Overview of Small Domain Estimation Research Outside the United States: An International Perspective

Richard W. Guldin

Senior Research Fellow, SAF

Principal, Guldin Forestry LLC

Overview Side-Boards

- **RECENT** SDE Research – Therefore, **since 2010**
- **OUTSIDE USA** – Therefore, **International**
- **PERTAINING TO FORESTS**
  - Ignoring all SDE research in socio-economic domains
  - Ignoring all SDE research in agriculture or other natural resources domains
  - Ignoring all basic statistical research on SDE

- *Therefore, an overview of applied SDE research published since 2010 about connecting national forest inventory field plot data from other countries to improve mean square errors of estimates for smaller geographic areas.*

What I'm presenting today is an overview of applied small domain estimation research—aka small area estimation research– that's been published since 2010. My focus is on SDE research conducted in other countries that connects their national forest inventory field plot data with auxiliary data to improve the mean square errors of estimates for smaller geographic areas.

I am ignoring all SDE research by USA authors—whether basic statistical research or applied research on domains outside of forests.

These sideboards restrict me to reporting on a small sliver of the published SDE research; but it's the sliver that those here

are most interested in.

Sources of Information

- Two key journals
  - *Remote Sensing of Environment* – 24 articles reviewed
  - *Remote Sensing* – 382 articles reviewed

- Google Scholar and ResearchGate
  - First 150 citations returned from Google Scholar
  - First 50 citations returned from ResearchGate

- Personal contacts with researchers from other countries

My data come from three sources:

- Two key journals—roughly 400 articles whose key words include "forest", "forest inventory" and "small area estim*.
- Two key online applications—the first a web-crawler and the second a members-only networking app
- Personal contacts with researchers from other countries—drawn from my personal knowledge and a couple of key contacts from Ron McRoberts.

Key Findings #1 – Best Sources of Info

- *Remote Sensing of Environment* was most helpful & informative journal – 24 publications from 2000 – 2019
- *Remote Sensing* journal's 382 articles were less helpful because of their strong focus on:
  - Sensor performance; very heavily into LiDAR and UAVs
  - Up-scaling from small areas to regions & national scale
- Google Scholar and ResearchGate added only a couple new articles on forest-related SAE beyond the journals
  - Less than 20 percent of citations were recent; since 2010
  - Strong focus on estimating above ground biomass from LiDAR point clouds. Recent work emphasizes UAVs and terrestrial sensors

The journal *Remote Sensing of Environment* had the articles that were most relevant to my assigned task. The second journals' citations were less helpful for two reasons. First, this journal is strongly focused on sensor performance; for forests, very heavily into airborne, and more recently terrestrial—LiDar and unmanned aerial vehicles—UAVs. Second, their focus on estimation was for upscaling from small areas to a regional or national scale.

Google Scholar and ResearchGate added only a couple of new articles on forest-related SDE beyond the two journals. The new articles came from the Canadian Journal of Forest Research. Surprisingly, the citations returned from these apps were not the most recent—probably because of journals' embargo policies. Again, the recent citations that were returned were strongly focused on estimating above-ground biomass from LiDAR point clouds or UAV pixels.

Key Findings #2 - Leading Topics from
Remote Sensing of Environment

- Imagery was the leading source of auxiliary data
  - Detailed field plot data were used to impute values to image pixels; values were then aggregated from all pixels to make SAEs.
  - K-Nearest Neighbor algorithms were used for imputation in 6 of the 24 articles. Many other types were tested in one or two articles.
- Attributes most frequently imputed were volumes of timber or growing stock and type of forest cover type
  - Two articles used imagery from two time periods to assess change
- The precision of various models and estimators was the focus of 8 papers
- Inaccuracy of field plot coordinates was common issue

Mt review of the 24 articles published in *Remote Sensing of Environment* since 2000 revealed several common threads.

