Flanders
State of the Art

Development of a semi-automated process to analyse data from the Flemish forest inventory

Toon Westra, Hans Van Calster, Pieter Verschelde, Leen Govaere and Paul Quataert

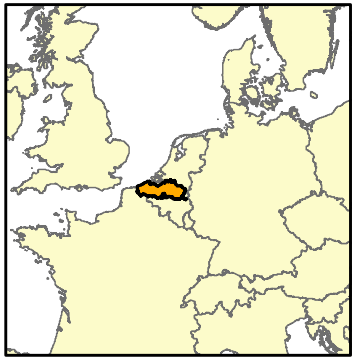
RESEARCH INSTITUTE
NATURE AND FOREST

3rd Science for the Environment Conference
Aarhus Denmark 1-2 October 2015

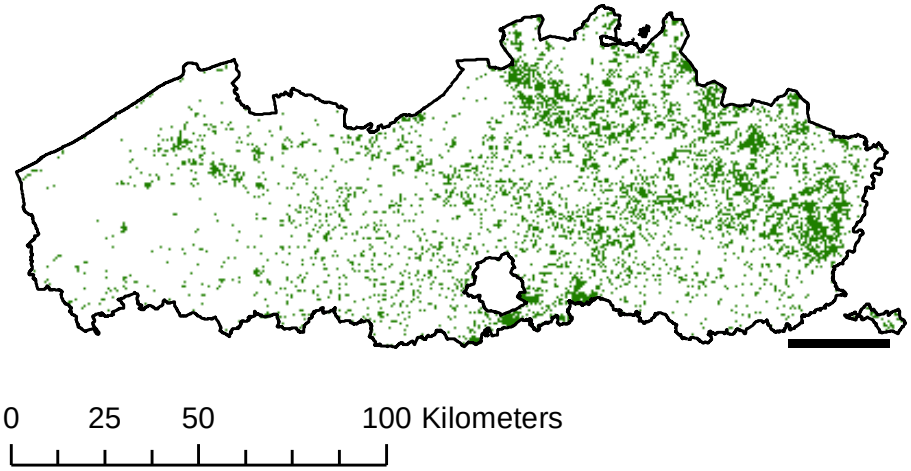
Overview

- Forest inventory in flanders: objectives and design
- Challanges for data analysis
- Semi-automated analysis process
  - Quality control of measured variables
  - Calculation of derived variables
  - Actual data analysis
- Conclusions

Flemish forest inventory



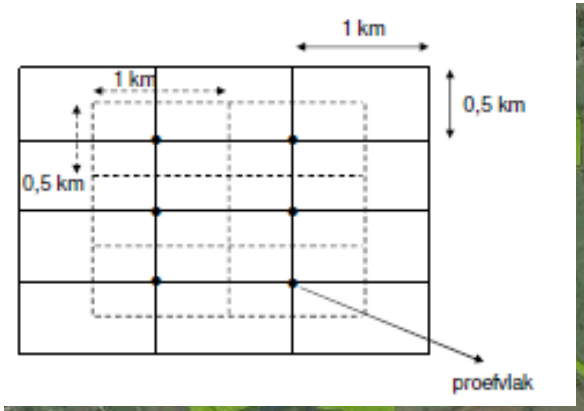
- Forest cover in Flanders
  - 185600 ha
  - 7,3 %



