Sampling for mapping and monitoring of soil carbon stocks

Dick Brus, Martin Knotters
Soil Science Centre
Wageningen University and Research Centre - Alterra
Introduction of the authors and their work


Clashing terms

Sample (n=5)

Sampling

Target universe, target population

Sampling unit

"Aliquot"
Aim

- To discuss sampling aspects of mapping, monitoring and validation
Motivation

- Selection process of sampling locations deserves more attention: ‘representative’ is often ill-defined

- ‘Start at the end, and reason backward’: integrated planning of data collection and data processing, with respect to the required information

Motivation

▶ Selection process of sampling locations deserves more attention: ‘representative’ is often ill-defined*

▶ ‘Start at the end, and reason backward’: integrated planning of data collection and data processing, with respect to the required information

Motivation

- Selection process of sampling locations deserves more attention: ‘representative’ is often ill-defined*
- ‘Start at the end, and reason backward’: integrated planning of data collection and data processing, with respect to the required information

Outline

1. Sampling for *mapping* of soil carbon stocks
2. Sampling for mapping of *changes* in soil carbon stocks (*monitoring*)
3. Sampling for *validation* of maps
Outline

1. Sampling for *mapping* of soil carbon stocks
2. Sampling for mapping of *changes* in soil carbon stocks *(monitoring)*
3. Sampling for *validation* of maps
Outline

1. Sampling for *mapping* of soil carbon stocks
2. Sampling for mapping of *changes* in soil carbon stocks (*monitoring*)
3. Sampling for *validation* of maps
Outline

1. Sampling for *mapping* of soil carbon stocks
2. Sampling for mapping of *changes* in soil carbon stocks (*monitoring*)
3. Sampling for *validation* of maps
1. Sampling for mapping: directed or random?

- Directed (targeted, purposive) sampling:
  - collection of data can be optimized for spatial interpolation (fair spatial coverage).

- Probability sampling:
  - enables model-free estimation of means or totals and their standard errors (design-based inference).
  - collected data are suitable for spatial interpolation if the sampling design guarantees fair spatial distribution of the selected sampling units.
1. Sampling for mapping: directed or random?

- Directed (targeted, purposive) sampling:
  - collection of data can be optimized for spatial interpolation (fair spatial coverage).

- Probability sampling:
  - enables model-free estimation of means or totals and their standard errors (design-based inference).
  - collected data are suitable for spatial interpolation if the sampling design guarantees fair spatial distribution of the selected sampling units.
1. Sampling for mapping: directed or random?

- Directed (targeted, purposive) sampling:
  - collection of data can be optimized for spatial interpolation (fair spatial coverage).
- Probability sampling:
  - enables model-free estimation of means or totals and their standard errors (design-based inference).
  - collected data are suitable for spatial interpolation if the sampling design guarantees fair spatial distribution of the selected sampling units.
1. Sampling for mapping: directed or random?

- Directed (targeted, purposive) sampling:
  - collection of data can be optimized for spatial interpolation (fair spatial coverage).

- Probability sampling:
  - enables model-free estimation of means or totals and their standard errors (design-based inference).
  - collected data are suitable for spatial interpolation if the sampling design guarantees fair spatial distribution of the selected sampling units.
1. Sampling for mapping: directed or random?

- Directed (targeted, purposive) sampling:
  - collection of data can be optimized for spatial interpolation (fair spatial coverage).

- Probability sampling:
  - enables model-free estimation of means or totals and their standard errors (design-based inference).
  - collected data are suitable for spatial interpolation if the sampling design guarantees fair spatial distribution of the selected sampling units.
1. Sampling for mapping: directed or random?

- **Directed (targeted, purposive) sampling:**
  - collection of data can be optimized for spatial interpolation (fair spatial coverage).
- **Probability sampling:**
  - enables model-free estimation of means or totals and their standard errors (design-based inference).
  - collected data are suitable for spatial interpolation if the sampling design guarantees fair spatial distribution of the selected sampling units.
Probability sampling

- *random* selection of elements
- selection probabilities are known, and $> 0$
- inference based on selection probabilities
Probability sampling

- random selection of elements
- selection probabilities are known, and $> 0$
- inference based on selection probabilities
Probability sampling

- random selection of elements
- selection probabilities are known, and $> 0$
- inference based on selection probabilities
Probability sampling

- random selection of elements
- selection probabilities are known, and $> 0$
- inference based on selection probabilities
Design types for probability sampling

- Simple random sampling
- Stratified simple random sampling
- Two-stage random sampling
- Cluster random sampling
- Systematic random sampling
- Stratified cluster random sampling
Design-based or model-based approach?

Design-based method best choice when:

- we want to estimate the distribution function or parameters thereof (mean, median, P90 etc.) for the area as a whole or for subareas;
- objective estimates of target properties are required, i.e. no subjective judgement on ‘representativeness’, no subjective model choices;
- objective estimates of estimation variance or confidence intervals are required, i.e. validity.
Design-based method best choice when:

- we want to estimate the distribution function or parameters thereof (mean, median, P90 etc.) for the area as a whole or for subareas;
- objective estimates of target properties are required, i.e. no subjective judgement on ‘representativeness’, no subjective model choices;
- objective estimates of estimation variance or confidence intervals are required, i.e. validity.
Design-based or model-based approach?

