Nation-wide mapping of fractional land cover with regression-based unmixing and Sentinel-2 imagery

David Frantz, Franz Schug

EO-Lab, Humboldt-Universität zu Berlin

Earth Observation Lab Colloquium, Humboldt-Universität zu Berlin, 20 July 2019, Berlin

Outline

1 Introduction Mat Stocks (DF)

2 Introduction Workflow (DF / I

3 Land Cover Mapping incl. OSI

4 Settlement Type Mapping (FS

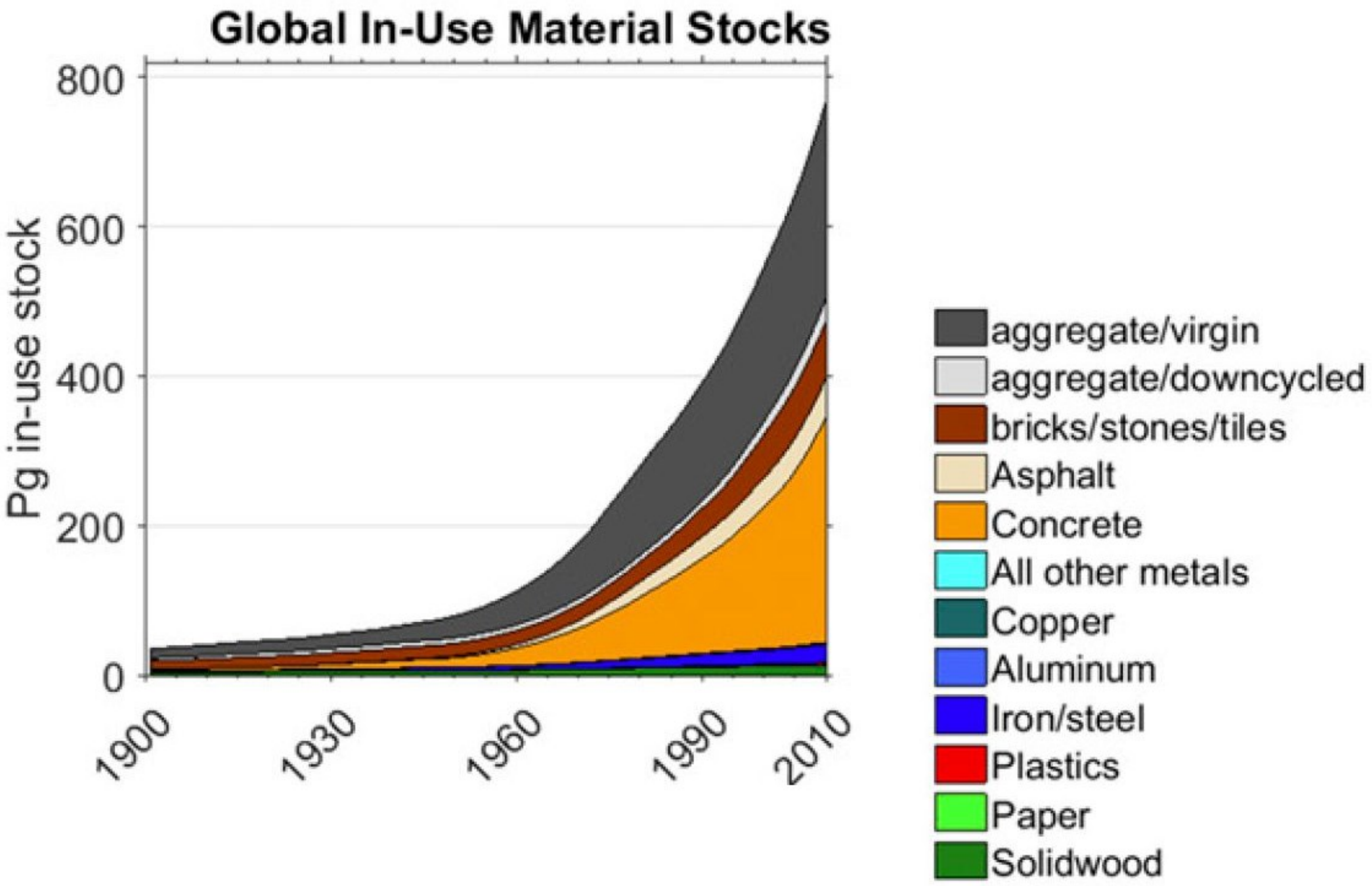
5 Height Mapping incl. Volume

6 Stocks Mapping (DF)

7 Other Countries (DF)

		FS	DF	Min	Min FS	Min
1	Intro		1	5	0	
2	Workflow		1	3	0	
3	Data & STMs		1	5	0	
4	Types	1		3	3	
5	LC incl. OSM	1		10	10	
6	Height		1	10	0	
7	Stocks		1	8	0	
8	Other	1		3	3	
9	Countries		1	3	0	
				50	16	

Background



Kraussmann et al. (2017), PNAS, 114

- Materials stocks are resources stored for **long term** periods
- **Construction materials** are the most used materials for stock accumulation.
- Global **material stock accumulation** since 1950 is important
- **Where** is the stock and how does it relate to **socio-economic factors**.
- Currently on a **nation-wide** level only.

Background

Inflow-driven approach

- **Deduces** stock estimation from (national) **statistics**.
- Can be easily applied to **large areas**.

e.g. Fishman et al. 2014, Wiedenhofer et al. 2015, Krausmann et al. 2017, …

Stock-driven approach

- **Induces** large area stock amounts from **single features**.
- Studies **small areas** and **aggregates** possible stock distribution.

e.g. Tanikawa et al. 2010, Kleemann et al. 2016, Kleemann et al. 2017,

Remote Sensing for material stocks mapping

Objectives and Workflow

- Three-step approach:

1. **Mapping human-made buildings and infrastructure**
2. Characterize human settlements
3. Approximate material stock patterns

Sentinel-2 MSI

Best temporal coverage: max 5/10 days since 2017

Spatial resolution: 10/20m

Spectral bands: 10 (13)

- **Research goal**: Map buildings and infrastructures through land cover detection and surface cover quantification
- **Key information**: Presence and density of artificial structures and human settlements as well as their physical surface composition.

Objectives and Workflow

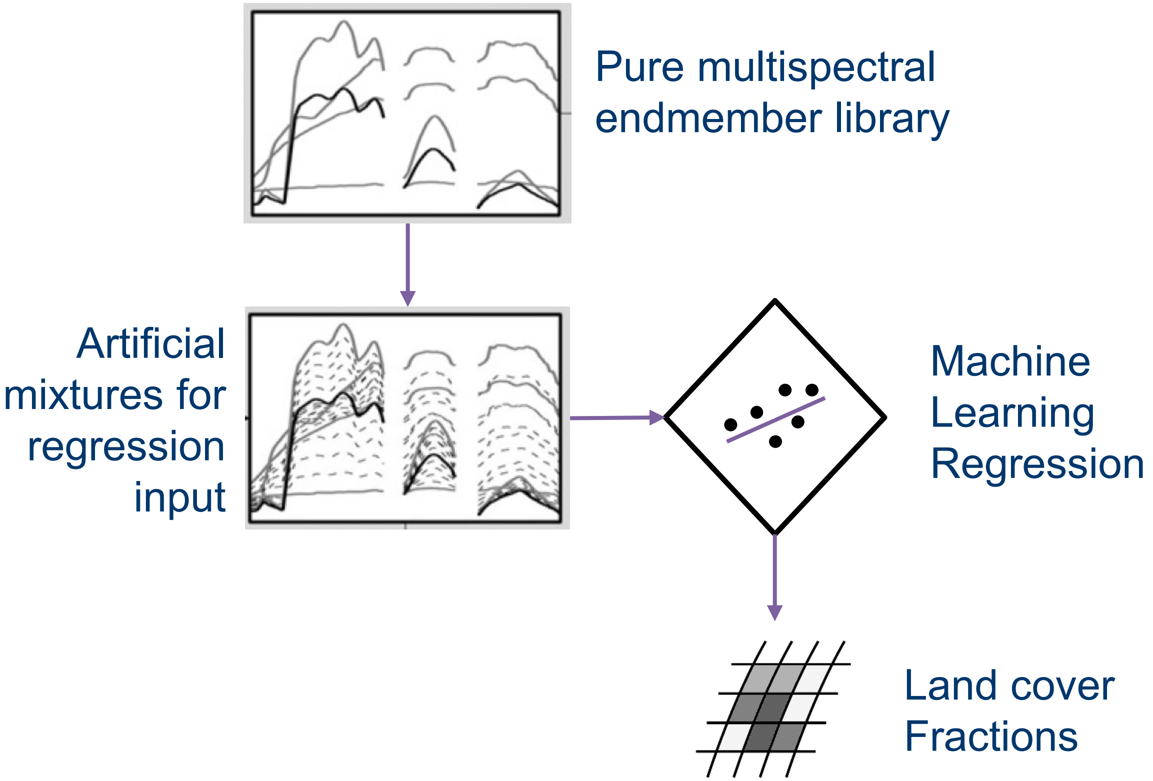
- → Quantify land cover over large areas
- → Preserve information on settlement gradients and gradual settlement characteristics
- → Use robust models with little seasonal impact (from, e.g., phenology)
  - → Seasonality is no indicator for stocks
  - → Seasonality is different across regions

