

Using integrated geospatial data in official statistics

Susie Jentoft

BaNoCoSS 19.06.2019



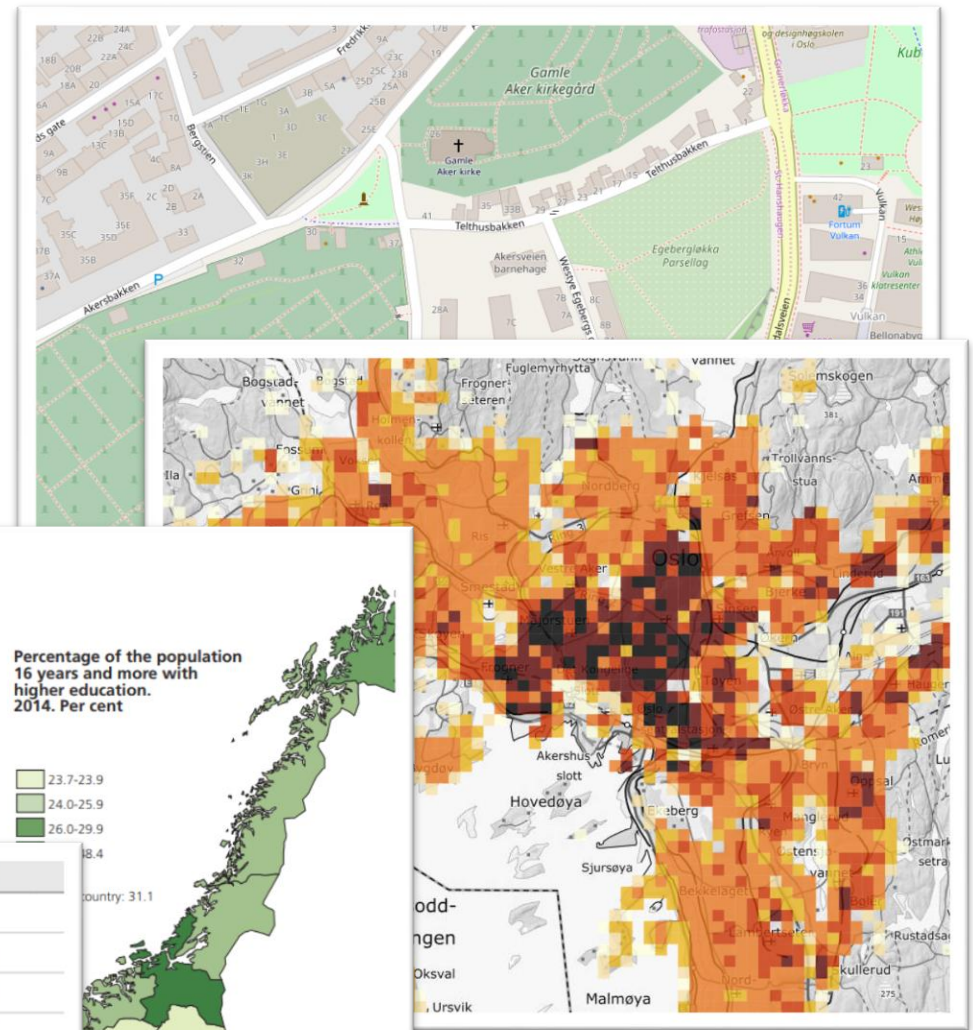
Statistisk sentralbyrå
Statistics Norway

What is geospatial data?

