# S2GLC – Sentinel-2 Global Land Cover
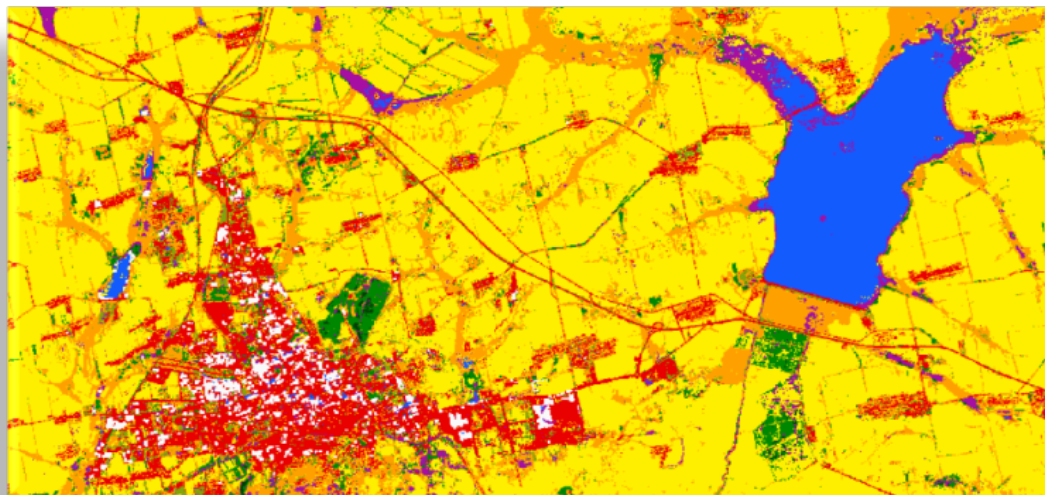
## Stanisław Lewiński, Ewa Gromny

*stlewinski@cbk.waw.pl, egromny@cbk.waw.pl*

## Earth Obserwation Department
## Space Research Centre of the Polish Academy of Sciences

# S2GLC – Sentinel-2 Global Land Cover
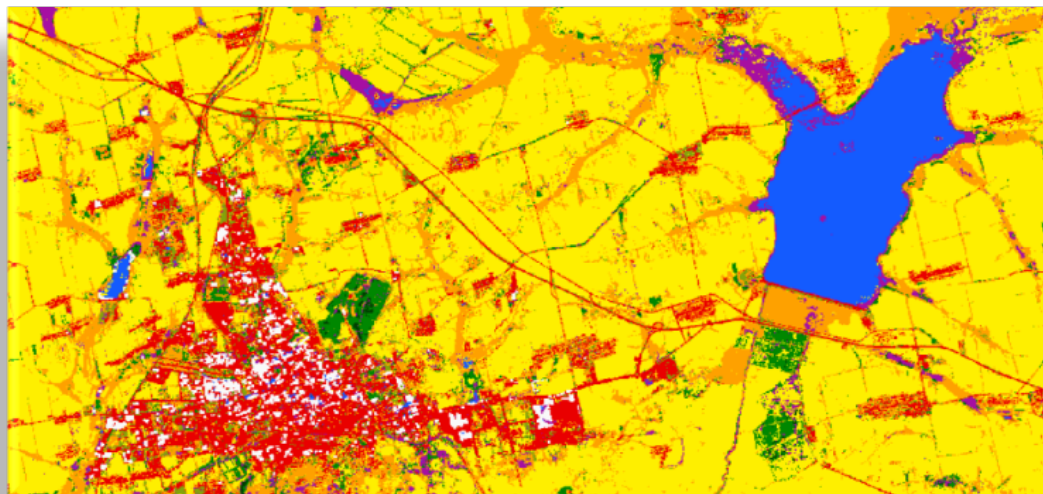## ESA project, processing Sentinel-2 for global mapping
### Scientific Exploitation of Operational Missions (SEOM)
### S2-4Sci Land and Water - Study 3: Classification