Imagery was the leading source of auxiliary data for making SDEs. Most of the early publications were about the K-NN approach. Six K-NN papers were published between 2006 and 2010. Erkki Tomppo (a Finn), Ron McRoberts, and/or Steen Magnussen (a Canadian) were the lead author or a coauthor in all six. The most recent K-NN paper was a review by Chirici et al. in 2016. Chirici is an Italian.

The attributes most frequently imputed were timber volume or growing stock volume or type of forest cover.

Eight papers examined the precision of various models and estimators.

Inaccuracy of field plot coordinates was a common issue in many papers. Tests of several "work-arounds" were reported.

Key Findings #3 - Small Network

- Research expanded from 1990 to 2010, largely due to a small network of researchers
  - Erkki Tomppo (Finland) was the early global leader
  - Ron McRoberts (USA) provided international leadership

Author/Coauthor	2010-2019	2000-2009
McRoberts (USA)	7	5
Tomppo (Finland)	3	6
Breidenbach (Norway)	3	0
Astrup (Norway)	3	0
Rahlf (Norway)	2	0
Chirici (Italy)	2	0
Waser (Switzerland)	2	0
Magnussen (Canada)	2	2
Finley (USA)	2	1
Stehman (USA)	1	2

- 2000-2009 totals
  - 11 articles
  - 9 authors mentioned once
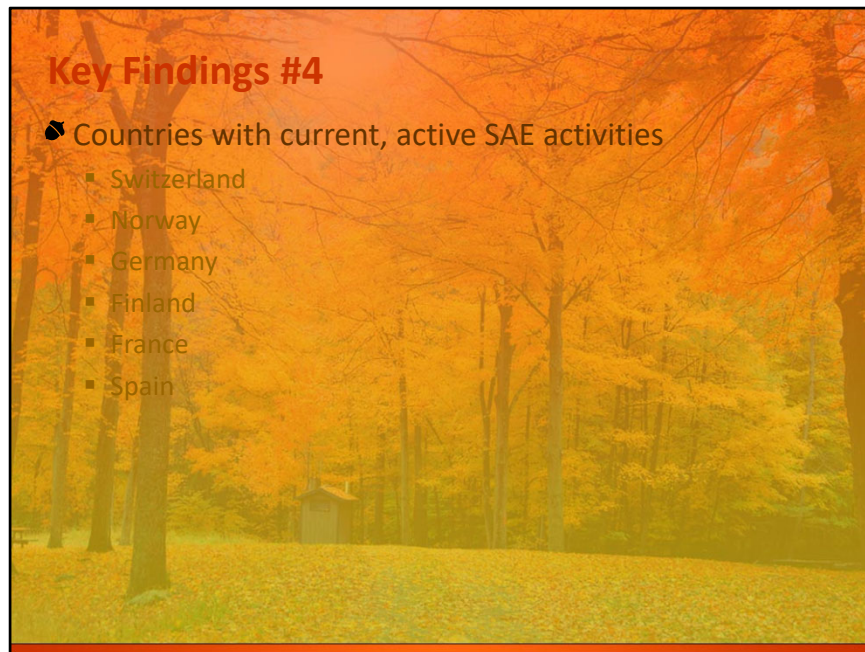 - 25 authors/co-authors
- 2010-2019 totals
  - 13 articles
  - 28 authors mentioned once
  - 55 authors/coauthors

Looking at all 24 articles in *Remote Sensing of Environment* between 2000 and 2019, Ron McRoberts was the author or co-author on half of them. Sole author on 4, lead author on 3, and coauthor on 5.

From a networking perspective, Tomppo and McRoberts were central figures. They coauthored six articles together.

McRoberts coauthored 2 articles with Magnussen, two with Chirici, and single articles with Breidenbach, Astrup, and Finley. Beyond his work with McRoberts, Erkki Tomppo published twice with Magnussen, Chirici, and Waser, and once with 12 others.

Note that although the number of articles was roughly the same for the two time periods—11 versus 13—the number of authors and coauthors doubled while the number of who published once tripled.