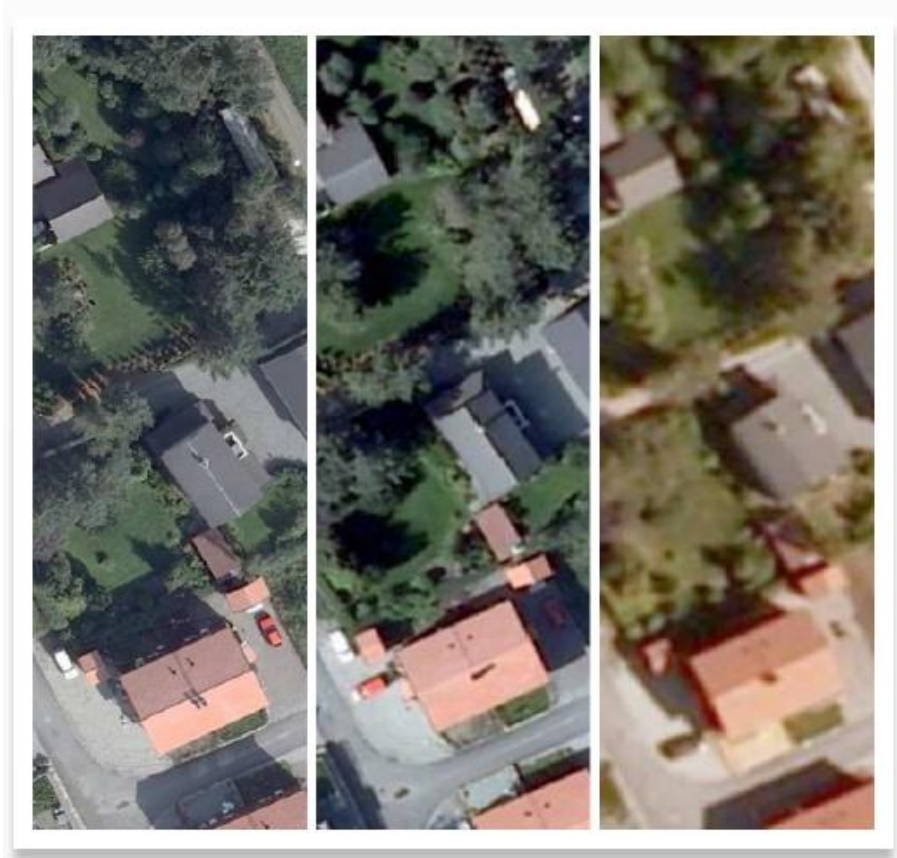
- Information with a geographic aspect



A	B	C
ID	Lat	Long
2141	59.911	10.756
2142	59.014	10.572
2143	58.798	10.783
2144	59.123	11.012
2145	59.292	10.834
2146	58.786	10.794
2147	58.903	10.934
2148	59.130	10.479
2149	59.495	11.050
2150	59.557	10.499



What are orthophotos?



- Geometrically corrected aerial photos
- Adjusted to a map projection
- Norway: country coverage
 - Orthophoto 10: 4-15cm (cities)
 - Orthophoto 20: 20cm (towns)
 - Orthophoto 50: 25-50cm (country)