**CBK PAN** (Space Research Centre of the Polish Academy of Sciences), Poland
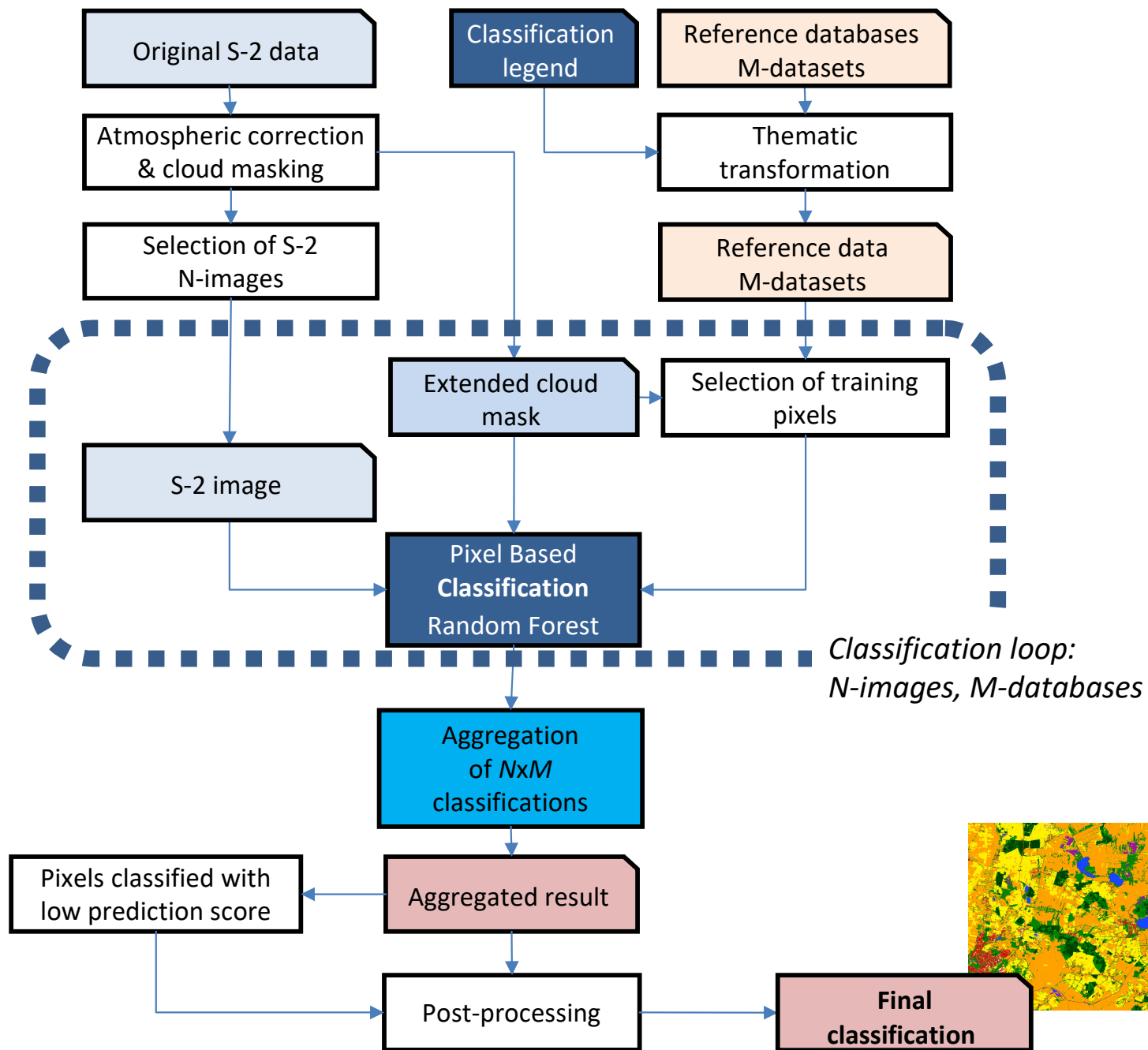**IABG** (Industrieanlagen Betriebsgesellschaft mbH), Germany
**EOXPLORE**, Germany
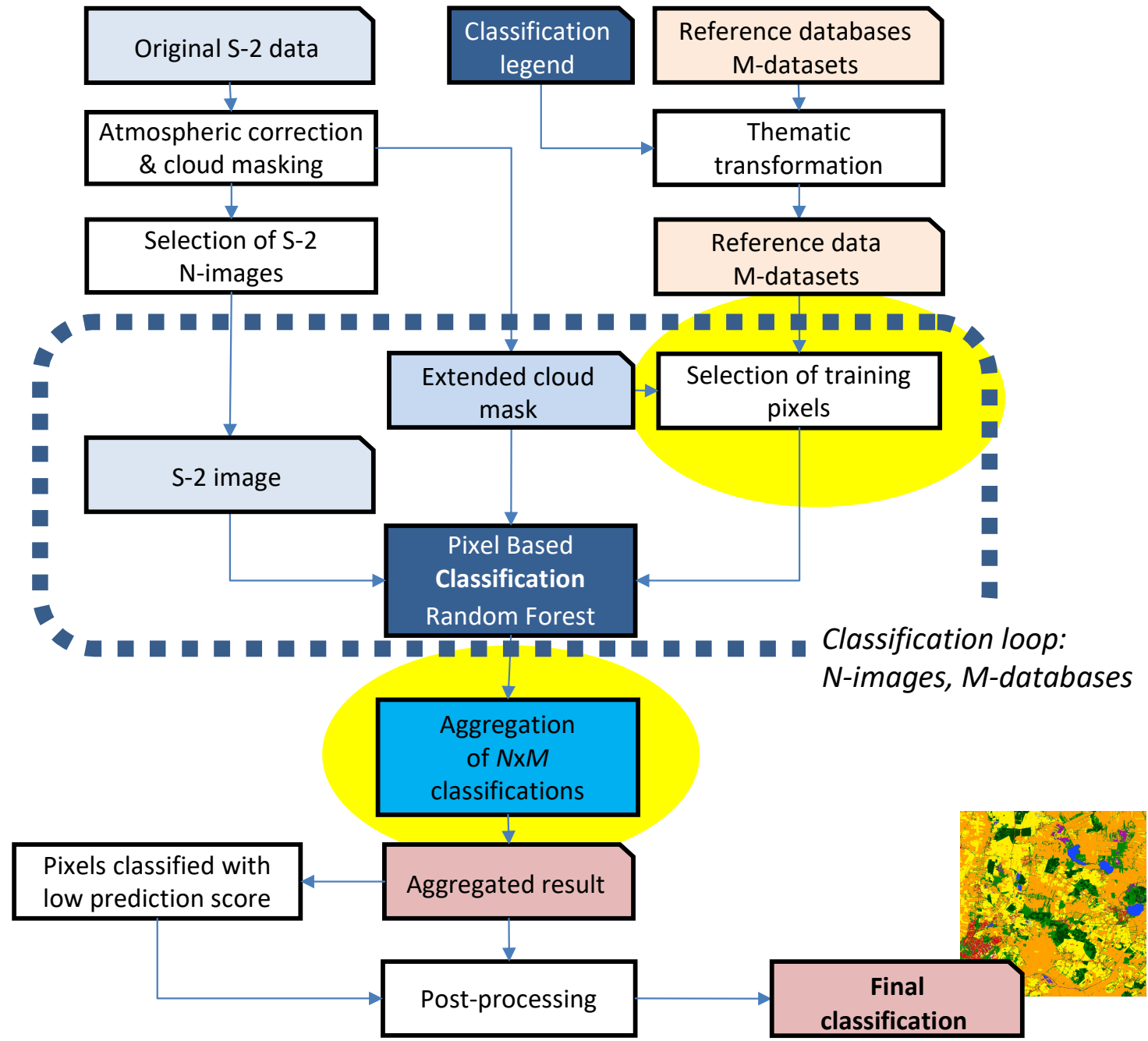**Friedrich-Schiller-Universität Jena**, Germany

*Stanisław Lewinski, Artur Nowakowski, Marcin Rybicki, Ewa Gromny,*
*Radosław Malinowski, Michał Krupiński, Elke Krätzschmar, Peter Hofmann,*
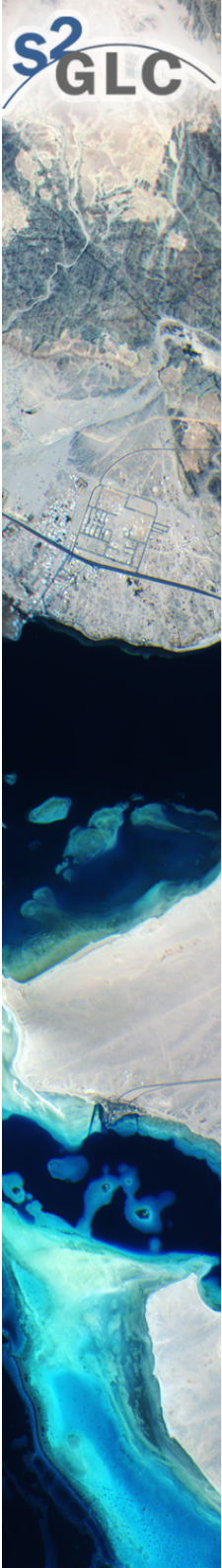*Conrad Bielski, Denise Dejon*