Flemish forest inventory

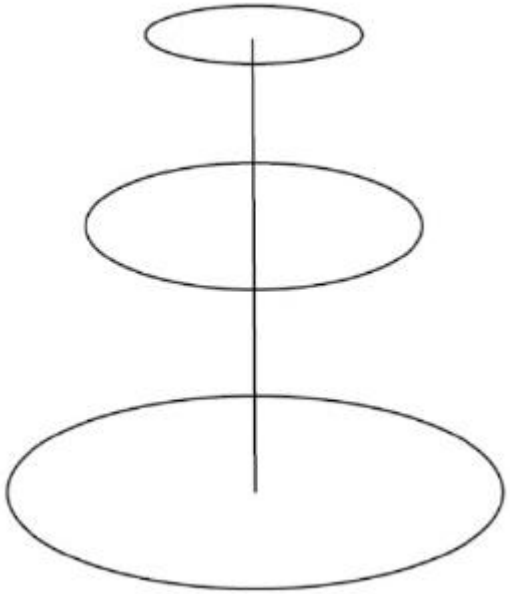
- Flemish forest inventory (VBI)
  - Information on status and trends of forest characteristics in Flanders
  - Managed by Agency of Nature and Forests (ANB)
  - 61 monitoring questions/ targets
  - 6 thematic categories
    - Forest distrubution
    - Tree species composition
    - Forest stand characteristics
    - Biodiversity indicators
    - Vegetation composition
    - Forest management and forest use (wood quantity and quality)

Flemish forest inventory



Flemish forest inventory

- Sample plot: stand characteristics and tree measurements



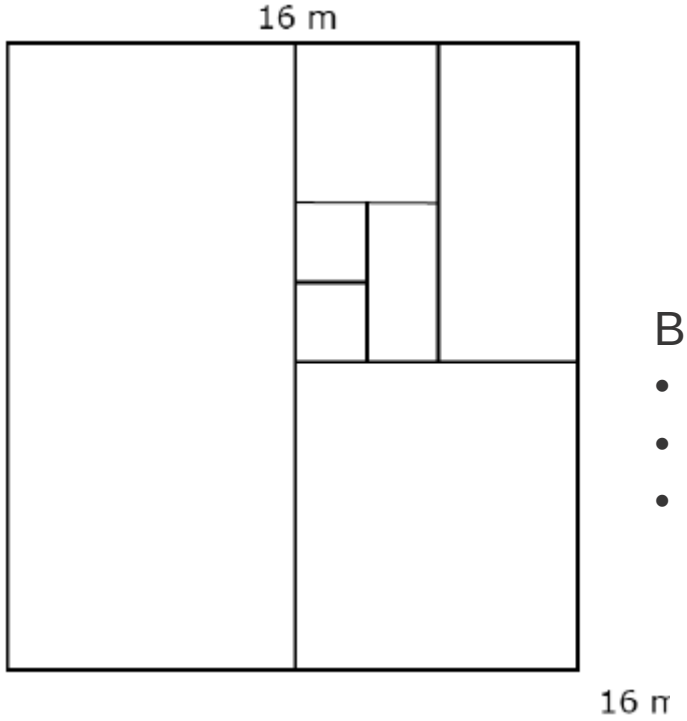
A2
- 4,5 m radius
- Trees with perimeter < 22 cm and height >= 2 m
- Number per species

A3
- 9 m radius
- Trees with 22 cm <= perimeter < 122 cm
- Species, status, perimeter, height, location

A4
- 18 m radius
- Trees with perimeter >= 122 cm
- Species, status, perimeter, height, location
- Stand characteristics