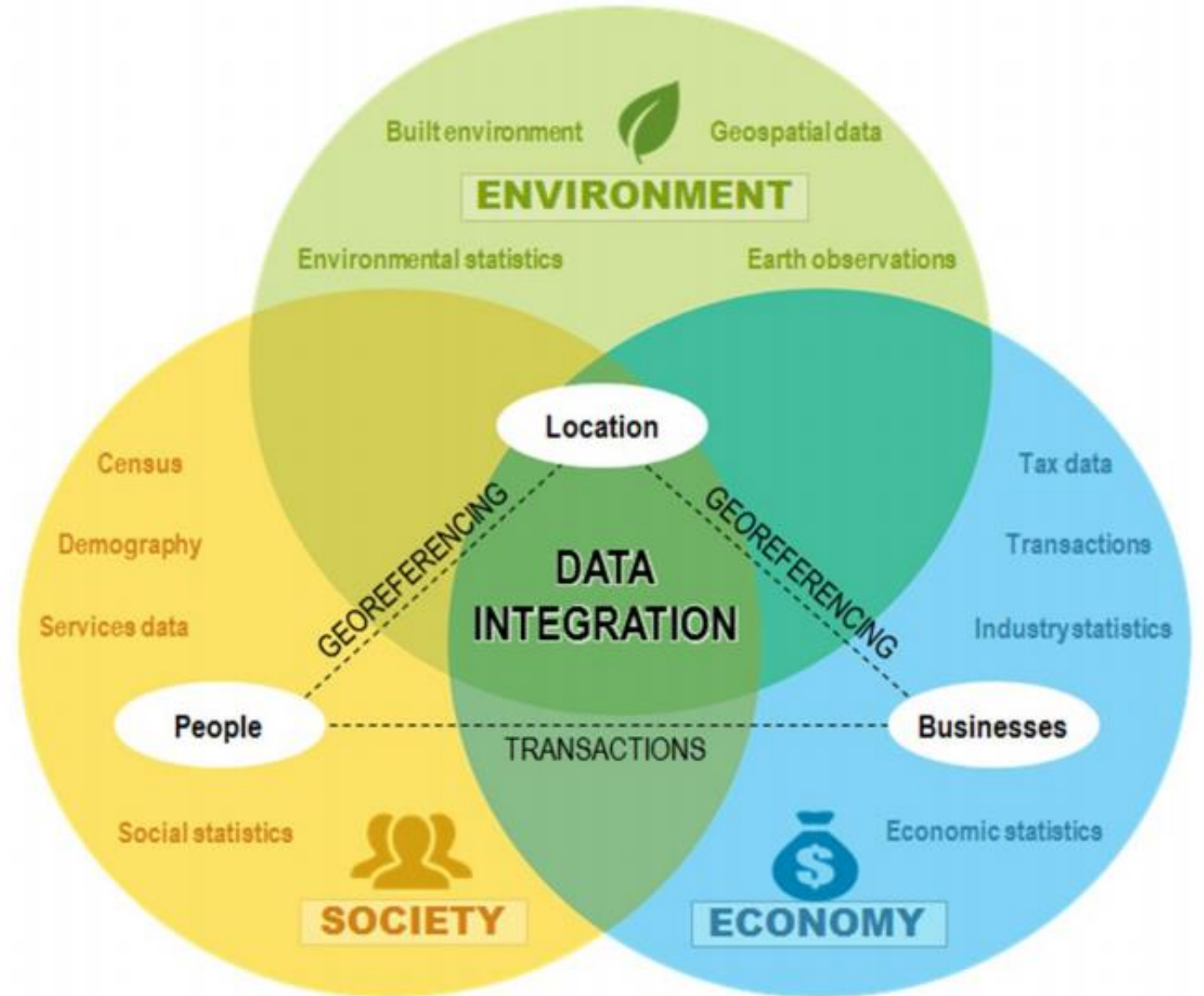
Geospatial data in official statistics

- Within survey paradigm
 - Sampling frame
 - Data collection
 - Editing
 - Post-stratification/calibration
 - Dissemination/visualisation
- Quality assessments
- New statistics