# S2GLC – Sentinel-2 Global Land Cover

Development of classification methodology and strategy
for global mapping using Sentinel-2 images

# S2GLC classification legend
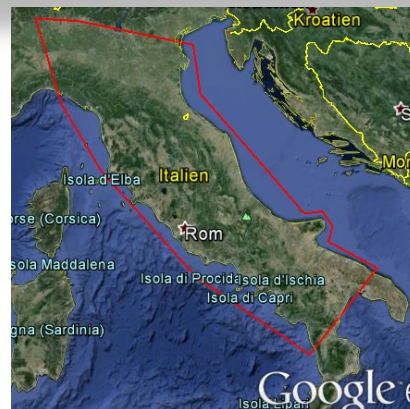
| S2GLC Legend | | |
|---|---|---|
| **Level 1** | **Level 2** | **Level 3** |
| **1.** Non-Vegetated surfaces | **1. 1.** Artificial surfaces and constructions | **1. 1. 1.** Artificial surfaces and constructions |
| | **1. 2.** Natural material surfaces | **1. 2. 1.** Consolidated areas |
| | | **1. 2. 2.** Un-Consolidated areas |
| **2.** Vegetated surfaces | **2. 1.** Woody vegetation | **2. 1. 1.** Evergreen broadleaf tree cover |
| | | **2. 1. 2.** Evergreen coniferous tree cover |
| | | **2. 1. 3.** Deciduous broadleaf tree cover |
| | | **2. 1. 4.** Deciduous coniferous tree cover |
| | | **2. 1. 6.** Bush, shrub |
| | **2. 2.** Low vegetation | **2. 2. 1.** Grass, herbaceous vegetation |
| | | **2. 2. 3.** Burnt areas |
| **3.** Cultivated and managed areas | **3. 1.** Cultivated and managed areas | **3. 1. 1.** Cultivated and managed areas |
| **4.** Inundated vegetation | **4. 1.** Inundated vegetation | **4. 1. 1.** Inundated vegetation |
| **5.** Water bodies | **5. 1.** Water bodies | **5. 1. 1.** Water bodies |
| **6.** Permanent snow covered surfaces | **6. 1.** Permanent snow covered surfaces | **6. 1. 1.** Permanent snow covered surfaces |
| **7.** Unclassified surfaces | **7. 1.** Surfaces permanently covered by clouds | **7. 1. 1.** Surfaces permanently covered by clouds |

# S2GLC classification workflow



Original S-2 data → Atmospheric correction & cloud masking → Selection of S-2 N-images

Classification legend → Thematic transformation

Reference databases M-datasets → Thematic transformation → Reference data M-datasets

Extended cloud mask → Selection of training pixels

S-2 image → Pixel Based **Classification** Random Forest

*Classification loop: N-images, M-databases*

Aggregation of *NxM* classifications → Aggregated result → Pixels classified with low prediction score → Post-processing → **Final classification**

# S2GLC classification workflow

Original S-2 data

Classification legend

Reference databases M-datasets

Atmospheric correction & cloud masking

Thematic transformation

Selection of S-2 N-images

Reference data M-datasets

Extended cloud mask

Selection of training pixels

S-2 image

Pixel Based **Classification** Random Forest

*Classification loop: N-images, M-databases*

Aggregation of *NxM* classifications

Pixels classified with low prediction score

Aggregated result

Post-processing

**Final classification**

# Selection of training pixels



Sentinel-2

Forest class

Existing LR database

Sentinel-2

# Aggregation rule for a single pixel classified at time series

| Time series 1, .., 7 | | Classification result | Prediction score |
|---|---|---|---|
| 1 | Classification 1 | Class 1 | 0.5 |
| 2 | Classification 2 | Class 2 | 0.6 |
| 3 | Classification 3 | Class 1 | 0.7 |
| 4 | Classification 4 | Cloud | |
| 5 | Classification 5 | Class 3 | 0.4 |
| 6 | Classification 6 | Class 3 | 0.8 |
| 7 | Classification 7 | Class 1 | 0.6 |

For Class 1  counter 1 = (0.5+0.7+0.6)/6 = 1.8/6 = **0.3**

For Class 2  counter 2 = 0.6/6= 0.1

For Class 3  counter 3 = (0.4+0.8)/6 = 1.2/6 = 0.2
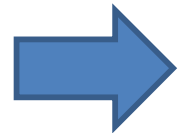
| Aggregated classification results | Class 1 |
|---|---|

