

Model assessment data

- Sources
  - FVA Freiburg, Germany:143 plots
  - Ecouves data, INRA, France: 12 plots
  - GIS Coopérative de données, France: 43 plots
- Measurements at irregular time intervals
  - Diameter on all the trees at all the measurement times
  - Height on (a subset of) the trees at some measurement times
  - Cause and time of removal on all the trees
    $\rightarrow$ exact information on mortality and thinnings
  - Co-ordinates of the trees not available in all the plots

- Plot selection criteria for the model assessment
  - Pure Douglas fir stands
  - Large variation in initial density
  - Large variation in intensity and timing of thinnings
  - Variation in site fertility (site index)

Challenges in model comparison

- Differences between models in their *logic / structure* and *thinning algorithm* may make simulation results incommensurable or difficult to compare with observed data

- Key differences in model structure
    - Input data requirements
    - Growth unit
    - Time step
    - Randomness
    - Nature of output data

- Key differences in thinning algorithm
  - Timing
  - Intensity
  - Selection of trees (size & spacing)
- Key questions
  - How to make simulations with different kinds of models commensurate?
  - How to enable comparison with observed data?
  - *What do we want to compare – simulators "in the raw", or growth mechanisms of simulators and their reaction to interventions, or ...?*

Challenges: input data requirements

- Difference between models in the *starting point of simulation*: age of 0 years *vs.* time point where diameters (measured / predicted) are available

    → "Advantage" for models starting latest
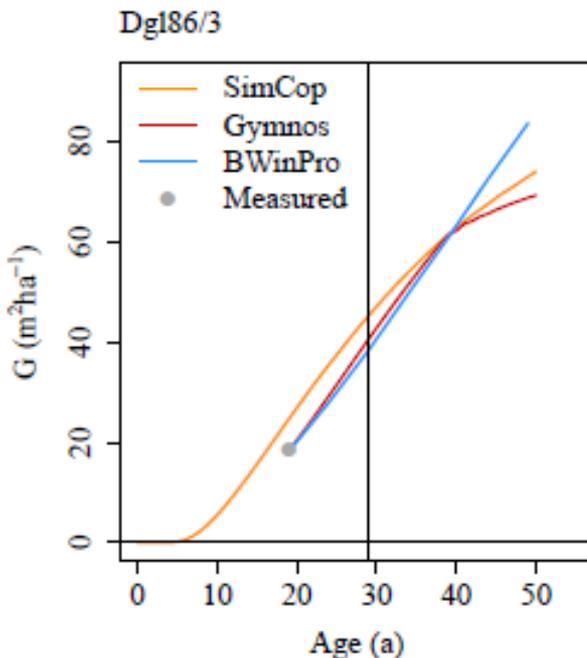
Simulations of the growth of a plot with initial density of 1000 stems/ha, without thinnings:
a) SimCop starts from age 0
b) Gymnos and BWinPro start from age 19, the first measurement time point (the first inventory file available)

→"Advantage" for Gymnos and BWinPro, as when they start SimCop already exhibits considerable error

Potential solution for a common starting point: use simulation results of a model starting from age of 0 years as input for models starting later

Tuning a model starting from age of 0 years to match observed values at first measurement time point: feasibility and consequences?

- *Unavailability of the input data required*: missing values predicted with imputation models (included in the simulation model or not)
  - Typical missing data: tree heights, tree co-ordinates

$\rightarrow$ Different values for the same input characteristics of the same plot for different simulation models, depending on the imputation models used

Potential solution for harmonising the input data: use exactly the same input data generated with the same imputation models

Imputation considered as a property of a simulation model?

Challenges: time step

- Difference between models in the *time step* of simulation: 1 year *vs.* 5 years

   → Problems with 5-year step in timing of thinnings

   → Problems with 5-year step in comparing simulation output with observed values (year of observation may not match in 5-year intervals)

   Usual solution for output comparison: interpolate (linearly) within 5-year intervals to get simulation result for the desired year

Challenges: randomness and output

- *Randomness* in models: output as distributions of stand characteristics within a plot

  - Typical processes involving randomness: generation of the initial state of a stand (tree locations, diameters and heights), thinning (selection of trees)

$\rightarrow$ Problems in comparing distributions (of scalars or time series) with single values (from data or other models)

Usual solution for output comparison: use mean / median of the distribution of the output characteristic

Comparison of time series: see if observed time series lies within $100(1-\alpha)$ % confidence band of simulated time series

- Comparing stand characteristics based on tree characteristics *derived from "emergent" tree characteristics*
  - Usual "emergent" tree characteristics: diameter and / or height
  - Typical tree characteristics derived from "emergent" characteristics: stem volume, biomass

$\rightarrow$ Extra variation between simulation models due to differences in the derivation methods

Potential solutions for output comparison:

(i) Use only stand characteristics based on "emergent" tree characteristics (N, G, diameter distribution, H_{dom})

(ii) Derive "non-emergent" tree characteristics using the same methods (volume models, biomass models) for all the simulation models

Challenges: thinning algorithm

- Timing
  - Age: models with different time steps may result in different timings

  - H_{dom}, RDI, N, G, ...: different models result in different timings, which are likely to deviate from that of the observed data

    $\rightarrow$ How to how to compare output time series involving different timings of thinnings with data or with each other?