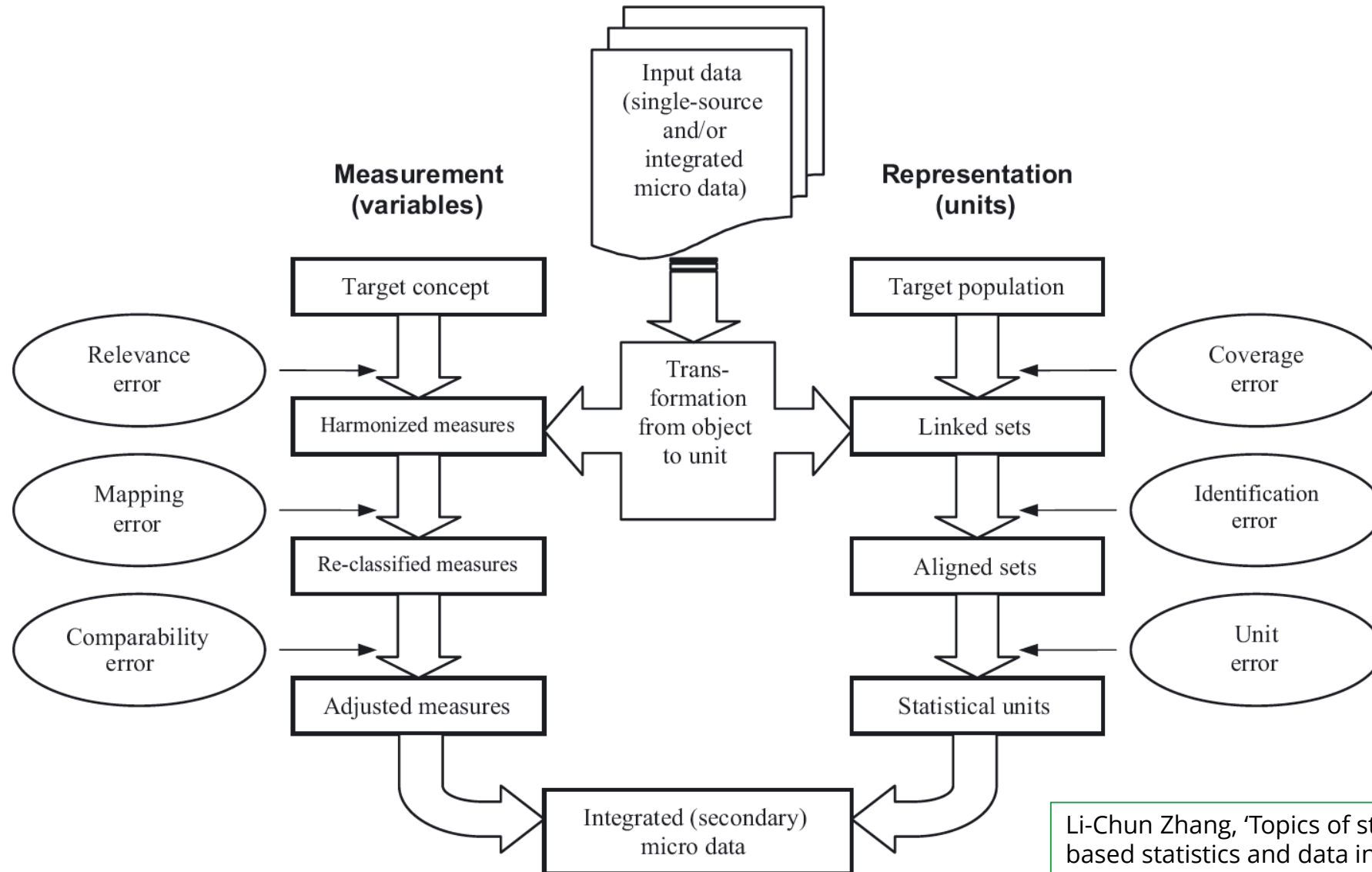


Data integration

Global statistical geospatial framework:
linking statistics and place (UN, 2018)



Two-phase life cycle model for integrated data



Li-Chun Zhang, 'Topics of statistical theory for register-based statistics and data integration', *Statistica Neerlandica* (2012) Vol. 66, Nr. 1, pp. 41-63

Case I: Holiday homes

- **Problem:** Every year, 'new' holiday homes are registered without a year of build.
- Two scenarios:
 - Reporting error
 - Delayed registration (under-coverage)
- **Goal:** Estimate/impute a build year for those missing.