Methods

‘Preserve gradual information‘

- Sub-pixel mapping: Regression-based unmixing with synthetically mixed endmember spectra

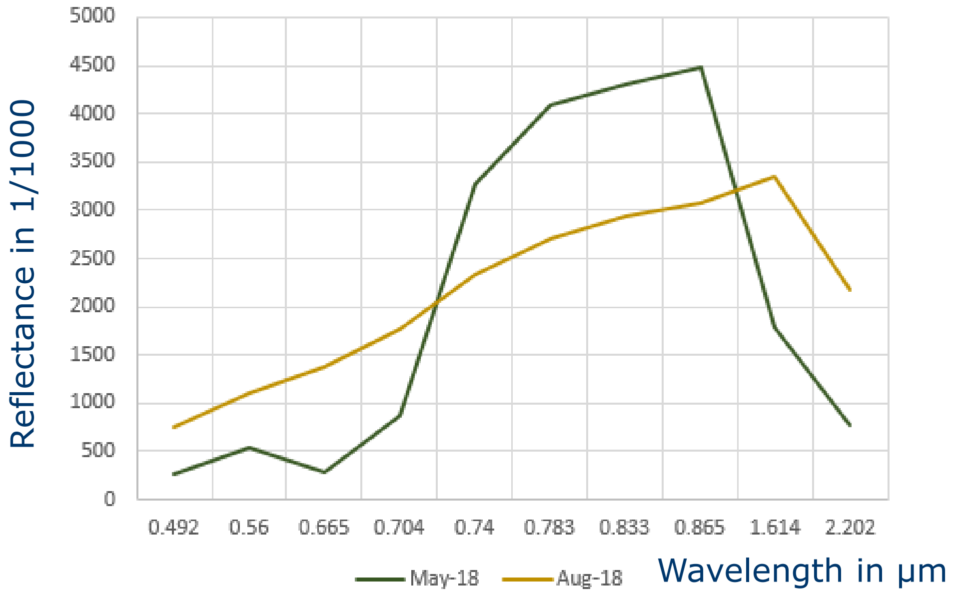
- Library: ca. 500 reference locations (Berlin/Brandenburg + selected sites) for 10 classes. Mapping is performed for 4 classes:
  - built-up surfaces and infrastructure
  - woody vegetation
  - non-woody vegetation
  - soil



Methods

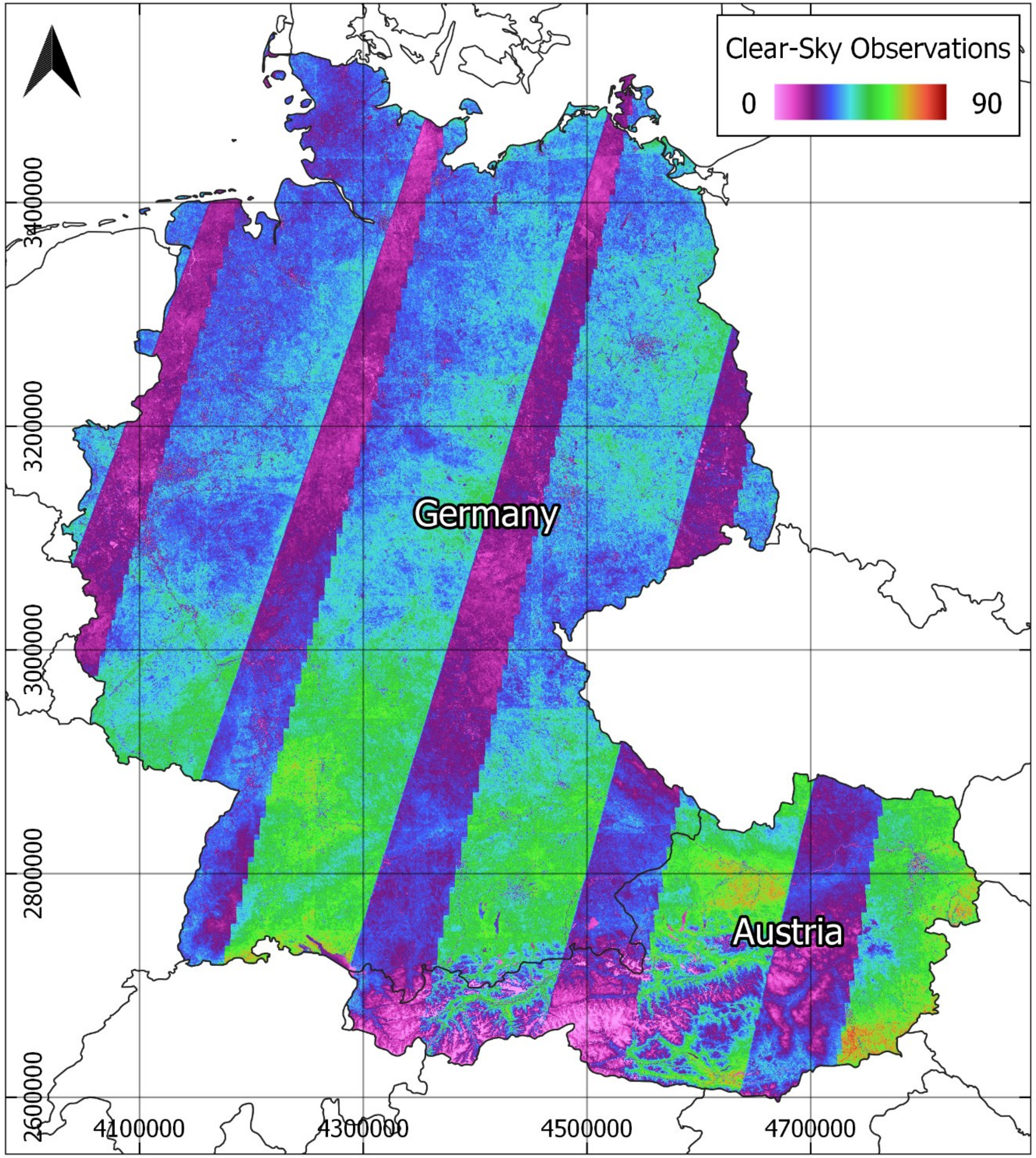
‚Robust models with little seasonal impact‘

- We use **intra-annual spectral-temporal metrics**
  - Spectral statistics, e.g. spectral mean / quartiles
  - No manual data selection process
  - Diminish seasonal effects, Decrease confusion of resembling surfaces



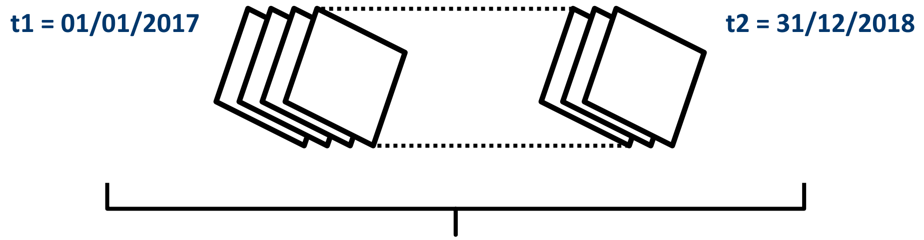
Methods

- Study area: Germany / Austria
- Data availability 20 to 60 observations per pixel