- Intensity
  - Target stand after thinning (N, G, V): corrects an overestimated level of before-thinning growth → "too good" results?

  - Removal in thinning (N, G, V): fair, no correction effect

  - Intricate unitless variable: how to determine from observed data?
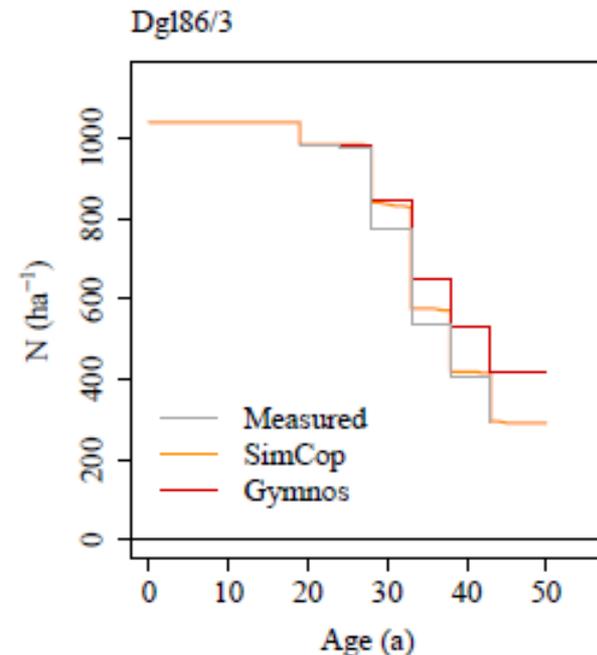
Simulations of the growth of a plot with observed thinnings:
a) SimCop: thinning into target density
b) Gymnos: list thinning
→"Too good" results for SimCop, as in the thinnings the stand density is always adjusted to the right level
NB. Only cut trees included in the thinnings, naturally dead trees excluded

Dg186/3

- Selection of trees: size
  - Individual selection (list of removed trees obtained from data)
  - Explicit allocation of removed trees into size classes
  - Qualitative variable describing thinning type
- Selection of trees: spacing
  - List of removed trees and tree co-ordinates obtained from data in a spatially explicit model: perfect description
  - Other: minimum distance between thinned trees, maximising nearest-neighbour distance for remaining trees

- Flexibility: different thinning algorithms have different ability to imitate observed (unconventional) thinnings

  Potential solutions to harmonise thinnings:

  (i) Use list of removed trees obtained from observed data (list thinning)

  (ii) Use some other common thinning algorithm for all models (Capsis)

How to compare model output with data

- Considering *errors of stand characteristics* with respect to data (observed – simulated)
  - Stand characteristics based on (i) "emergent" tree characteristics or on (ii) tree characteristics derived with the same methods from the "emergent" characteristics

- Errors vary in many directions
  - Models with no randomness: variation (i) over plots and (ii) over time within a plot
    - Plot-wise time series of errors of stand characteristics
  - Models with randomness: additional variation over the distribution(s) of the random element(s)
    - Collection (distribution) of plot-wise time series of stand characteristics

- Difficulties in constructing error time series within a plot: differences between time points of observations and simulations (5-year time step, timing of thinnings)

- Difficulties in comparing time series between plots: differences between time points of observations / simulations

- Different ways to compute means, variances and RMSEs of errors
  - (i) Over plots and over time (ignoring the time series structure) $\rightarrow$ general bias, error variance and RMSE for each model
  - (ii) Over time within a plot $\rightarrow$ bias, error variance and RMSE for each model in each plot
  - (iii) Over plots in age classes $\rightarrow$ "time series" of bias, error variance and RMSE for each model
  - For models involving randomness, use mean time series of errors in these considerations

Concluding remarks

- Model comparison against data is not a straightforward task

- Decisions taken to make simulations with different models commensurate and to imitate in the best way the observed thinnings can have drastic influences on the results

- *"Necessary conditions"* for meaningful model comparison?

Acknowledgements

Access to growth models and help with implementation:

T. Bronner & J.-M. Ottorini (SimCoP)
A. Albrecht (BWinPro)
J. Perin & G. Ligot (Gymnos)

Thank you for your attention!

Challenges: technical issues

- Technical difficulties in using models with different types of interface

  - Data input and output: text files of different formats, database files of different formats

  - Interactive *vs.* batch mode: facility to perform a large number of simulations

  - Documentation

  - Access to the code: modifiablity, extensibility

  - Potential solutions: common modelling platform, standardisation of input and output