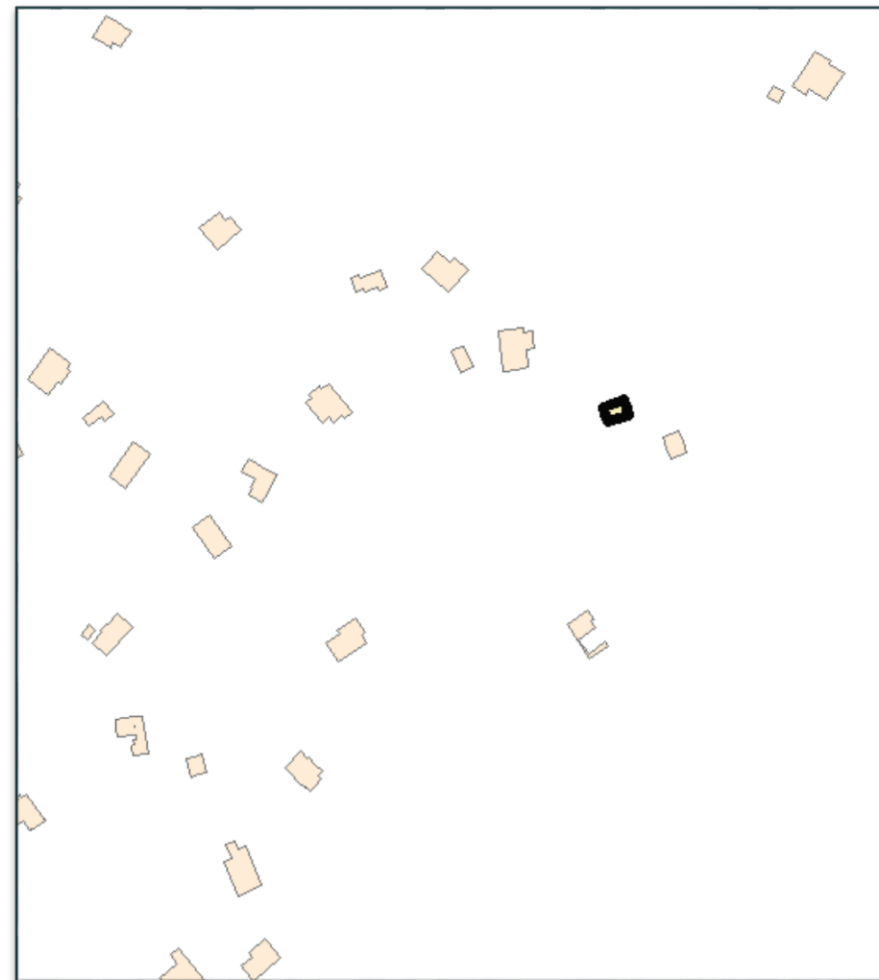
Project outline

- Sample (200) among those newly registered missing build year
- Linked sample with historic orthophotos
- Determine date of build (manually)
- Estimate under-coverage and establish imputation routine



Data

- Based on new holiday homes in the previous year as of 01.01.2017
 - 823 with unknown build year (4500 with build year)
- Building register (Matrikkel) linked to Building map database (FKB-bygg)
- ArcGIS and WMS