study area: clear-sky observations per pixel

Methods



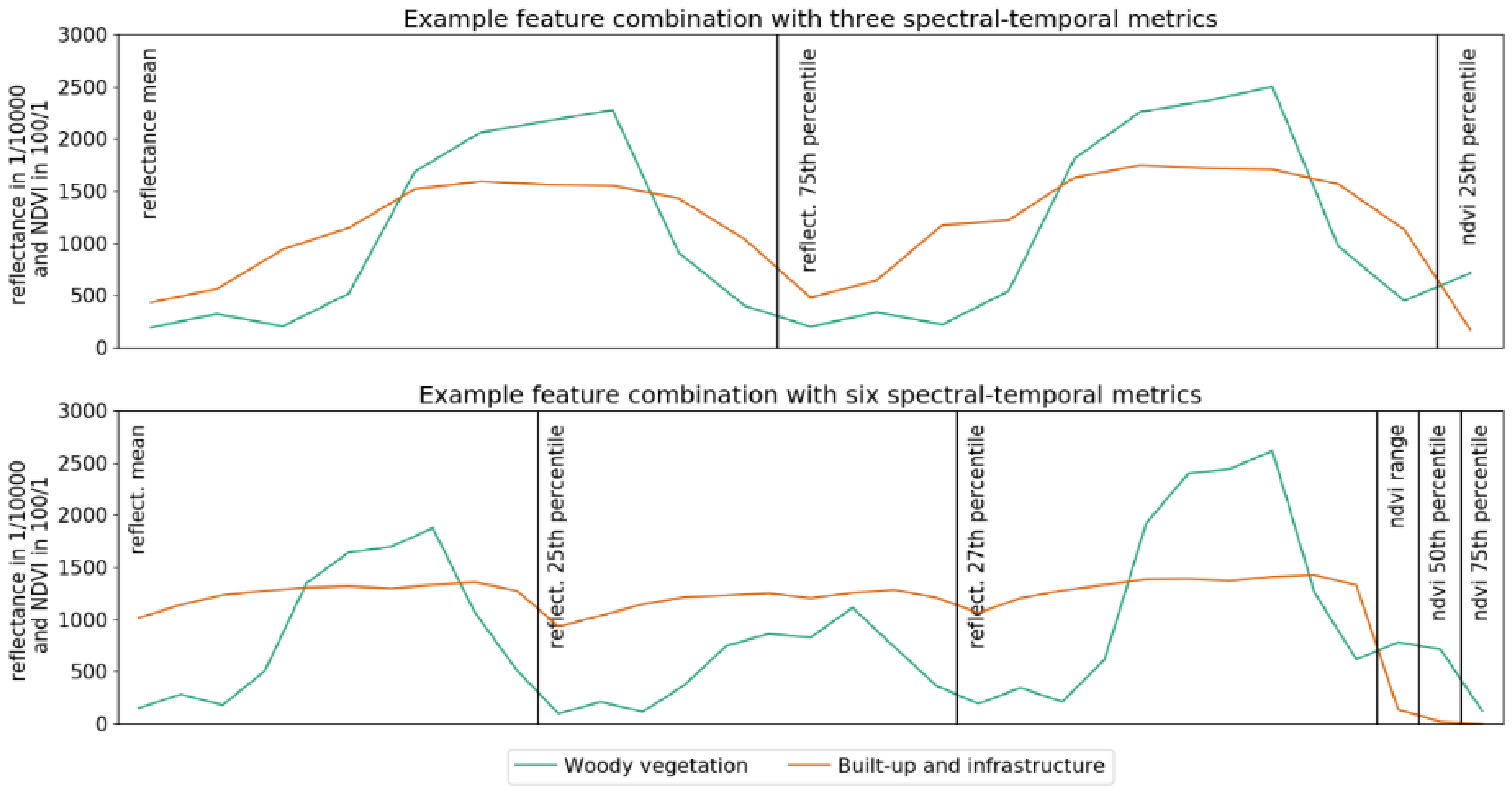
Spectral-temporal metrics:
Standard deviation
Inter-quartile range
Range
Mean
Median
Minimum
Maximum
25th percentile
75th percentile

NDVI:
Standard deviation
Inter-quartile range
Range
Mean
Median
Minimum
Maximum
25th percentile
75th percentile

All combinations of 18 elements = 262.143 models

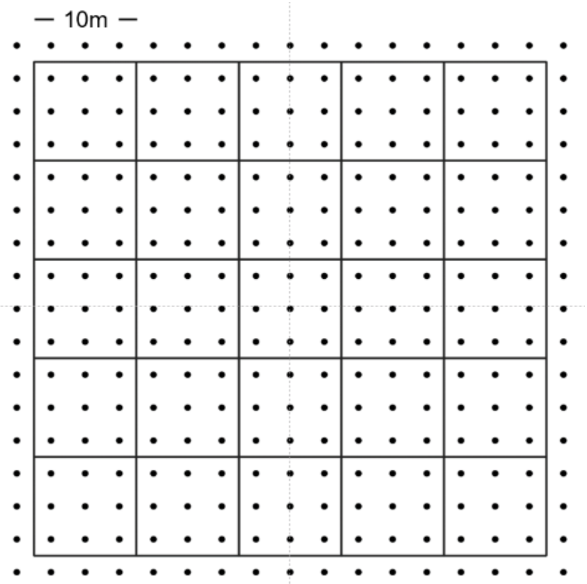
- Are there features or feature combinations particularly suitable as an input to regression-based modelling?
- Group and randomly select metrics as data input
- Compare a number of 1,275 models with different respective input data

Methods

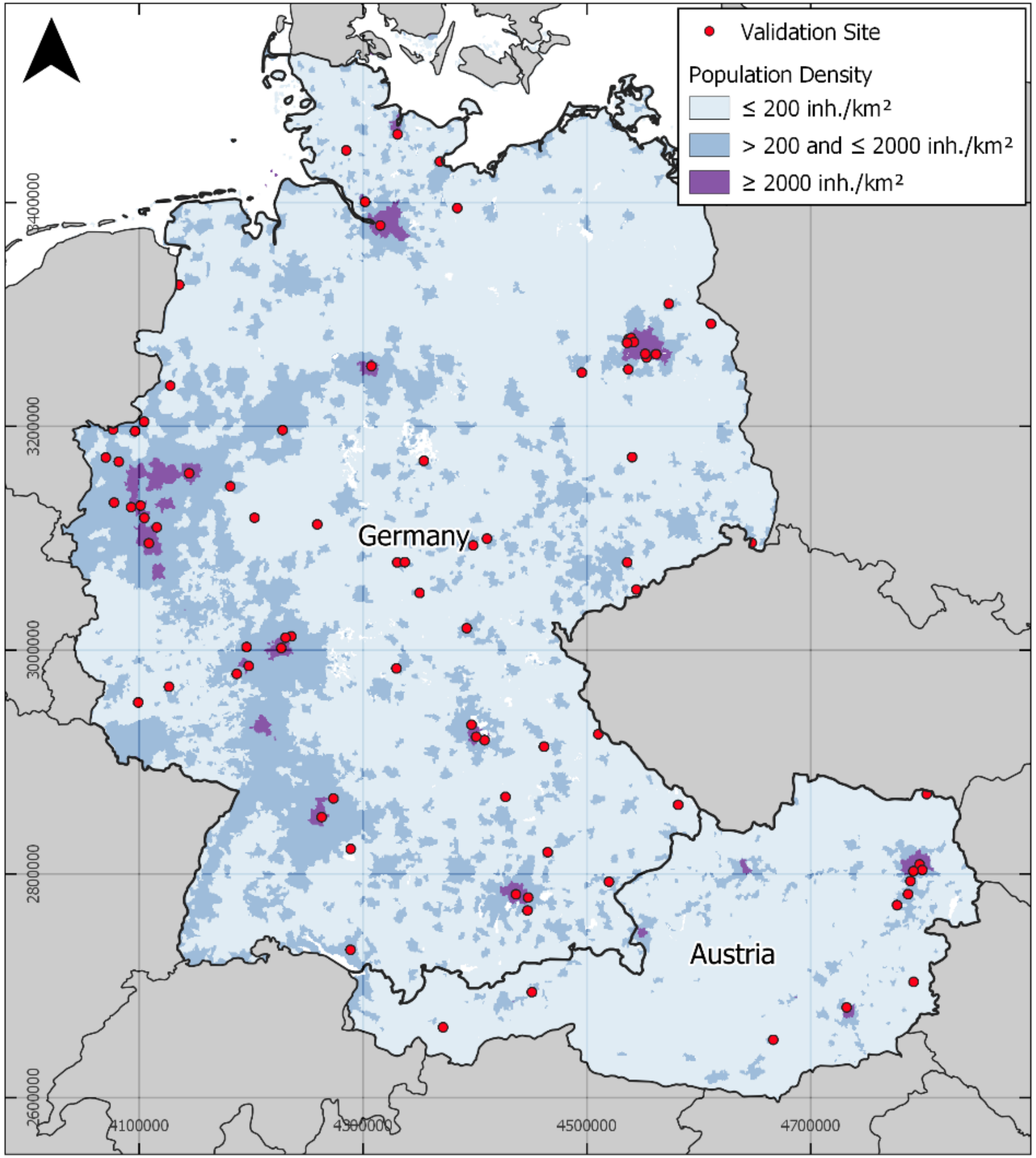

Example feature combination with three spectral-temporal metrics
reflectance in 1/10000 and NDVI in 100/1
3000
2500
2000
1500
1000
500
0
reflectance mean
reflect. 75th percentile
ndvi 25th percentile
Example feature combination with six spectral-temporal metrics
reflect. mean
reflect. 25th percentile
reflect. 27th percentile
ndvi range
ndvi 50th percentile
ndvi 75th percentile
Woody vegetation
Built-up and infrastructure

Methods

- Validation: 84 sites, stratified by population density (3 classes) AND Copernicus Imperviousness (0-0.25, 0.25-0.5, 0.5-0.75, 0.75-1)
- Validation grid at 10m, 20m and 50m spatial resolution
- 24,276 pts
- Fractions