Flemish forest inventory

- Sample plot: vegetation

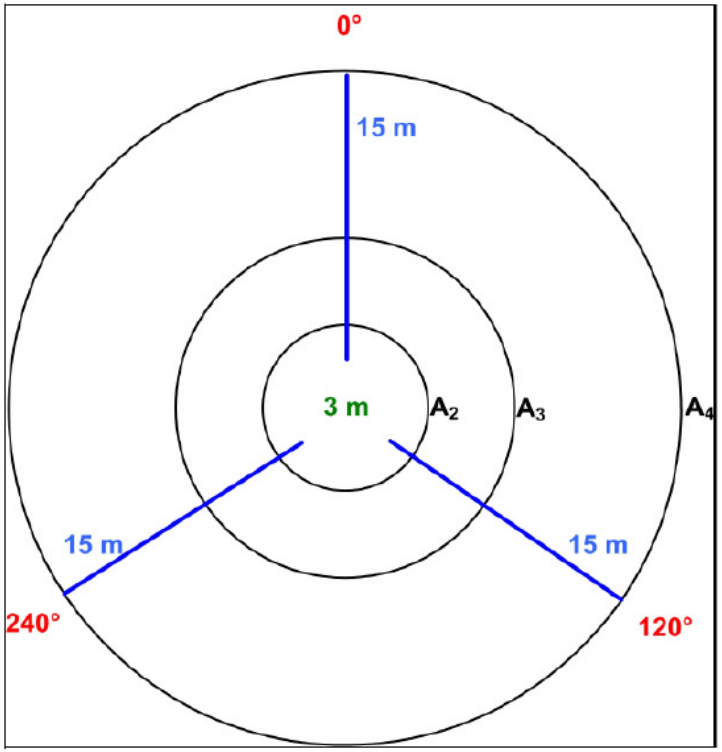


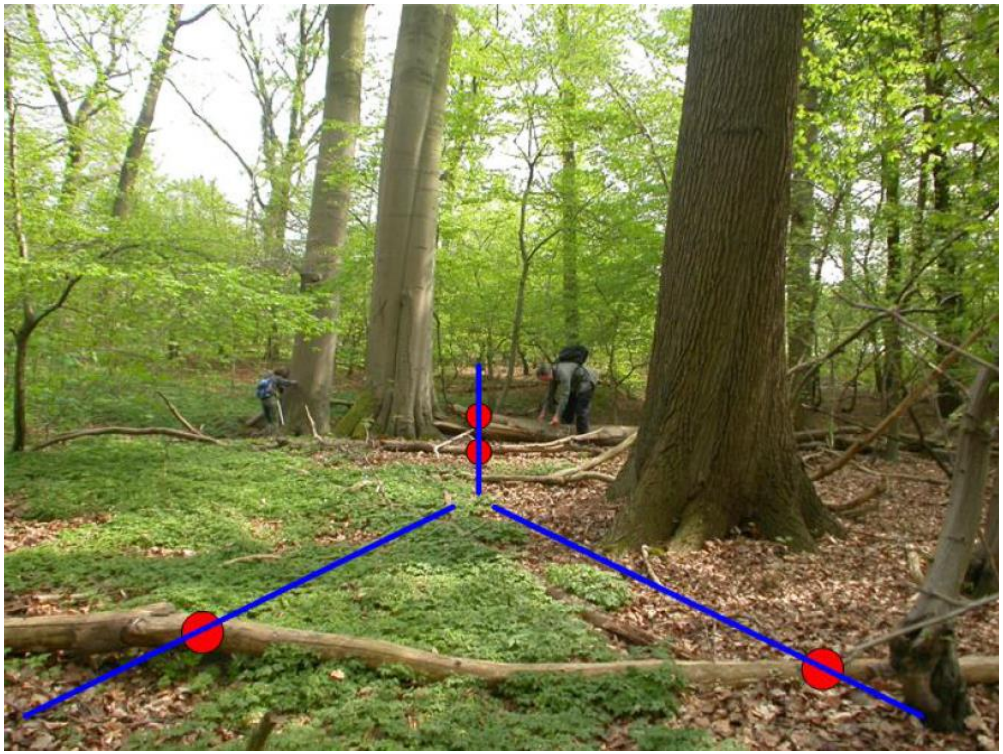
Braun-Blanquet cover/abundance

- Herb layer
- Shrub layer
- Tree layer

Flemish forest inventory

- Sample plot: line intersect sampling for laying dead wood





Flemish forest inventory

- 1st inventory: 1997 – 1999
- 2nd inventory: 2009 – ongoing
  - 12 year cycle
  - Approximately 50 % of marked sample plots could be retrieved → mixture of permanent and tempory sample plots
  - Various differences in measurement protocol
    - 1st inventory: plots on forest edge or containing different stand types removed to get homogeneous forest plot
    - 2nd inventory: plots are put on original position