Key Findings #4

- Countries with current, active SAE activities
  - Switzerland
  - Norway
  - Germany
  - Finland
  - France
  - Spain

Six European countries have active, SDE activities—recently published and still underway.

Based on the leading authors identified from the published articles, I reached out to a number of them for additional details. My sense of the responses is that Switzerland, Norway, and Germany have robust SDE research programs, usually with collaboration between NFI and university researchers.   There have also been one or two recent studies since 2010 in Finland, France, and Spain. Some of these researchers and their publications have referenced additional forest-based SDE

research in Sweden, Austria, and Denmark, but I couldn't find recent publications from researchers in those three countries.
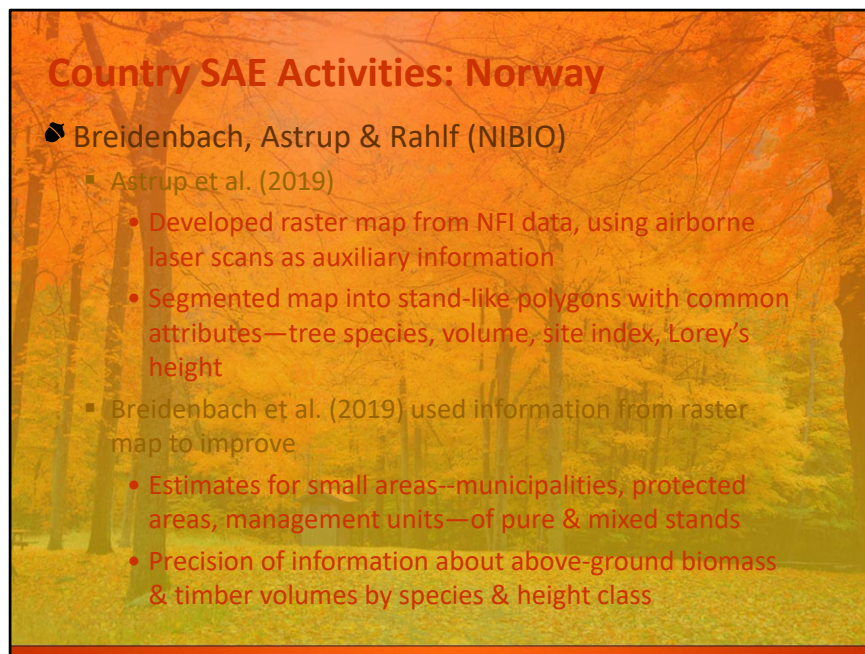
Let me touch very briefly on what's happening in these countries before my time expires. I know some of you will want more details. Give me your business card with email address and I'll email you my manuscript with the details and citations.

Country SAE Activities: Switzerland

- Prof. Daniel Mandallaz & students (ETH-Zurich)
  - Design-based, model-assisted approach
  - Developed two-phase simple random sampling with post-stratification estimation based on auxiliary info
- Pullkinen, Lanz & Zell (WSL Birmensdorf)
  - Combining NFI data with digital aerial imagery and LiDAR-based terrain models to estimate forest area, total growing stock volume, and above ground biomass for cantons, forest districts, and municipalities

Swiss colleagues sent me a four-page summary of their recent publications and ongoing research activities. The challenge in the steep alpine terrain is to use SDEs to provide forest management information that can replace costly, forest management inventories.

I must note that Daniel Mandallaz and his graduate students have become widely known and respected for their recent research. Much of it is being used in other countries, such as Germany, and influencing researchers in other countries.

ETH-Zurich is a science, technology, engineering, and mathematics university in Zurich. WSL is the Swiss Federal Institute for Forest, Snow and Landscape Research, located just outside Zurich in Birmensdorf.