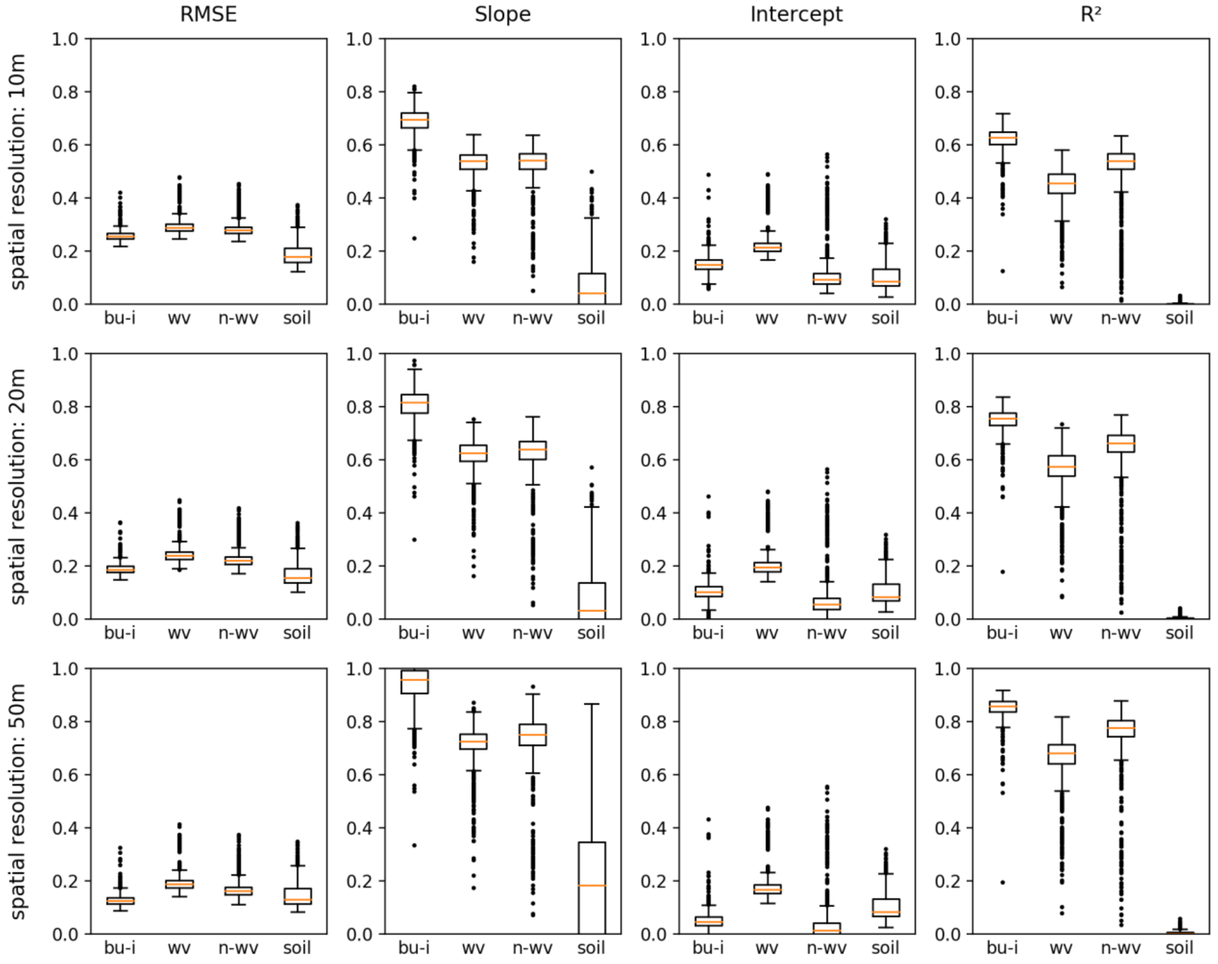
validation grid



validation Sites Density from SEDAC Gridded Population of the World POPGRID v4

Results

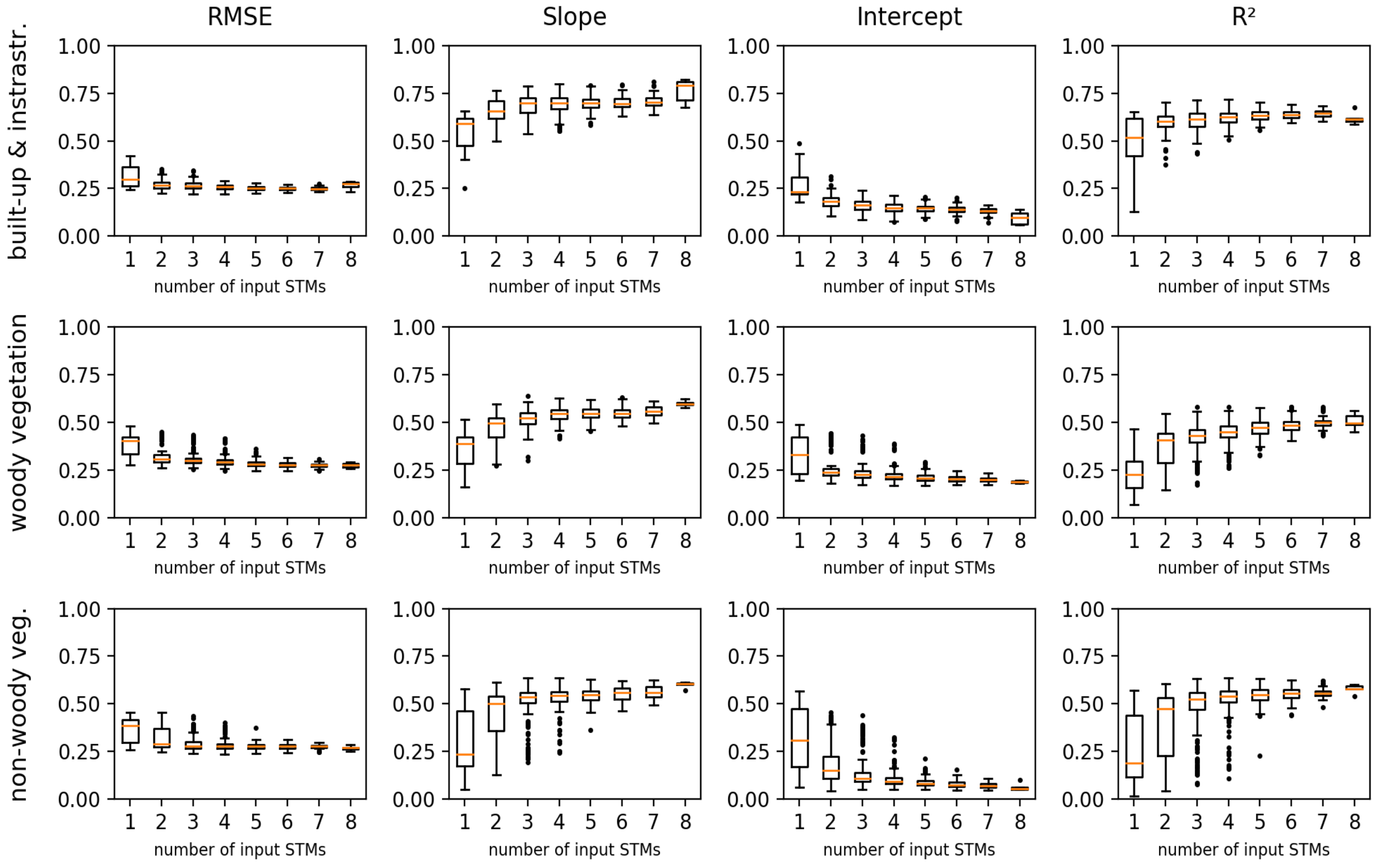
- Quality of 1,275 models by spatial resolution and class



- Distribution of RMSE is rather stable.
- Built-up and non-woody vegetation class reach better slopes and coefficients of determination.
- Quality increases with decreased resolution, patterns remain.

Results

- Quality of 1,275 models by number of input metrics

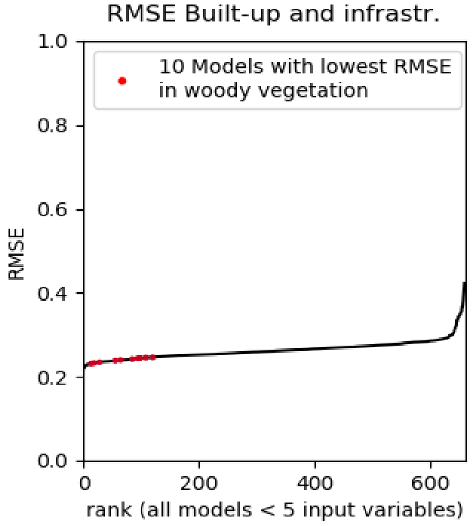


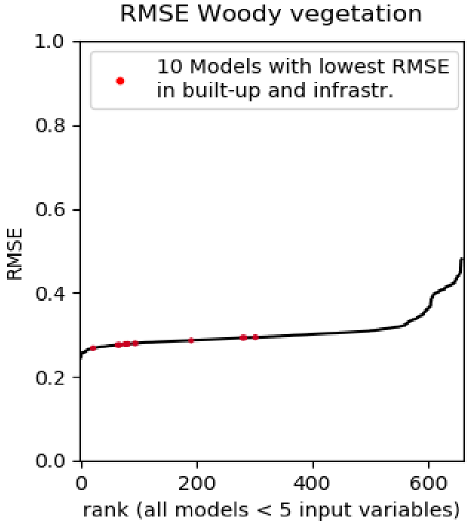
- A number of three to four input metrics is usually sufficient to reach best model quality.