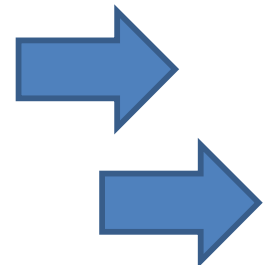# Aggregation of K-dates (time series) and N-databases
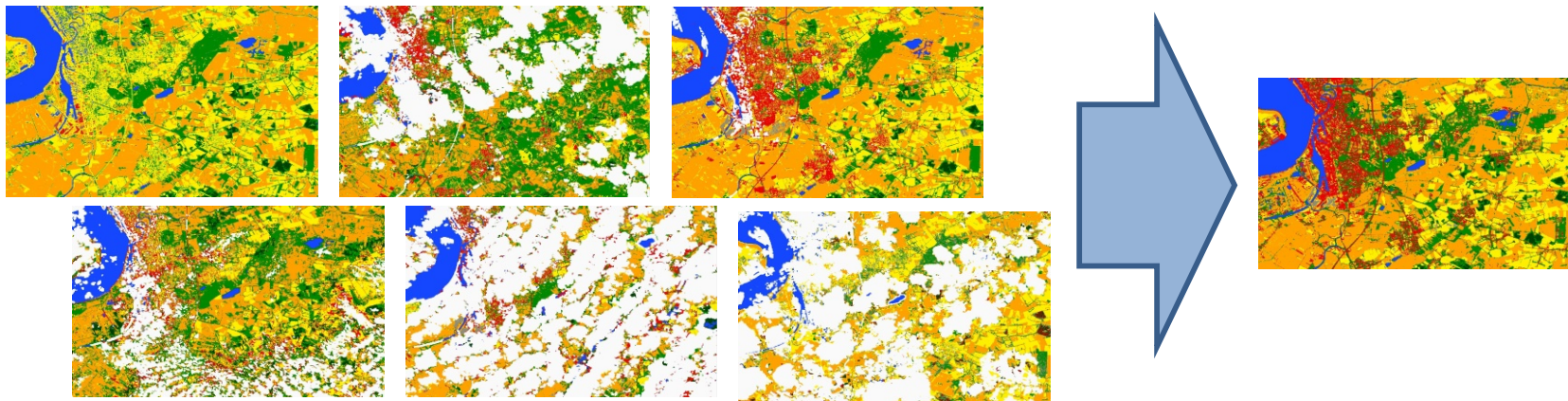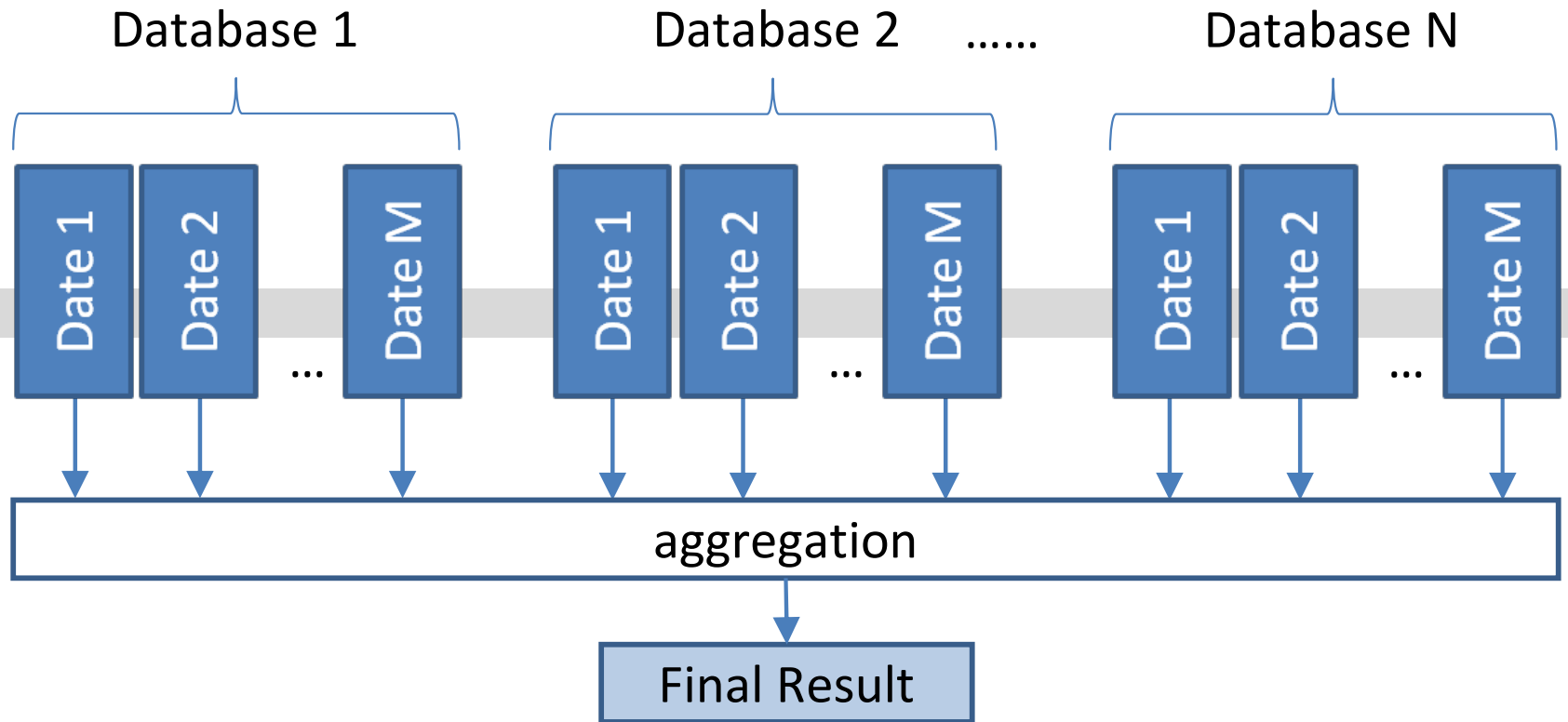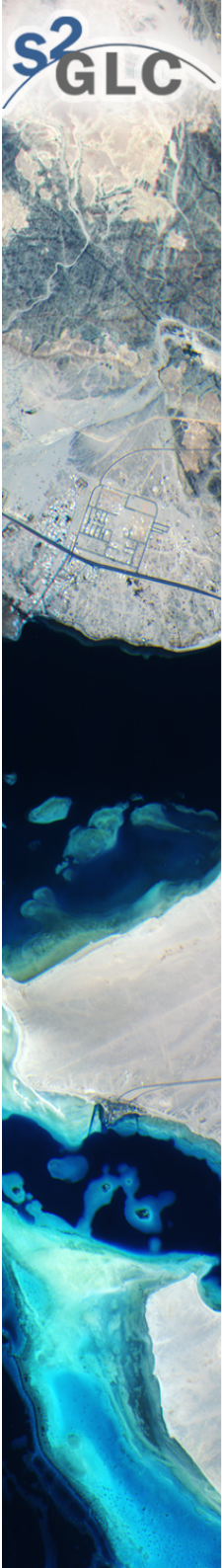
**Sentinel-2**, 10m,         Germany, test side 53°40´N, 8°45´E