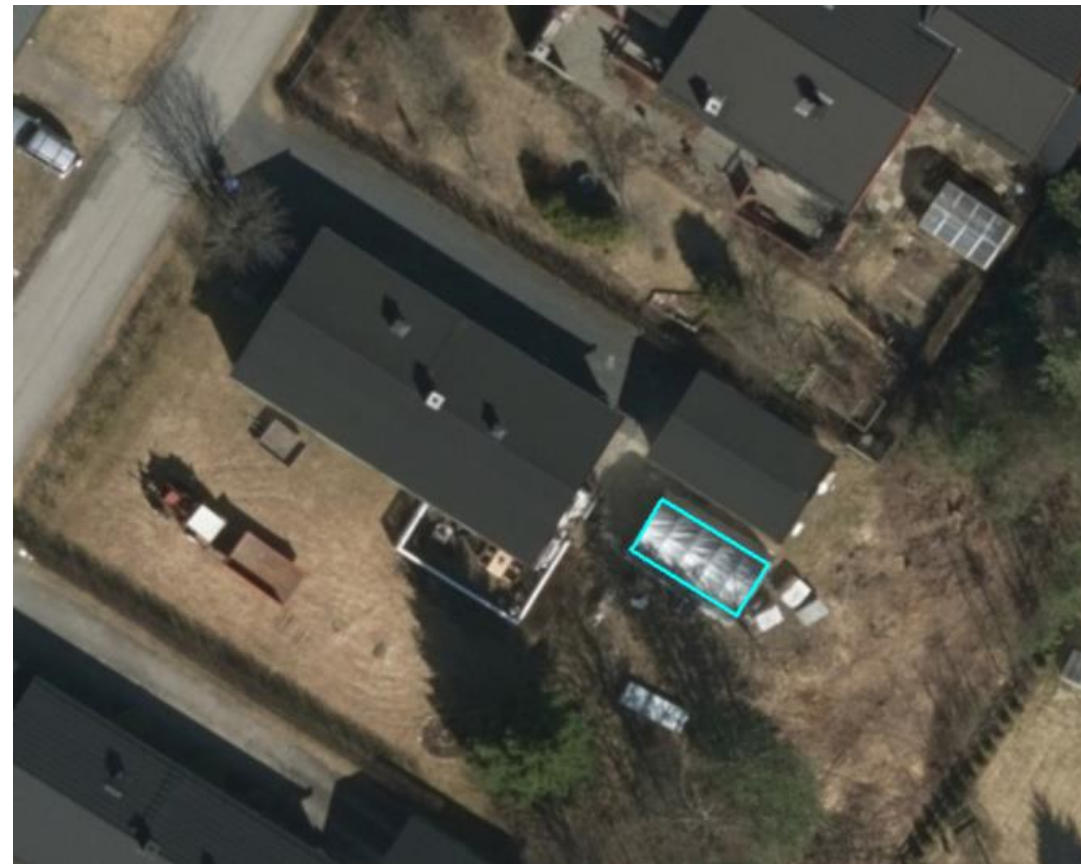
Varying picture quality



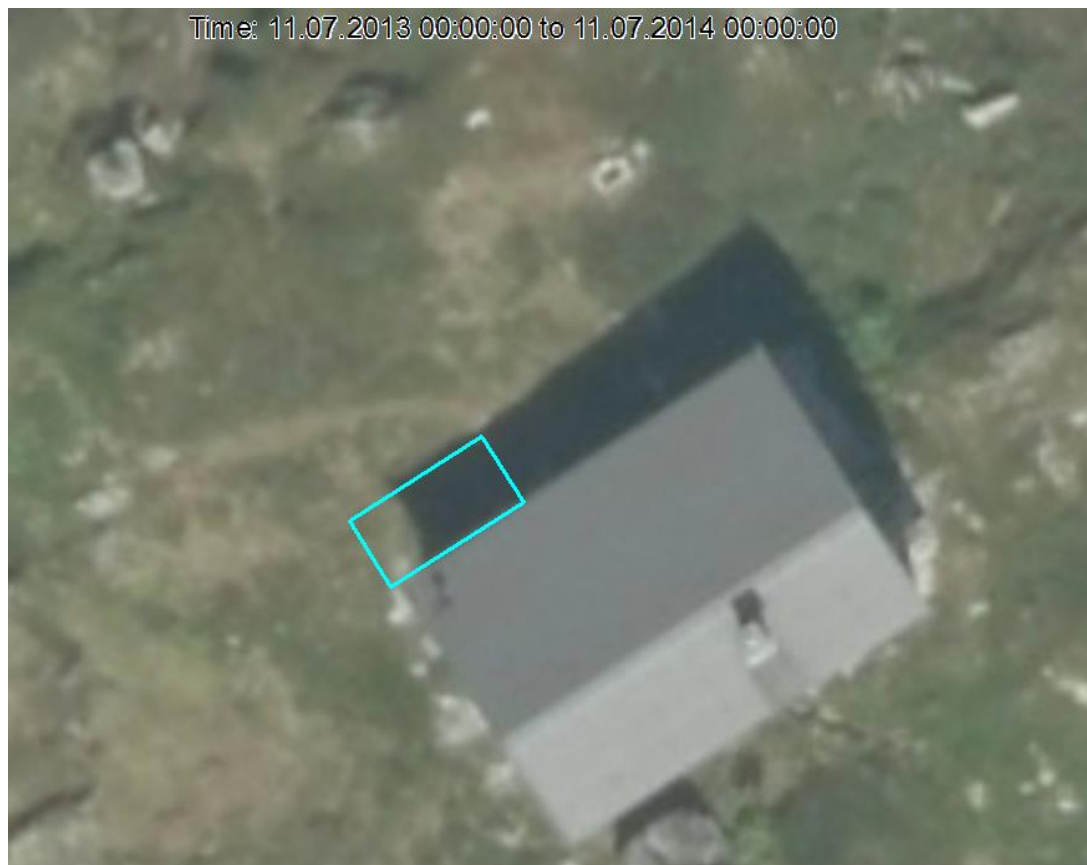
Changes



Examples of errors

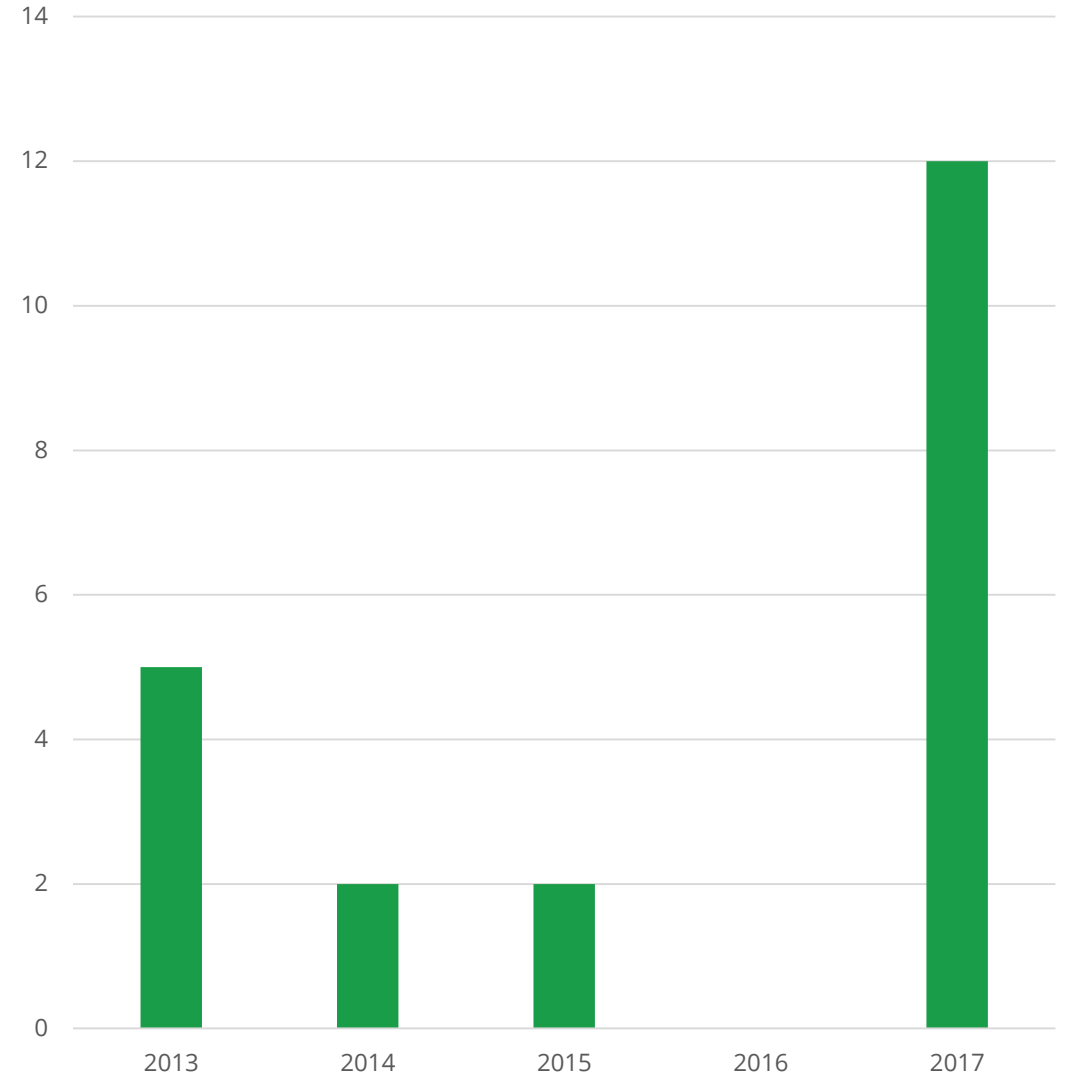


Examples of errors



Some results

- 5% not linked to a holiday home
- Only 10% built in previous 5 years
- Imputation for new/old variable developed



Case II: Land use of previous agricultural areas



?



Statistisk sentralbyrå
Statistics Norway

Changes in land use

- What is the land used for?
- Stratified sample ($n = 600$, $N = 6067$) of de-registered agricultural areas.
- Sample linked to orthophotos
- Classified as garden, wilderness, storage, road, fields or sporting areas (manually)
- Classified as agricultural (manually)



Results

- Most common was gardens (40% of areas)
- 10% contained areas appearing to be used for agricultural.



Thank you!

susie.jentoft@ssb.no



Statistisk sentralbyrå
Statistics Norway