STM:

Results

- Which input spectral-temporal metrics perform well in all classes and all quality metrics?



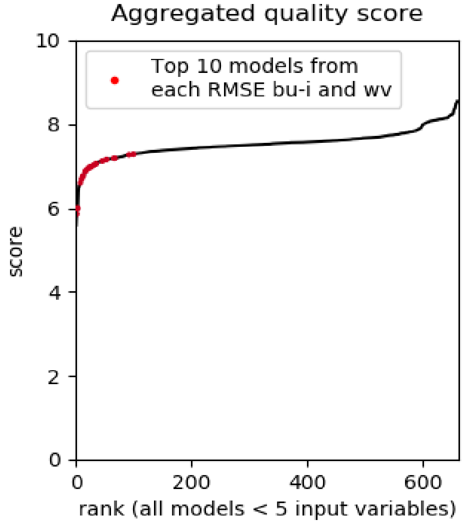


Transformation function that penalizes low performance:

$$f(x) = \frac{1}{4}\log_{10}(x + 0.001) + 1$$

with x as the normalized RMSE, $1/R^2$ and |(1 - Slope)| for each model.

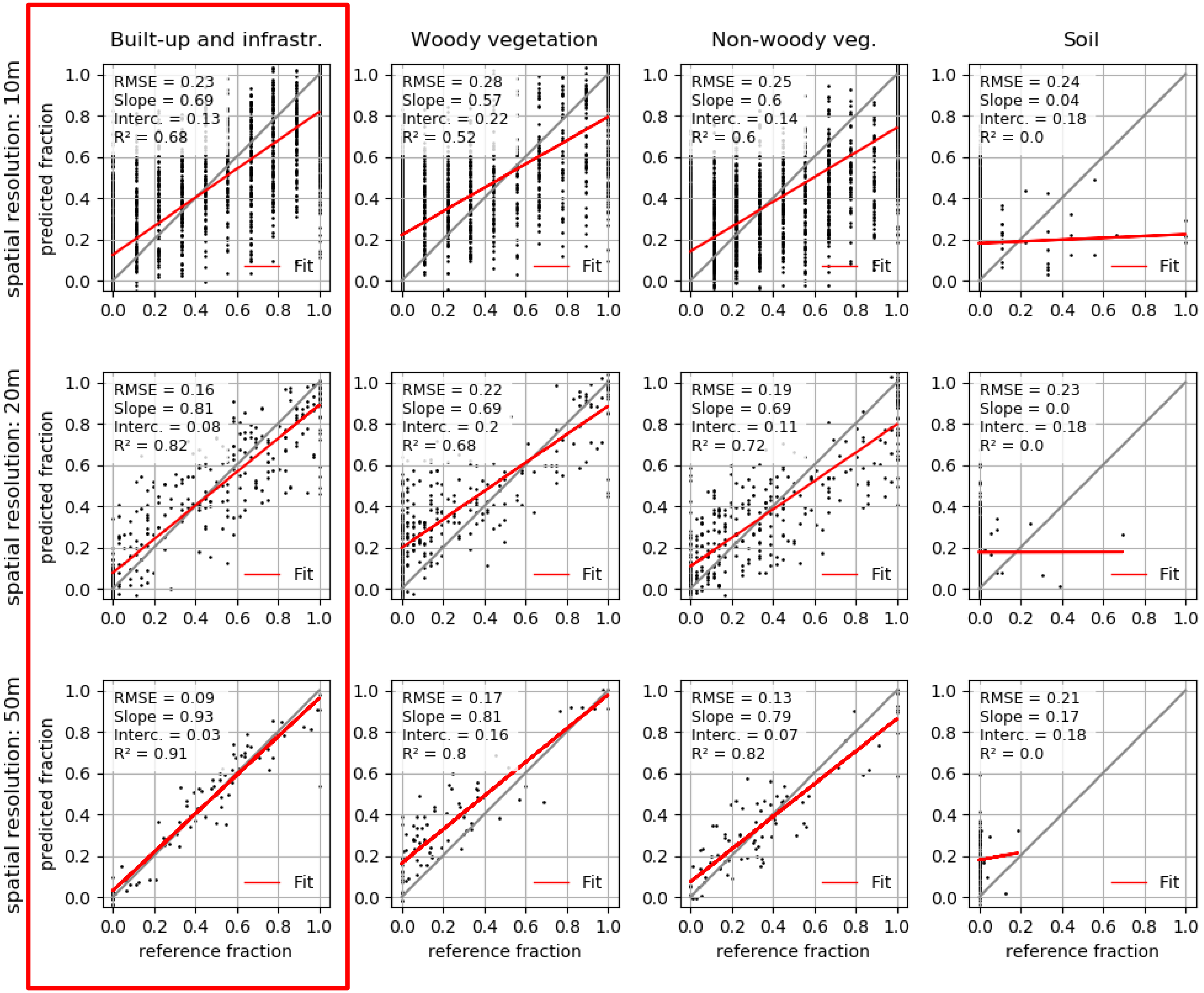
Sum of f(x) for three classes and three quality metrics = quality score.