Design-based method best choice when:

- we want to estimate the distribution function or parameters thereof (mean, median, P90 etc.) for the area as a whole or for subareas;
- **objective** estimates of target properties are required, i.e. no subjective judgement on ‘representativeness’, no subjective model choices;
- **objective** estimates of estimation variance or confidence intervals are required, i.e. validity.
Design-based or model-based approach?

Design-based method best choice when:

- we want to estimate the distribution function or parameters thereof (mean, median, P90 etc.) for the area as a whole or for subareas;
- **objective** estimates of target properties are required, i.e. no subjective judgement on ‘representativeness’, no subjective model choices;
- **objective** estimates of *estimation variance* or *confidence intervals* are required, i.e. validity.
Model based method best choice when:

- we want to map the target property (making predictions for unvisited locations);
- sample size large enough for calibrating a model of variation (e.g. variogram: $n > 100$);
- strong autocorrelation exists, from which we may profit in mapping;

Model based method best choice when:

- we want to map the target property (making predictions for unvisited locations);
- sample size large enough for calibrating a model of variation (e.g. variogram: \( n > 100 \));
- strong autocorrelation exists, from which we may profit in mapping;

Design-based or model-based approach?, continued

Model based method best choice when:

- we want to map the target property (making predictions for unvisited locations);
- sample size large enough for calibrating a model of variation (e.g. variogram: $n > 100$);
- strong autocorrelation exists, from which we may profit in mapping;

Model based method best choice when:

- we want to map the target property (making predictions for unvisited locations);
- sample size large enough for calibrating a model of variation (e.g. variogram: $n > 100$);
- strong autocorrelation exists, from which we may profit in mapping;

Model based method best choice when:

- we want to map the target property (making predictions for unvisited locations);
- sample size large enough for calibrating a model of variation (e.g. variogram: \( n > 100 \));
- strong autocorrelation exists, from which we may profit in mapping;

2. Sampling for mapping of changes in soil carbon stocks (*monitoring*)

---

**Static**

**Synchronous**
2. Sampling for mapping of changes in soil carbon stocks (monitoring)
2. Sampling for mapping of *changes* in soil carbon stocks (*monitoring*)

Serially alternating

Supplemented
Changes: differences, trends or effects?

difference,
\[ d_{10-0} = -0.6 \text{ kg m}^{-2} \]
\[ d_{20-10} = 4.9 \text{ kg m}^{-2} \]

linear trend,
0.33 kg m\(^{-2}\) year\(^{-1}\)

effect of policy,
intervention at \( t = 10 \),
step decay model,
effect = 6.0 kg m\(^{-2}\) after 10 years
### 2. Sampling for monitoring

#### Possible selection modes for various aims of monitoring

<table>
<thead>
<tr>
<th>space</th>
<th>time</th>
</tr>
</thead>
<tbody>
<tr>
<td>purposive</td>
<td>purposive</td>
</tr>
<tr>
<td>purposive</td>
<td>space-time mapping</td>
</tr>
<tr>
<td>random</td>
<td>estimation of trends in spatial means, totals, areal proportions*</td>
</tr>
<tr>
<td>random</td>
<td>model-free estimation of space-time means</td>
</tr>
</tbody>
</table>

3. Sampling for validation

- **Validation**: testing whether a map or a model satisfies its purpose
- **Objectivity** is crucial in validation (so that the validation procedure cannot get the blame for bad results)
- Collection of additional data by probability sampling is therefore recommended

3. Sampling for validation

- **Validation**: testing whether a map or a model satisfies its purpose
- Objectivity is crucial in validation (so that the validation procedure cannot get the blame for bad results)
- Collection of additional data by probability sampling is therefore recommended

3. Sampling for validation

▶ **Validation**: testing whether a map or a model satisfies its purpose

▶ **Objectivity** is crucial in validation (so that the validation procedure cannot get the blame for bad results)

▶ Collection of additional data by probability sampling is therefore recommended

3. Sampling for validation

- **Validation**: testing whether a map or a model satisfies its purpose
- **Objectivity** is crucial in validation (so that the validation procedure cannot get the blame for bad results)
- Collection of additional data by probability sampling is therefore recommended

Concluding remarks

- Sampling for mapping (spatial prediction): directed sampling or probability sampling, striving for fair spatial distribution.
- Sampling for monitoring of trends in spatial means: directed sampling in time, probability sampling in space. Revisiting of locations is recommended.
- Sampling for validation: collection of additional data by probability sampling is recommended.
Concluding remarks

- Sampling for mapping (spatial prediction): directed sampling or probability sampling, striving for fair spatial distribution.
- Sampling for monitoring of trends in spatial means: directed sampling in time, probability sampling in space. Revisiting of locations is recommended.
- Sampling for validation: collection of additional data by probability sampling is recommended.
Concluding remarks

- Sampling for mapping (spatial prediction): directed sampling or probability sampling, striving for fair spatial distribution.
- Sampling for monitoring of trends in spatial means: directed sampling in time, probability sampling in space. Revisiting of locations is recommended.
- Sampling for validation: collection of additional data by probability sampling is recommended.
Concluding remarks

- Sampling for mapping (spatial prediction): directed sampling or probability sampling, striving for fair spatial distribution.
- Sampling for monitoring of trends in spatial means: directed sampling in time, probability sampling in space. Revisiting of locations is recommended.
- Sampling for validation: collection of additional data by probability sampling is recommended.
Thank you!