Using R for small area estimation in the Norwegian National Forest Inventory

Johannes Breidenbach

Ronald E. McRoberts, Rasmus Astrup

Norwegian Institute of Bioeconomy Research

Climate Center and National Forest Inventory
  P.O. Box 115, NO-1431 Ås
  Tel: +47 9988 1892; JOB@nibio.no
Contents

Introduction
   NFI
   Small area estimation in the NFI

Methods

Results

Summary
The Norwegian National Forest Inventory

- Since 1919
- >22,000 sample plots
- 1/5th of plots measured every year
The Norwegian National Forest Inventory

- Circular sample plots, 250 m²
- Species, diameter, location, height
- Ecological parameters, harvests
- Timber volume, biomass
- Ca 150 different parameters
Why NFI?

- Political decision processes
- National estimates of forest parameters
- Reporting requirements (e.g. Kyoto protocol, FAO)
- Strategic planning
Small area estimation

- Too few NFI plots for a reliable estimate
- Estimates on stand level → operational forest management decisions
- Combination of field plots with RS
- Design-based → model-based
Aim

- Overview of synthetic estimators
- Empirical coverage of CIs
- Implementation: JoSAE package
Contents

Introduction
NFI
Small area estimation in the NFI

Methods

Results

Summary
Regression synthetic estimator – mean

- Systematic sample
- Linking model

\[ y = X\beta + \varepsilon, \quad \varepsilon \sim N(0, \sigma^2_\varepsilon) \]

- Mean of the model predictions for the small area \( m \)

\[ \hat{Y}_m = \frac{1}{N_m} \sum X^T_m \hat{\beta} = \bar{X}_m^T \hat{\beta} \]

- Biased – but useful
Regression synthetic estimator – Var I

“...easily obtained...”

Est. of superpopulation parameters

\[
\widehat{\text{Var}}_S(\hat{Y}_m) = \bar{X}_m^T \widehat{\text{Cov}}(\hat{\beta}) \bar{X}_m
\]

- Independent of stand size
- Residual error ignored since \(\varepsilon \sim N(0, \sigma_\varepsilon^2)\)
Regression synthetic estimator – Var II

For small stands

Pred. finite population elements

Including heteroskedasticity

\[
\widehat{\text{Var}}_{SH}(\hat{Y}_m) = \widehat{\text{Var}}_S(\hat{Y}_m) + \frac{1}{N^2_m} \sum_{i} N_m \hat{\sigma}_i^2
\]
Regression synthetic estimator – Var III

Including spatial autocorrelation

\[
\hat{\text{Var}}_{SSS}(\hat{Y}_m) = \hat{\text{Var}}_S(\hat{Y}_m) + \frac{1}{N^2_m} \sum_{i}^{N_m} \sum_{j}^{N_m} \hat{\sigma}_i \hat{\sigma}_j \hat{\rho}_{ij}
\]

- Requires an estimate of spatial autocorrelation
Empirical coverage proportion (ECP)

- Sample plots clustered in stands
- Info on spatial autocorrelation
Contents

Introduction
NFI
Small area estimation in the NFI

Methods

Results

Summary
Observed/estimated mean
CI(Y_synth)
CI(Y_synth−small−het)
CI(Y_synth−small−spat, rho=55)
Contents

Introduction
  NFI
  Small area estimation in the NFI

Methods

Results

Summary
Summary

\[
\hat{\text{Var}}_S(\hat{Y}_m) = \hat{\mathbf{X}}_m^T \hat{\text{Cov}}(\hat{\beta}) \hat{\mathbf{X}}_m
\]

\[
\hat{\text{Var}}_{SS}(\hat{Y}_m) = \hat{\text{Var}}_S(\hat{Y}_m) + \frac{\hat{\sigma}_\varepsilon^2}{N_m}
\]

\[
\hat{\text{Var}}_{SSS}(\hat{Y}_m) = \hat{\text{Var}}_S(\hat{Y}_m) + \frac{1}{N_m^2} \sum_{i}^{N_m} \sum_{j}^{N_m} \hat{\sigma}_i \hat{\sigma}_j \hat{\varrho}_{ij}
\]

- Random effects on stand level
  ⇒ Careful interpretation!
- R package JoSAE will be updated (paper accepted), other SAE packages available!
Thank you for listening!