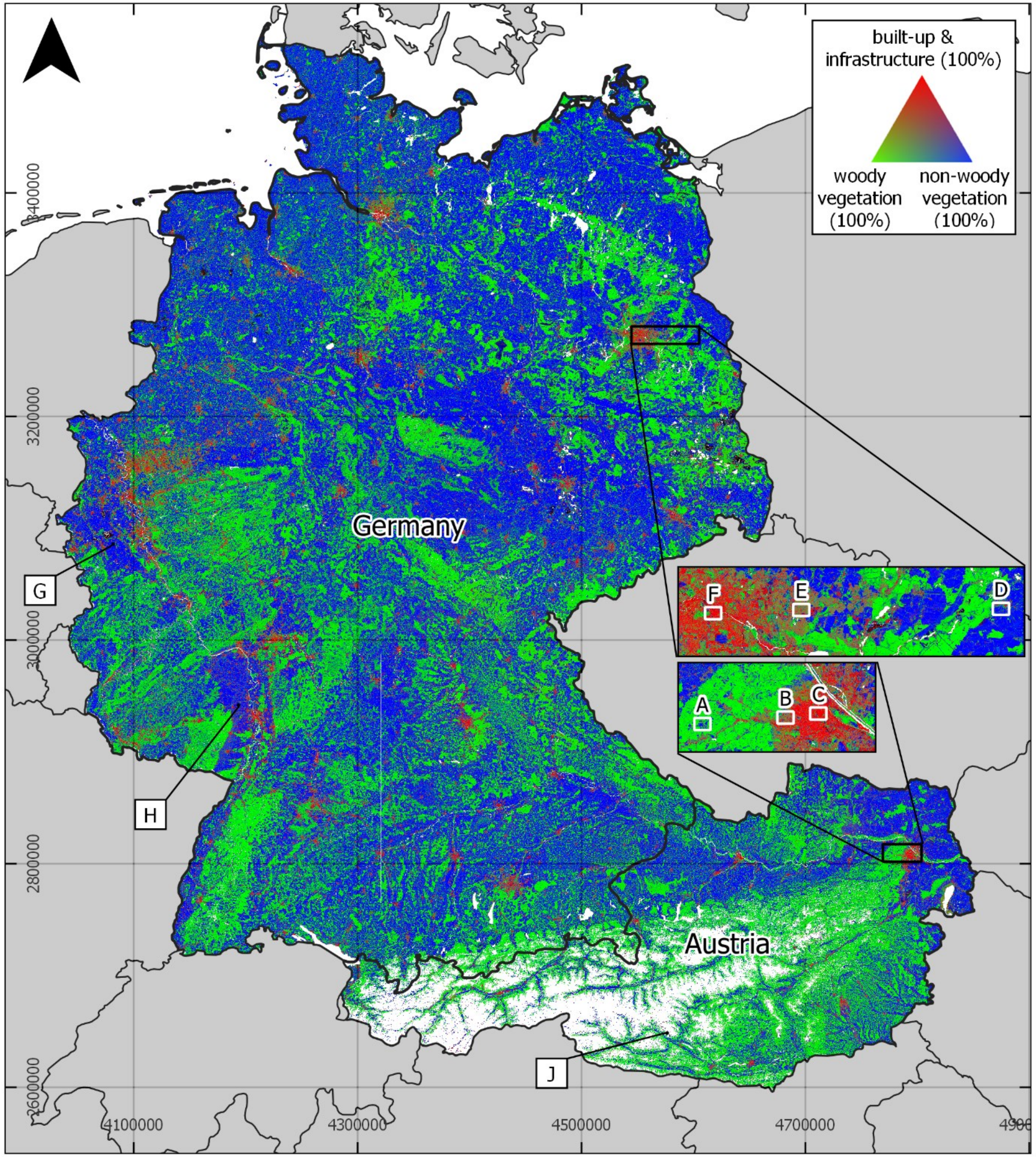


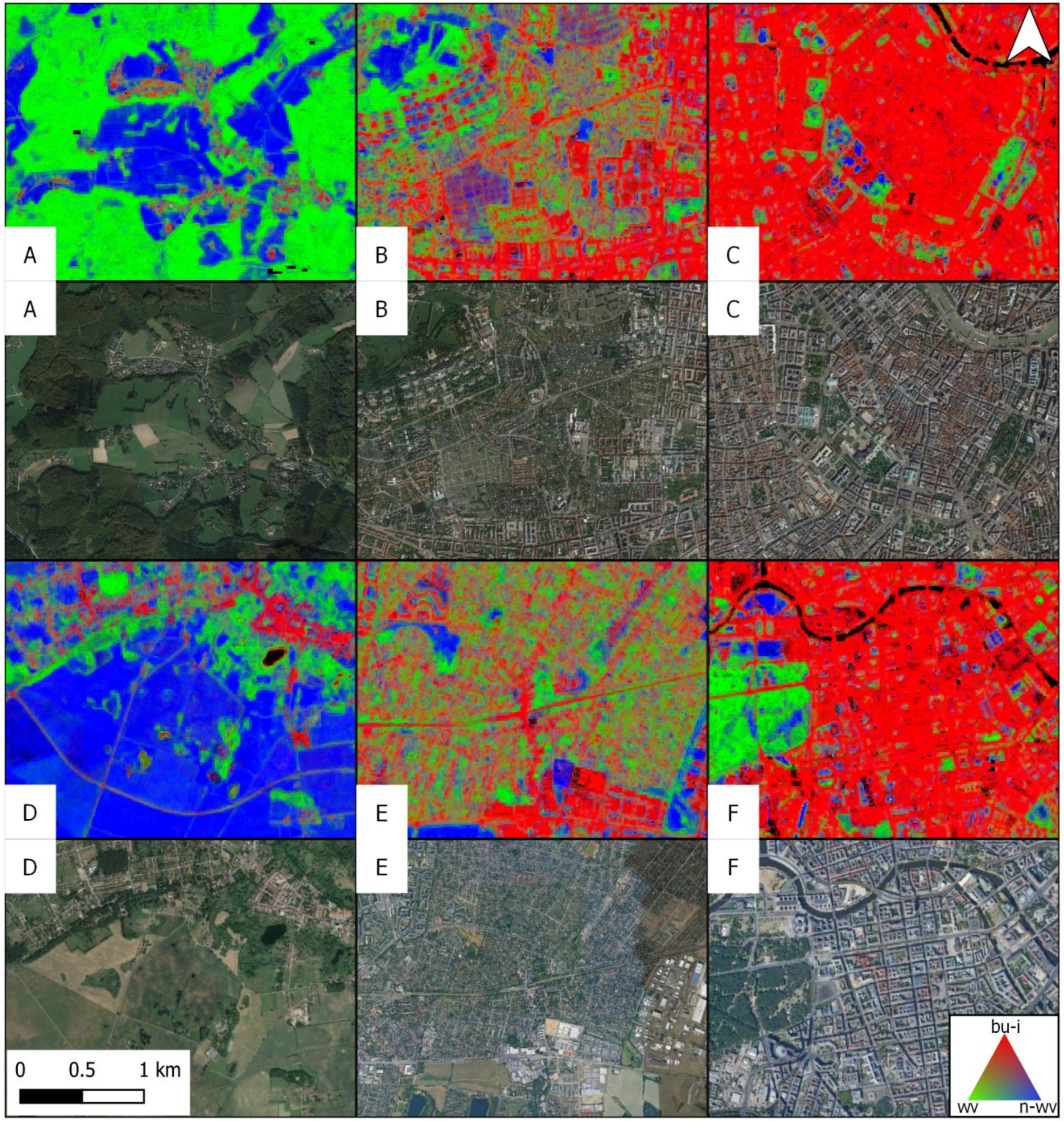
Based on the top-ranked models we chose a metrics combination for the large area mapping:
Median reflectance, 25th & 75the percentile of reflectance and max. NDVI