# CCI Land Cover, 300m, database used for training,



**Legend:**
- Cropland, rainfed
- Herbaceous cover
- Mosaic cropland (>50%) / nat. veg. (<50%)
- Mosaic nat. veg. (>50%) / cropland (<50%)
- Tree cover, broadleaved (>15%)
- Tree cover, needleaved (>15%)
- Tree cover, mixed
- Mosaic tree, shrub (>50%) / herbaceous (<50%)
- Grassland
- Shrub or herbaceous cover
- Urban areas
- Water bodies
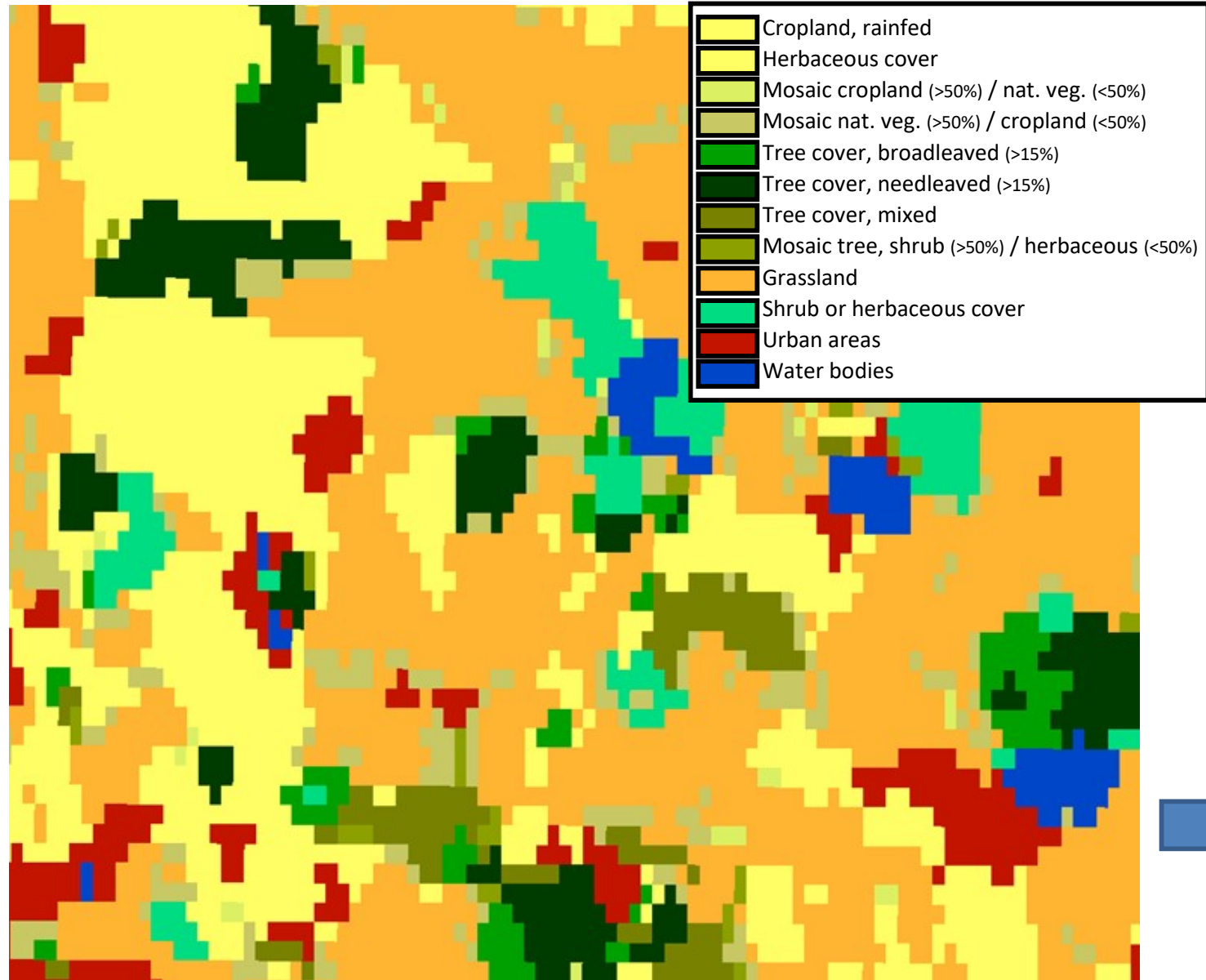
## S2GLC classification, 10m, aggregated result



Legend:
- Artificial surfaces
- Consolidated areas
- UnConsolidated areas
- Vegetated surfaces
- Evergreen coniferous tree cover
- Deciduous broasleaf tree cover
- Bush, shrub
- Grass, hearbaceous vegetation
- Cultivated and manager areas
- Inundated vegetation
- Water bodies
- Clouds

# Processed images and utilised LC databases
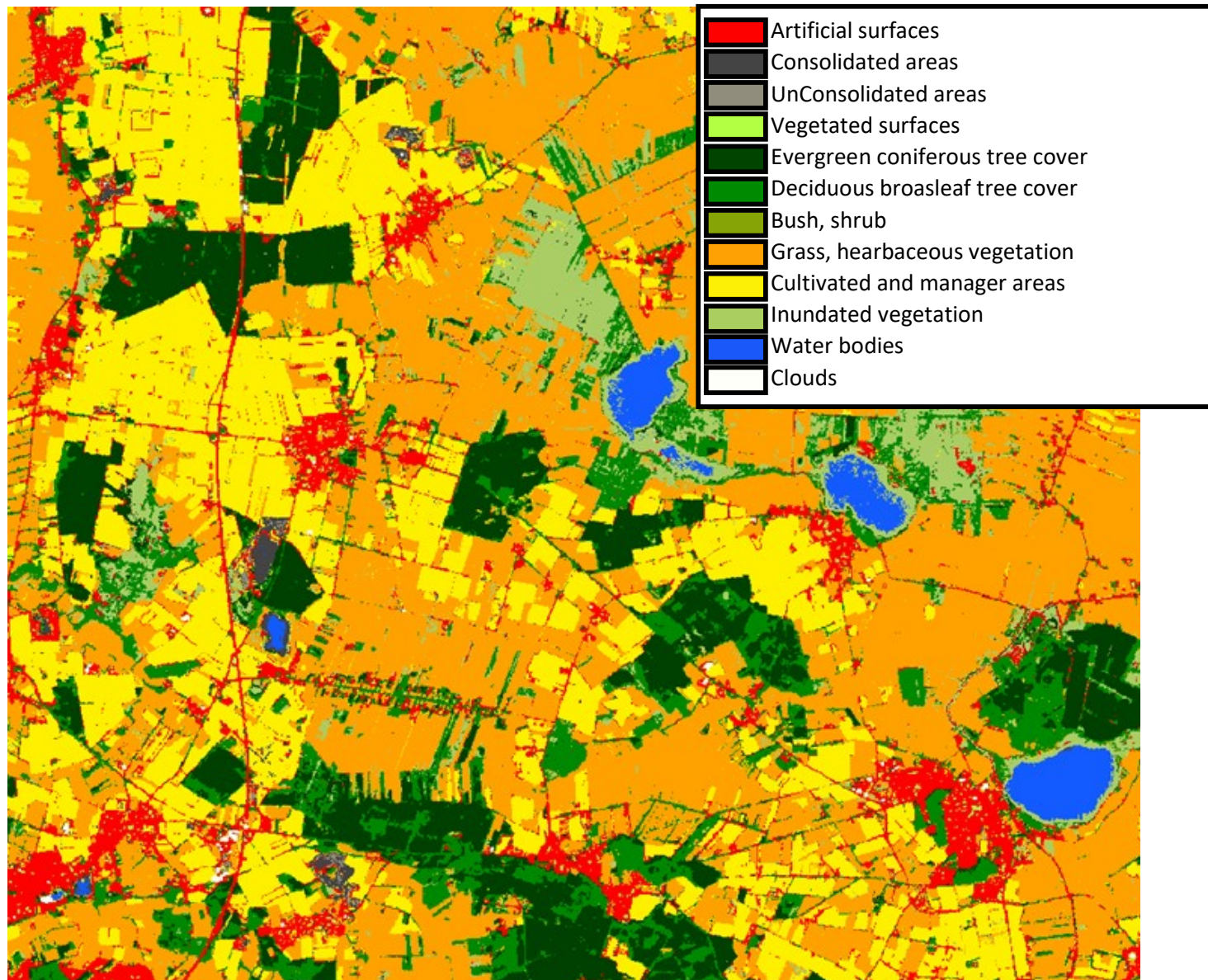
| Test site | Area km² | Sentinel-2 images | | Databases used for training | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| | | tiles | images per tile | global | regional | Auxiliary | |
| Germany | 357 375 | 56 | 10 | CCI Land Cover GlobeCover 2009 | CORINE Land Cover | GUF | NDWI |
| Italy | 301 230 | 63 | 10 | CCI Land Cover GlobeCover 2009 | CORINE Land Cover | GUF | NDWI |
| China | 200 750 | 31 | 6-10 | CCI Land Cover GlobeCover 2009 GlobeLand30 | | GUF | NDWI |
| Namibia | 235 345 | 32 | 10 | CCI Land Cover | LC for Africa | GUF | NDWI |
| Colombia | 211 705 | 30 | 3-10 | CCI Land Cover GlobeCover 2009 | Colombian LC | GUF | NDWI |
| Summary: | 1 306 405 | 212 | | | | | |