Country SAE Activities: Norway

🌰 Breidenbach, Astrup & Rahlf (NIBIO)

▪ Astrup et al. (2019)
  • Developed raster map from NFI data, using airborne laser scans as auxiliary information
  • Segmented map into stand-like polygons with common attributes—tree species, volume, site index, Lorey's height

▪ Breidenbach et al. (2019) used information from raster map to improve
  • Estimates for small areas--municipalities, protected areas, management units—of pure & mixed stands
  • Precision of information about above-ground biomass & timber volumes by species & height class

SDE research in Norway is centered in the Norwegian Institute of Bioeconomy Research. This is the centennial year of Norway's NFI; a major program of NIBIO. Several of us here were able to attend and present papers at the centennial celebration in Norway this past May.

Three NIBIO researchers—Johannes Breidenbach, Rasmus Astrup, and Johannes Rahlf—are a strong core of expertise on SDE. They have combined NFI and airborne laser data to develop a detailed raster map and then used information from that map to make estimates for quite small areas. They have teased out differences

among several conifer species and mixed conifers and hardwoods to really improve the precision of estimates for biomass and timber volumes by species and height classes.

Read their articles and follow their work.

Country SAE Activities: Germany & Finland

- **Germany** – SAE for Forest Management
  - NFI information to make estimates in Rhineland-Palatinate state (8,400 km² of forest) at forest district (~45,000 ha), and sub-district (~ 4,500 ha) levels
  - R-P scientists partnered with Mandallaz & ETH students to test 3 types of design-based, small area regression estimators in a double-sampling approach
  - Data from Thuringia & Mecklenburg-Pomerania are next
- **Finland** – Efficiency of Post- Stratification for SAE
  - Haakana et al (2019) reported results from southern municipalities & provinces (5,700 – 921,000 ha)
  - Used designed-based (K-NN) approach to estimate growing stock volume by species for Landsat TM pixels, after Tomppo
  - Updating base maps for thinnings, regeneration, & final harvests and better municipality boundary files would improve SDEs.

In Germany, scientists from the NFI and researchers from ETH-Zurich, Gottigen and Freiburg have pooled their talents to work with state agency experts in southwestern Germany to improve estimates for small management areas—the 45 districts and 455 sub-districts. The recent successes there are leading to invitations from other state forestry agencies to replicate the SDE work there.

In southern Finland, there has been recent work to make acceptable estimates of growing stock volume for municipalities and provinces. What Haakana and

colleagues have found is that even older regional forest management maps can help provide auxiliary information to the K-NN imputation process. Again, Tomppo's earlier work undergirds all they do.

Country SAE Activities: France & Spain

- **France** – 3D Auxiliary Data
  - Used 3D variables (from airborne laser scans and digital aerial photos) to generate canopy height models for 2008 & 2014, a forest type map, vegetation indices, and Landsat 8 as auxiliary information for French NFI data
  - Used 3D height change estimates in K-NN method to estimate changes in basal area & several types of volumes
- **Spain** – Integrated Airborne Laser & NFI Data
  - Used Spanish NFI data, airborne laser & MODIS in a two-stage up-scaling procedure to estimate above-ground biomass
  - Unbiased ways to update NFI estimates to desired reporting year
  - Tested a "random forests" modeling approach & "bootstrap" estimators to construct change maps with specific uncertainties

Both France and Spain are taking some innovative steps to asses changes in forest conditions over time.

Again, the focus on height changes flows from the desire to make estimates of changes in above-ground biomass from volume equations that are a function of height.

Four summary points:

Spatial domains dominate. Socio-economic domains—popular in other disciplines—are absent in forest sector work.

Design-based, model-**assisted** approaches are most popular, although hybrid model-**based**, design-based approaches are actively being explored.

Recent applications show influence of Breidenbach and his fellow Norwegians and Mandallaz and his students—

a broader pool of European expertise.

Using SAEs to cut costs by replacing local forest management inventories and to have local forest estimates that are harmonize with NFI information are major stakeholder needs. A big emerging need is solid above-ground biomass estimates for carbon accounting and national carbon reporting.