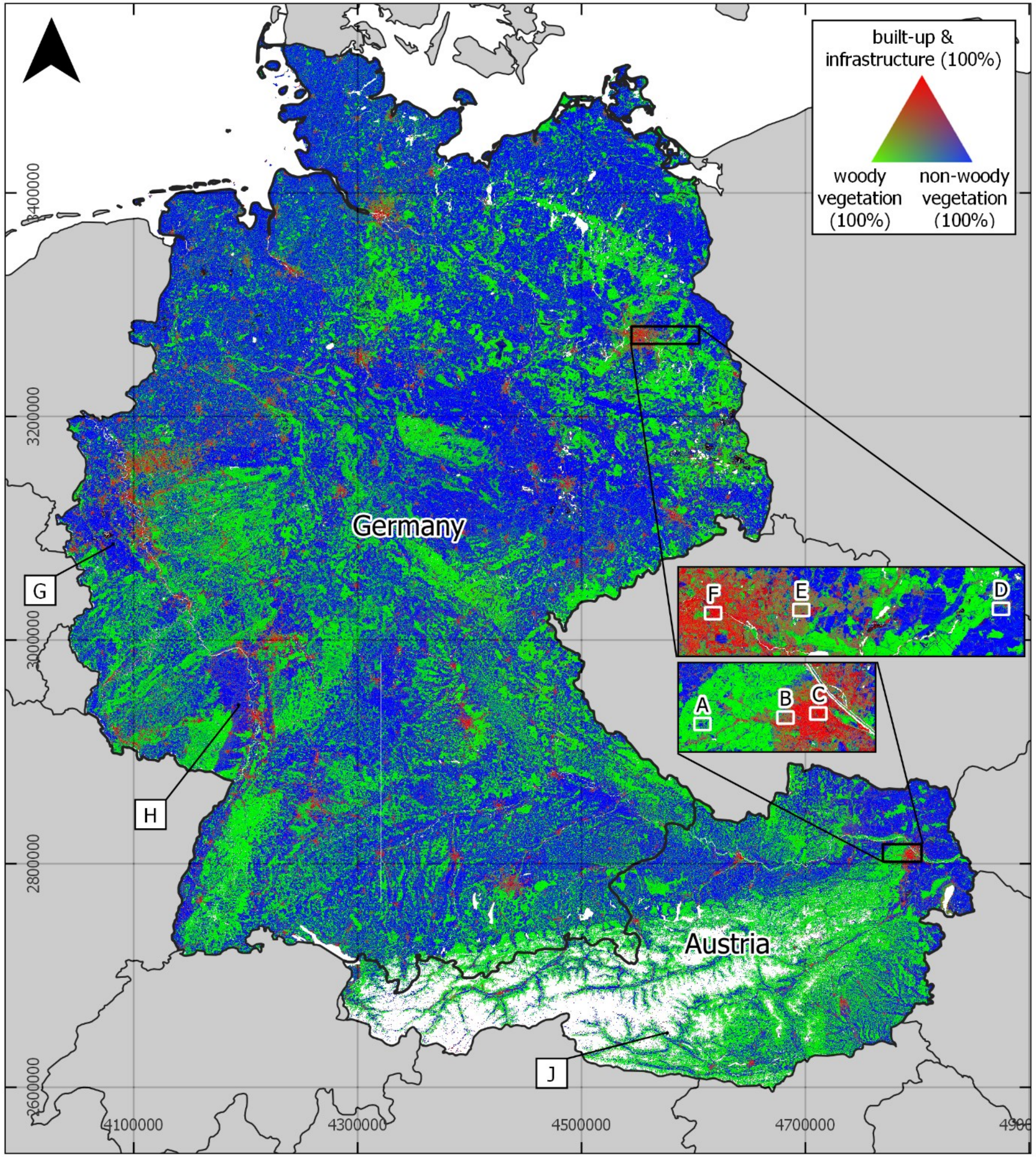
Results

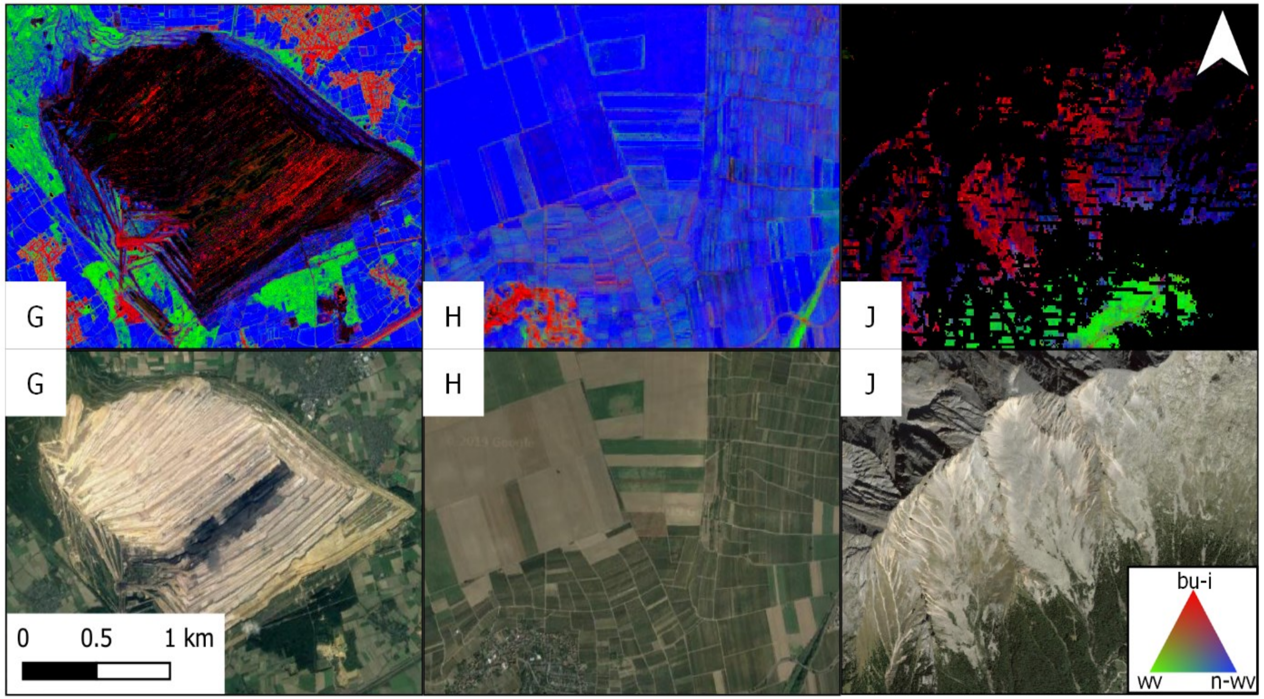
- Selected model quality:
  - Four classes
  - Three spatial resolutions
- Good model performance in three out of four classes




built-up & infrastructure (100%)
woody vegetation (100%)
non-woody vegetation (100%)
Germany
Austria
G
H
J
A
B
C
D
E
F
3400000
3200000
3000000
2800000
2600000
4100000
4300000
4500000
4700000

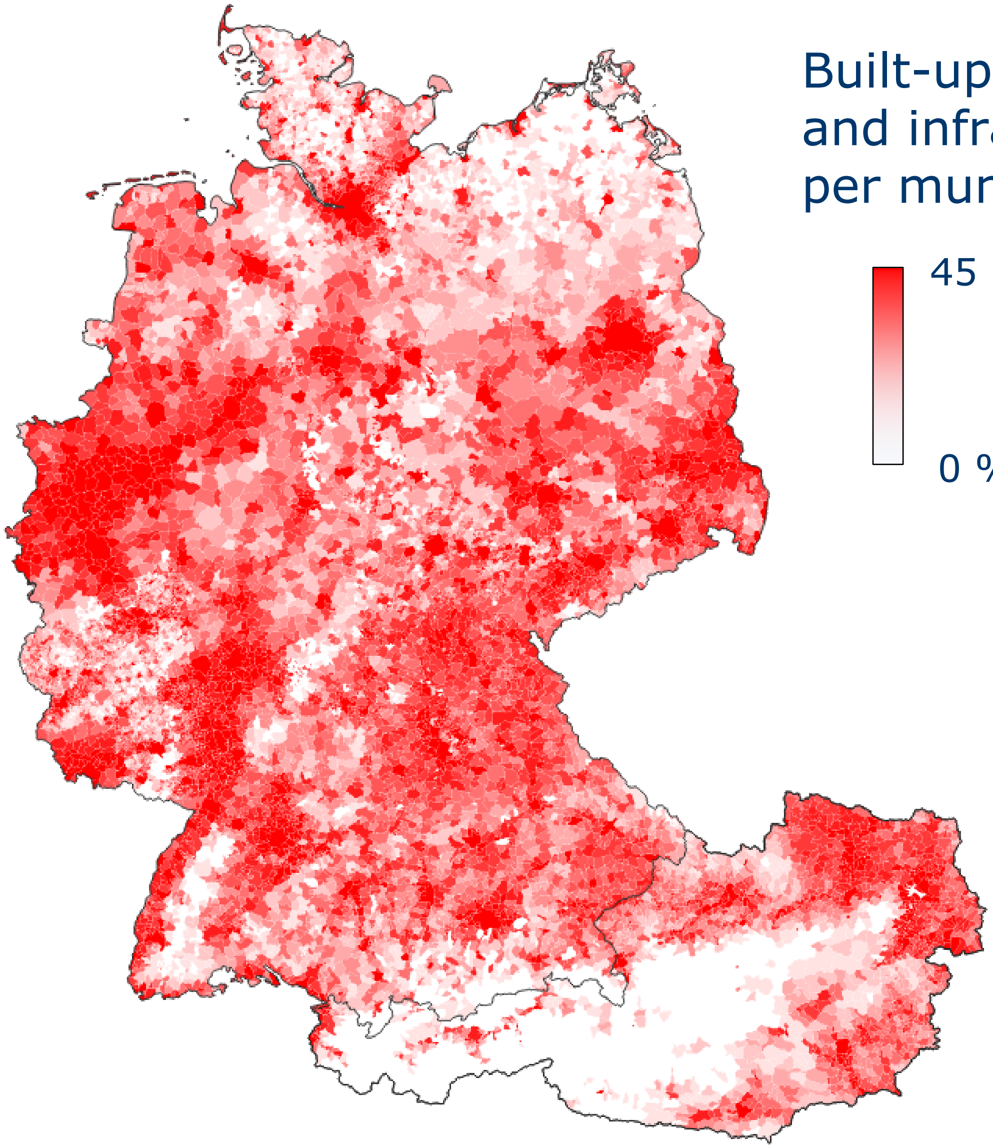

A
B
C
D
E
F
0 0.5 1 km
bu-i
wv
n-wv





- About 9.6 % of the overall area in Germany (6.7 % in Austria) is covered with built-up features and infrastructure.
- Copernicus Land Cover Product 2012, *artificial area*: Germany 9%, Austria 6%.

Results



Built-up surface and infrastructure per municipality

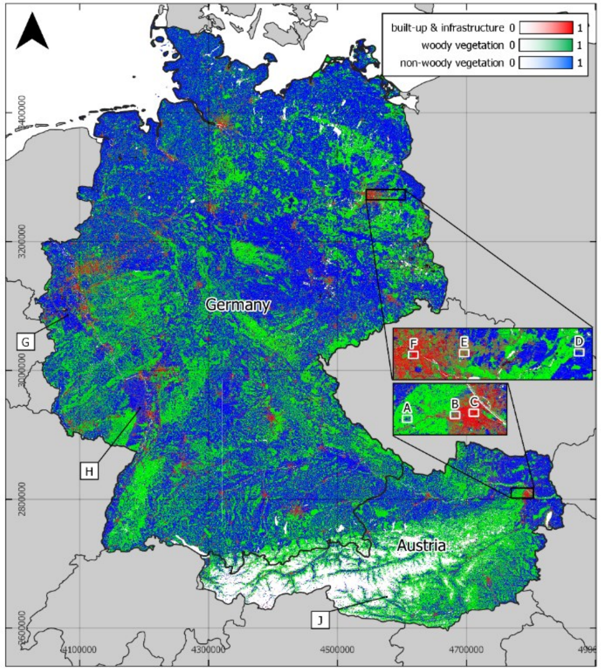
45 %

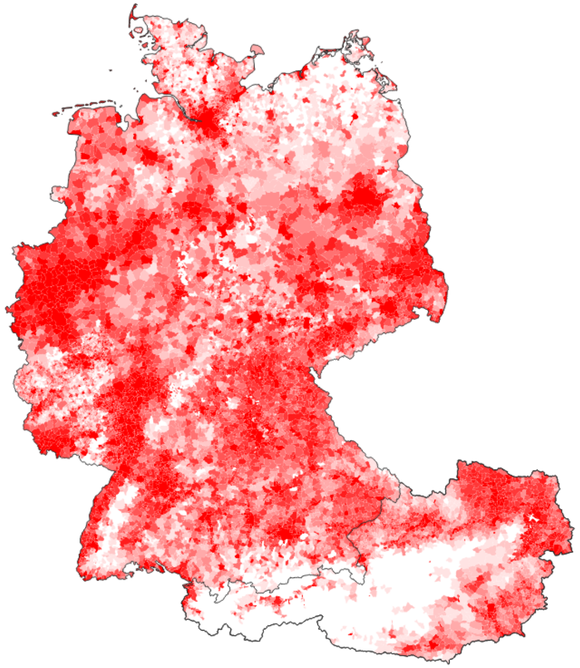
0 %

Municipality	Mean built-up in %
Ottobrunn b. München	48.29
Oberhausen	47.49
Weißenthurm/Rhein	46.76
Herne	46.52
Eichwalde	45.80
Ludwigshafen/Rhein	44.13
Nürnberg	43.47
Neulußheim (BW)	42.82
Gröbenzell b. München	41.83
München	41.25
Bischofsheim	41.18
Siershahn	40.88
Gelsenkichen	40.48
Mannheim	40.20
Asperg	40.12