GUF - Global Urban Footprint (DLR)
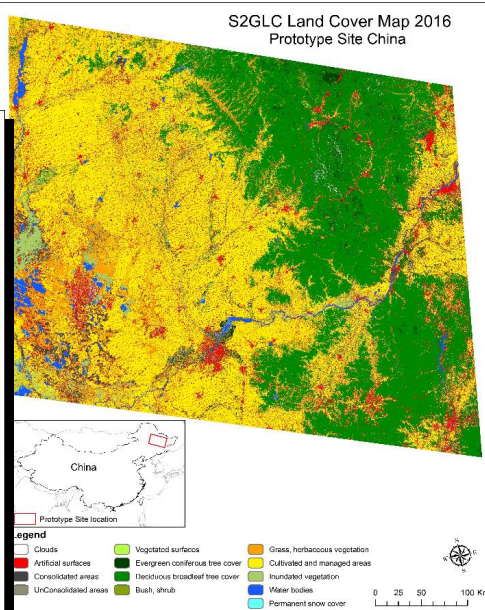NDWI - Normalized Difference Water Index

S2GLC Land Cover Map 2016
Prototype Site Germany
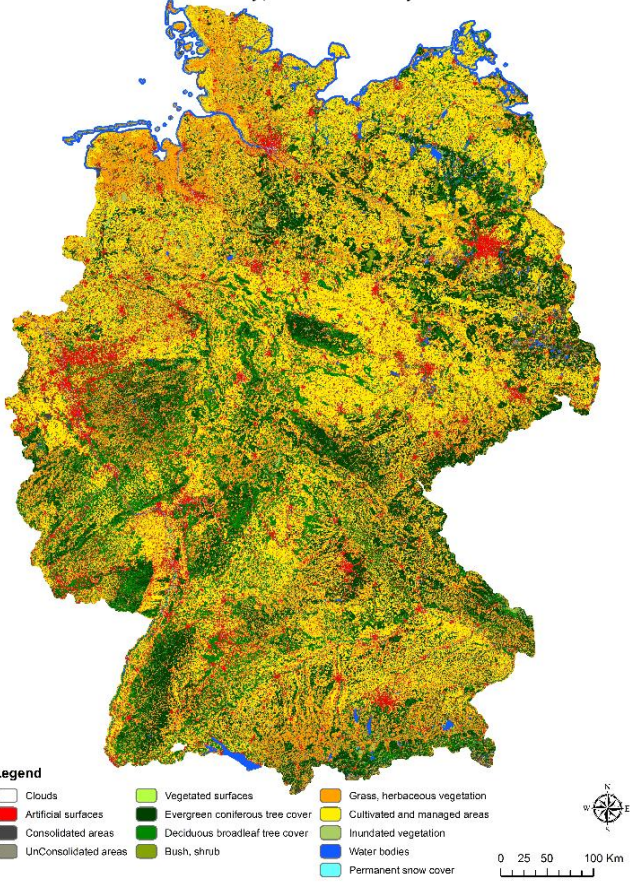
S2GLC Land Cover Map 2016
Prototype Site Italy

**Legend**
- Clouds
- Artificial surfaces
- Consolidated areas
- UnConsolidated areas
- Vegetated surfaces
- Evergreen coniferous tree cover
- Deciduous broadleaf tree cover
- Bush, shrub
- Grass, herbaceous vegetation
- Cultivated and managed areas
- Inundated vegetation
- Water bodies
- Permanent snow cover