Analysing data from the Flemish forest inventory

- Goal: periodical reporting of results based on standardized data analysis

- Challanges

  - Many types of data collected in the field
  - Many variables derived from raw data (for example: wood volume, biodiversity indices, Ellenberg values,…)
  - Many monitoring questions to be answered

→ Good organization and documentation of dataset

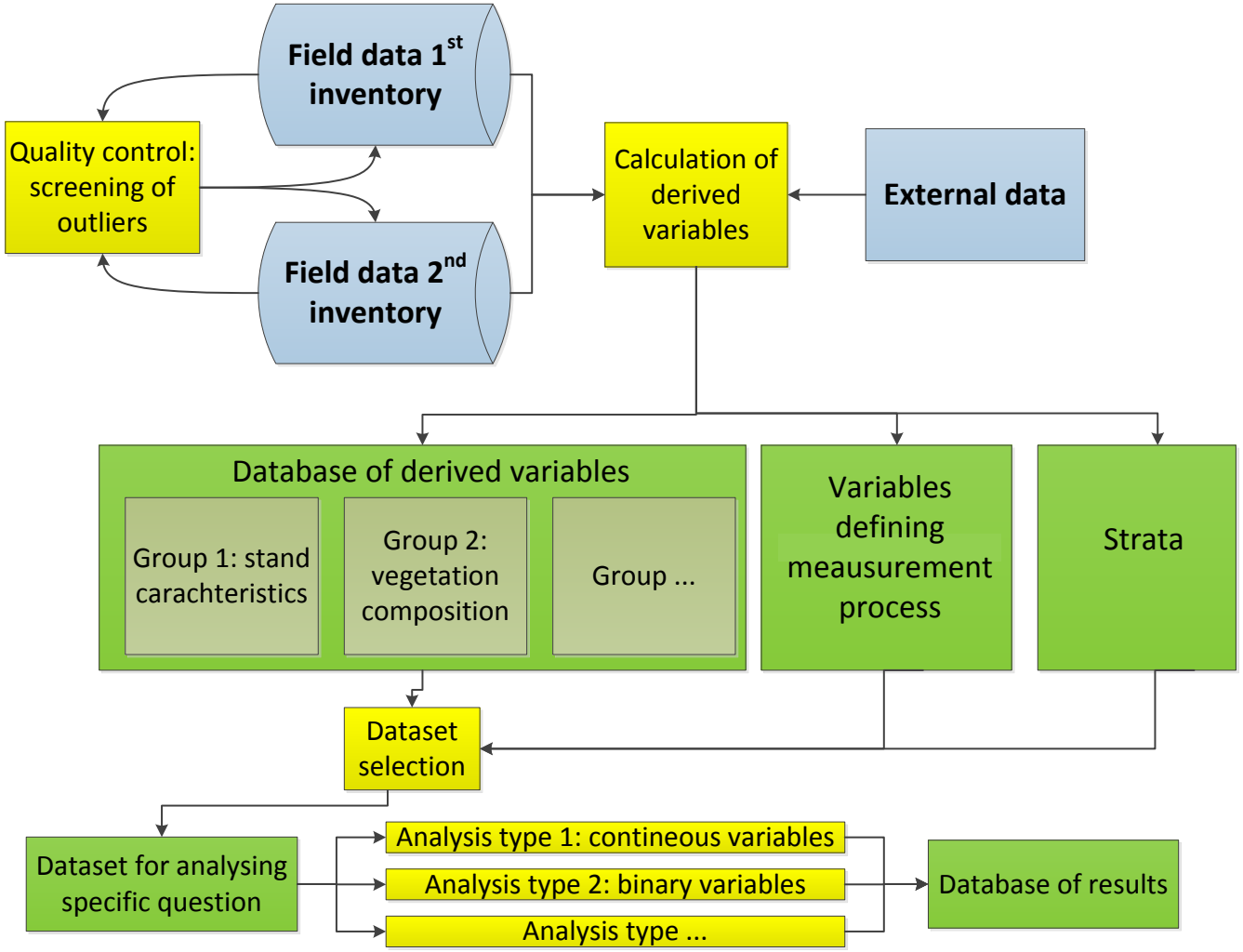
  - Mixture of permanent and temporary plots
  - Some plots are only partially covered by forest