Conclusion

- We can accurately map fractions of built-up surfaces and infrastructure, woody and non-woody vegetation with spectral-temporal metrics from Sentinel-2 time series
- Spectral-temporal metrics offer a very robust input to regression-based unmixing. No feature combination particularly stands out.
- We use a sub-pixel information, multi-class, high resolution, large area approach!
- Challenges exist in areas where surfaces appear like those used for built-up features
- Further research will be required on regional model transferability and library composition





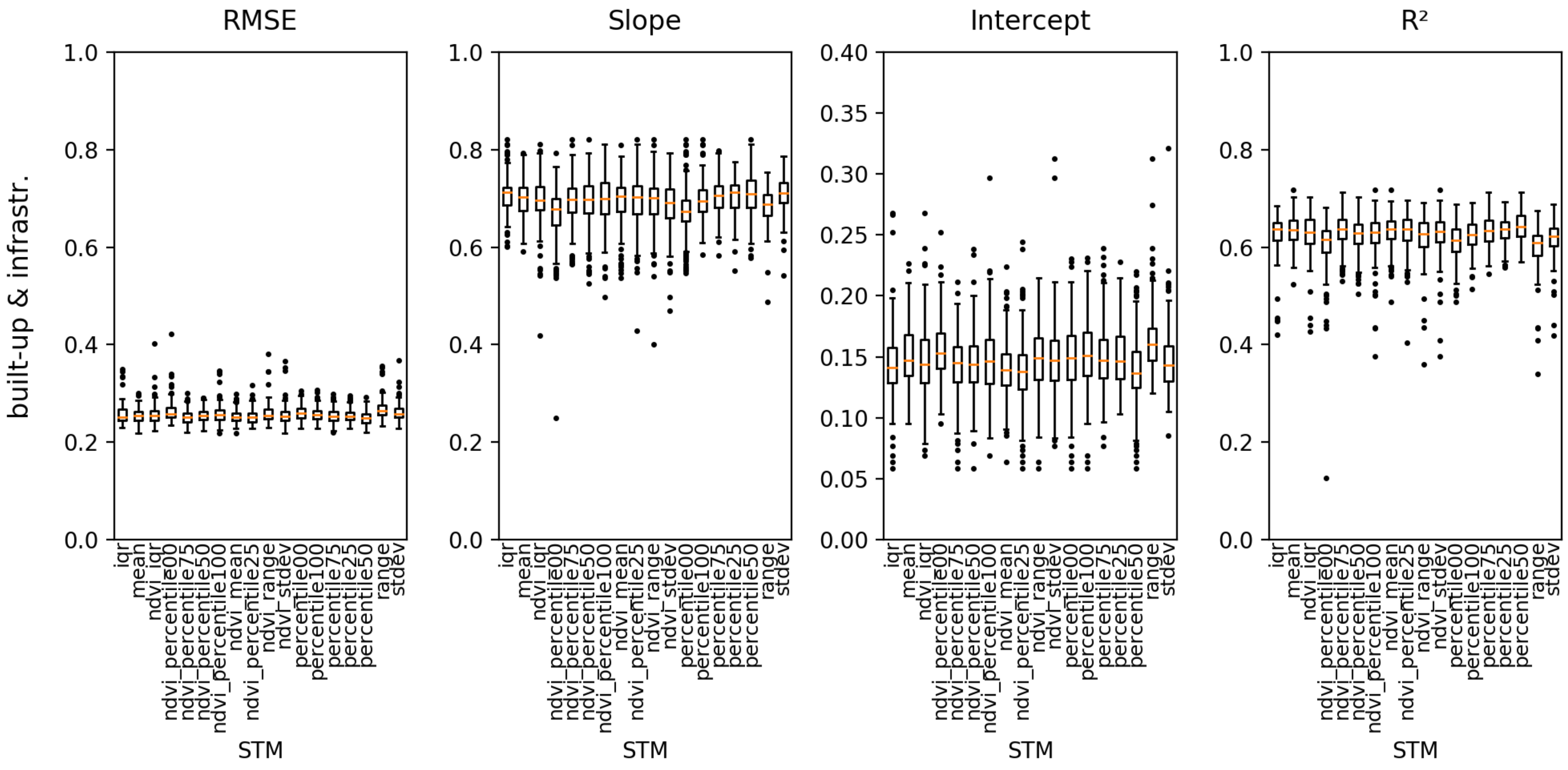
Thank you

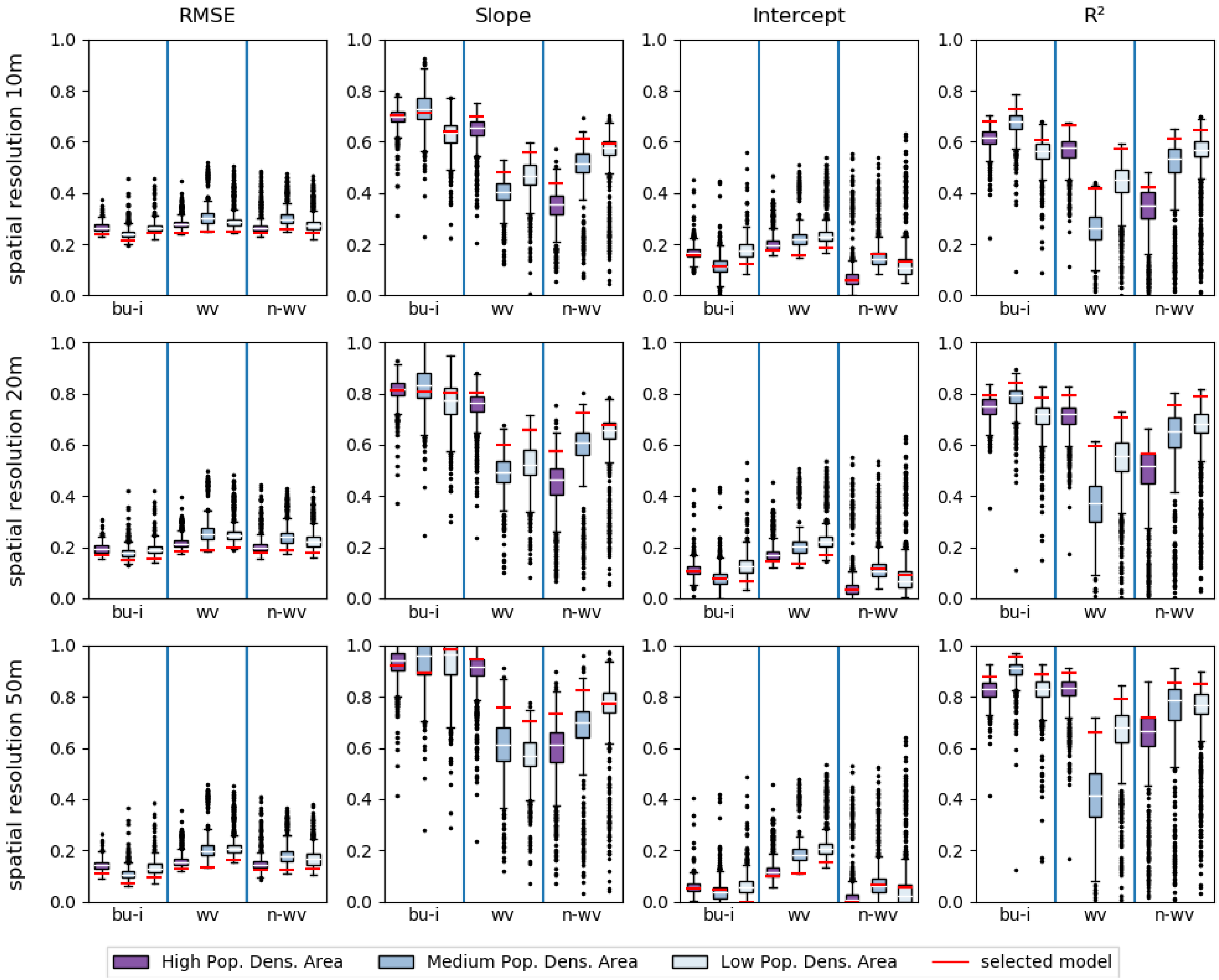
Franz Schug

Nation-wide mapping of fractional land cover with regression-based unmixing and Sentinel-2 imagery

European Research Council
Established by the European Commission

"This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 741950)".


RMSE
Slope
Intercept
R²
built-up & infrastr.
STM
iqr
mean
ndvi_iqr
ndvi_percentile00
ndvi_percentile75
ndvi_percentile50
ndvi_percentile100
ndvi_mean
ndvi_percentile25
ndvi_range
ndvi_stdev
percentile00
percentile100
percentile75
percentile25
percentile50
range
stdev


RMSE
Slope
Intercept
R²
spatial resolution 10m
spatial resolution 20m
spatial resolution 50m
1.0
0.8
0.6
0.4
0.2
0.0
bu-i
wv
n-wv
High Pop. Dens. Area
Medium Pop. Dens. Area
Low Pop. Dens. Area
selected model