S2GLC Land Cover Map 2016
Prototype Site Namibia

S2GLC Land Cover Map 2016
Prototype Site Colombia

S2GLC Land Cover Map 2016
Prototype Site China
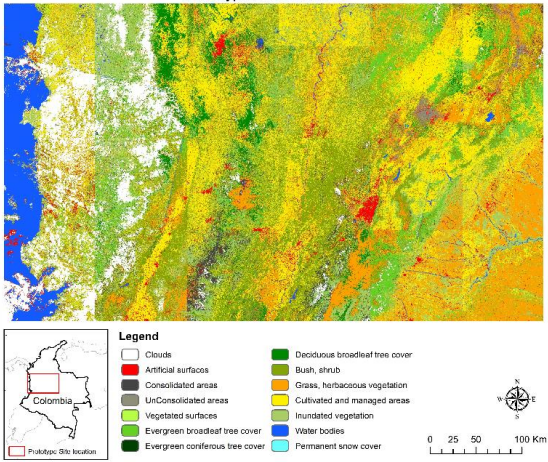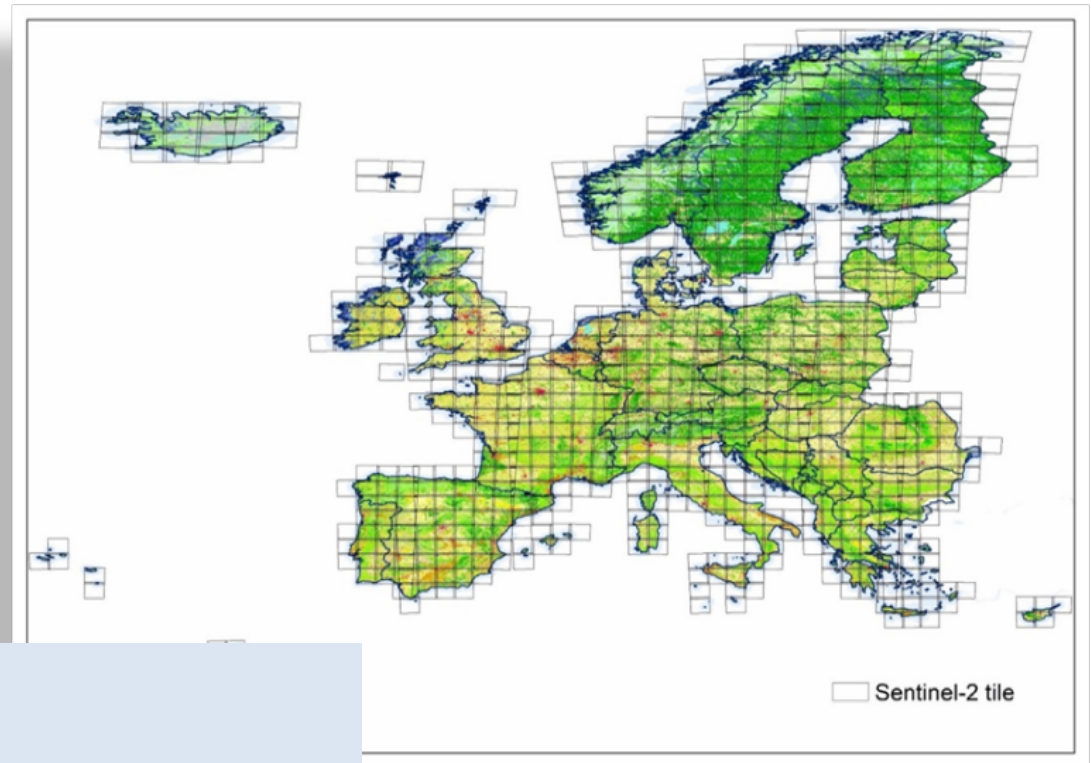
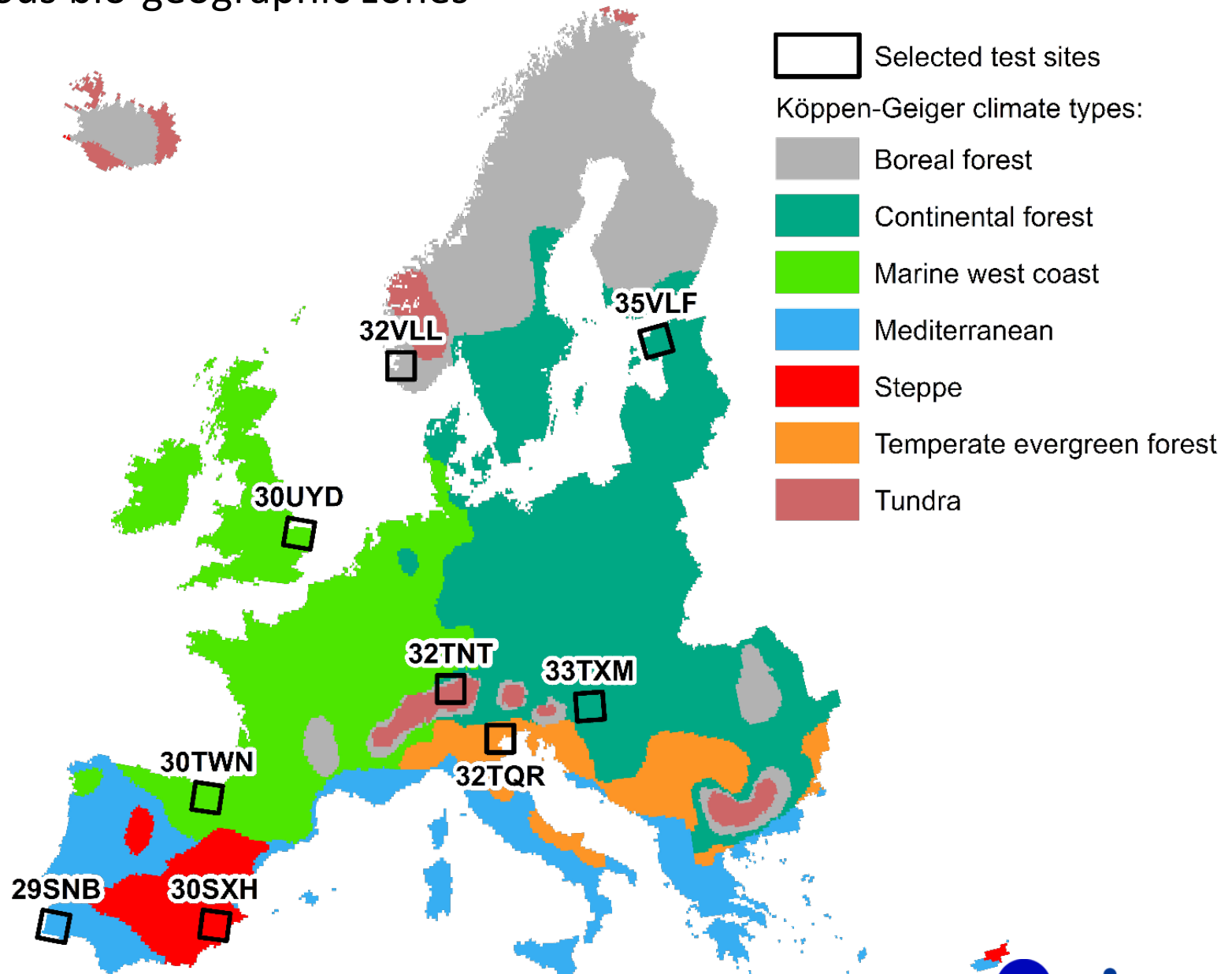| Test site | Overall accuracy | Kappa coefficient | *CORINE LC* (LUCAS points) |
|-----------|------------------|-------------------|----------------------------|
| Germany | **85.2** | 0.85 | *82.8* |
| Italy | **72.5** | 0.72 | *76.0* |
| China | 72.0 | 0.72 | |
| Namibia | 56.1 | 0.49 | |
| Colombia | 52.5 | 0.46 | |
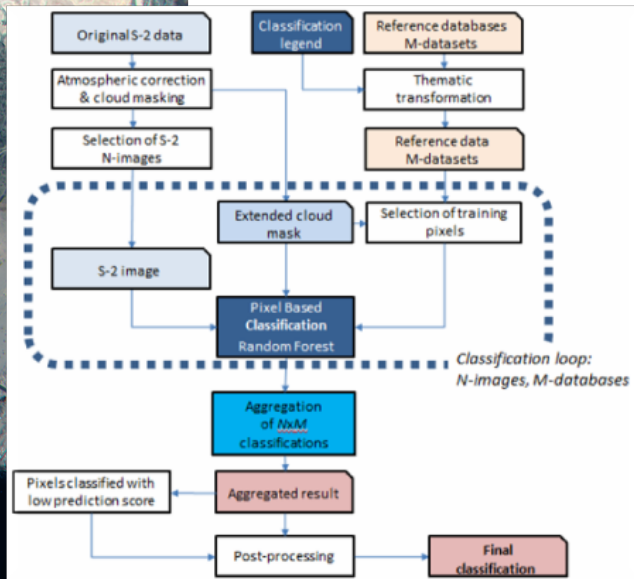
- ❑ application of the S2GLC approach for the classification of the whole of Europe
- ❑ The whole process performed on CreoDIAS using software developed by CBK PAN