→ Need for appropriate statistical techniques

Analysing data from the Flemish forest inventory

- Quality assurance of data analysis is essential
- INBO developped a semi-automated process to support data analysis and reporting of results
- Analysis process is implemented by a series of R-scripts
- The actual analysis will be performed by ANB

Overview of analysis process



Analysis process: quality control

- Comparison of tree measurements in the 1st and 2nd inventory
- Only for trees that have been remeasured in 2nd inventory (35% of the trees)

- Detection of anomalies and group them into different types
  - Zombies: status changes from dead to alive
  - Shifters: tree changes to another species
  - Movers: tree changes position
  - Outliers in perimeter growth
  - Outliers in height growth

- Validation
  - Change value of measured variable
  - Consider the trees in the 1st and 2nd inventory as different trees → this impacts the further analysis
  - Confirm measurements

Analysis process: database of derived variables

- Most questions can not directly be answered based on the original measured variables → derived variables have to be calculated

- For example
  - What ‘s the volume of dead standing trees per ha?
  - Status, height and perimeter per tree → volume per dead tree → volume dead trees per plot type → volume dead trees per ha

- In total 90 derived variables for 61 monitoring questions/targets

- Variables describing the measurement process are needed to perform a proper statistical analysis
  - Measurement date
  - Plot weight
  - Permanent plot of temporary plot
  - …

Analysis process: database of derived variables

- Derived variables, measurements characteristics and strata information are organised in a database

- Documentation
  - Which derived variables are needed to answer a certain monitoring questions, which external data is used
  - Clear description of how derived variables are calculated

Analysis process: actual data analysis

- Example scripts for detailed analysis of monitoring questions
  - Different types of derived variables: continuous, binary, count data, categorical data
  - Thematically different types of questions
- Data analysis consists of different parts
  - Data exploration
  - Design-based parameter estimation
  - Model-based parameter estimation
- Results can be stored in a database
  - Parameter estimate and confidence intervals
  - Date of analysis
  - Characteristics of analysing technique
  - Characteristics of the dataset on which analysis was performed
- Documentation
  - How to interpret data exploration plots
  - How to interpret models
  - How to validate results

Analysis process: model – based parameter estimation

- Estimating differences between both inventories is not straightforward
  - In only part of the locations plots were remeasured, in the other part new plots were established
  - Measurements in permanent plots are more similar

- Our approach → Generalized linear mixed effects models (GLMM)
  - Plot ID is used as random effect
  - GLMM can handle different data distributions
  - GLMM can handle plot weights
  - Strata can be incorporated
  - Nlme package (Pinheiro et al., 2015)

Analysis process: model – based parameter estimation

- Example analysis (preliminary results)
  - VolumeSnags ~ InventoryID + (1 | PlotID)
  - Model output
    - Fixed effects:

```
             Estimate Std. Error t value
(Intercept)  4.0238    0.3122    12.889
InventoryID2 5.1140    0.5310     9.631
```

  - Estimated standing dead wood volume
    - 1st inventory: 4,0 m^3/ha
    - 2nd inventory: 4,0 + 5,1 = 9,1 m^3/ha
    - Difference: 5,1 m^3/ha
    - 95% confidence interval of difference: 4,1 – 6,2 m^3/ha

Conclusions

- Quality assurance of data analysis process is an essential part of long term monitoring programs

→ Standardized, statistically sound, repeatable analysis of data and reporting of results

- Approach for Flemish forest inventory
  - Conceptual scheme of analysis process
  - Well documented R-scripts
  - Parameter estimation based on GLMM
  - Documentation + manual

Thank you