Sentinel-2 tile

**40** Virtual machines  -  eo2.2xlarge,

8 processor cores, 32GB RAM, 128GB permanent memory

# Test sites for Europe classification

Classification modified based on nine test sites representing various bio-geographic zones

**Selected test sites**

**Köppen-Geiger climate types:**

- Boreal forest
- Continental forest
- Marine west coast
- Mediterranean
- Steppe
- Temperate evergreen forest
- Tundra

Test site labels: 32VLL, 35VLF, 30UYD, 32TNT, 33TXM, 30TWN, 32TQR, 29SNB, 30SXH

| Land Cover Class Name | Training Source | Filtration rules | | | | |
|---|---|---|---|---|---|---|
| | | NDWI | HRL Imp | HRL TCD | HRL GRA | NDVI |
| Artificial surfaces and constructions | HRL IMD >70% | ✓ | - | >10% | ✓ | |
| Cultivated | CLC | ✓ | >30% | >10% | ✓ | |
| Vineyards | CLC | ✓ | >30% | >10% | ✓ | |
| Herbaceous | HRL GRA | ✓ | >30% | >10% | | |
| Deciduous | HRL DLT ∩ CLC | ✓ | >30% | | ✓ | |
| Coniferous | HRL DLT ∩ CLC | ✓ | >30% | | ✓ | |
| Moors and Heathland | CLC | ✓ | >30% | >10% | | |
| Sclerophyllous vegetation | CLC | ✓ | >30% | >10% | ✓ | |
| Natural material surfaces | CLC | ✓ | >30% | >10% | ✓ | ✓ |
| Permanent snow, glaciers | CLC | ✓ | >30% | >10% | ✓ | |
| Marshes | CLC | ✓ | >30% | >10% | ✓ | |
| Peatbogs | CLC | ✓ | >30% | >10% | ✓ | |
| Water bodies | NDWI | - | - | - | - | - |

NDWI – Normalized Water Index,   HRL Imp – Impervious, HRL

TCD – Tree Cover Density,   HRL GRA – Grass, HRL

DLT – Dominant Leaf Type,   NDVI – Normalized Vegetation Index
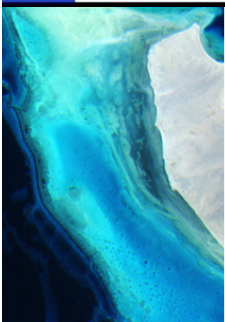
# Validation performer based on test sides
# Error Matrix

| | producer's accuracy | | user's accuracy | |
|---|---|---|---|---|
| | Total | % | Total | % |
| Artificial surfaces and constructions | 338 | **78.4** | 361 | **73.4** |
| Cultivated and managed areas | 2840 | **85.6** | 2649 | **91.7** |
| Vineyards | 272 | **75.4** | 415 | **49.4** |
| Herbaceous vegetation | 1749 | **76.8** | 1748 | **76.9** |
| Moors and Heathland | 136 | **66.2** | 308 | **29.2** |
| Broadleaf tree cover | 1521 | **86.1** | 1516 | **86.3** |
| Coniferous tree cover | 1129 | **86.0** | 1140 | **85.2** |
| Sclerophyllous vegetation | 632 | **48.3** | 371 | **82.2** |
| Natural material surfaces | 173 | **60.1** | 162 | **64.2** |
| Marshes | 120 | **74.2** | 217 | **41.0** |
| Peatbogs | 77 | **90.9** | 199 | **35.2** |
| Water bodies | 2244 | **95.1** | 2146 | **99.4** |

OA      82.9%

Kappa   0.798

(13 classes)

# Validation performer based on test sides
## Error Matrix

| | producer's accuracy | | | | user's accuracy | | | |
|---|---|---|---|---|---|---|---|---|
| | Total | % | | | Total | % | | |
| Artificial surfaces and constructions | 338 | **78.4** | | | 361 | **73.4** | | |
| Cultivated and managed areas | 2840 | **85.6** | 89.0 | | 2649 | **91.7** | 90.4 | |
| Vineyards | 272 | **75.4** | | | 415 | **49.4** | | |
| Herbaceous vegetation | 1749 | **76.8** | 81.4 | | 1748 | **76.9** | 74.6 | |
| Moors and Heathland | 136 | **66.2** | | | 308 | **29.2** | | |
| Broadleaf tree cover | 1521 | **86.1** | | | 1516 | **86.3** | | |
| Coniferous tree cover | 1129 | **86.0** | | | 1140 | **85.2** | | |
| Sclerophyllous vegetation | 632 | **48.3** | | | 371 | **82.2** | | |
| Natural material surfaces | 173 | **60.1** | | | 162 | **64.2** | | |
| Marshes | 120 | **74.2** | 94.4 | | 217 | **41.0** | 44.7 | |
| Peatbogs | 77 | **90.9** | | | 199 | **35.2** | | |
| Water bodies | 2244 | **95.1** | | | 2146 | **99.4** | | |

OA          82.9%                    85%

Kappa    0.798                    (10 classes)

(13 classes)

Legend
- Clouds
- Artificial surfaces and constructions
- Cultivated areas
- Vineyards
- Broadleaf tree cover
- Coniferous tree cover
- Herbaceous vegetation
- Moors and Heathland
- Sclerophyllous vegetation
- Marshes
- Peatbogs
- Natural material surfaces
- Permanent snow cover
- Water bodies

0    1000    2000    3000    4000 km

# Summary of Europe LC classification

- Classification legend defined based on existing databases.
    - Individual rules for training samples selection for each LC class.

- Classification approach characteristic:
    o Pixel-based
    o Classification features – spectral bands (10 m, 20 m) and spectral indices
    o Random Forest classifier
    o Final result calculated by aggregation (single classifications/prediction score)
- Full automation of all processing steps, processing of about 16000 images
- All processing steps deployed as dedicated software developed by CBK PAN.
- Overall accuracy for test sites: 83% (13 classes), 85% (10 classes)

Classification approach can be adjusted to other training data.

# Roma , Italy
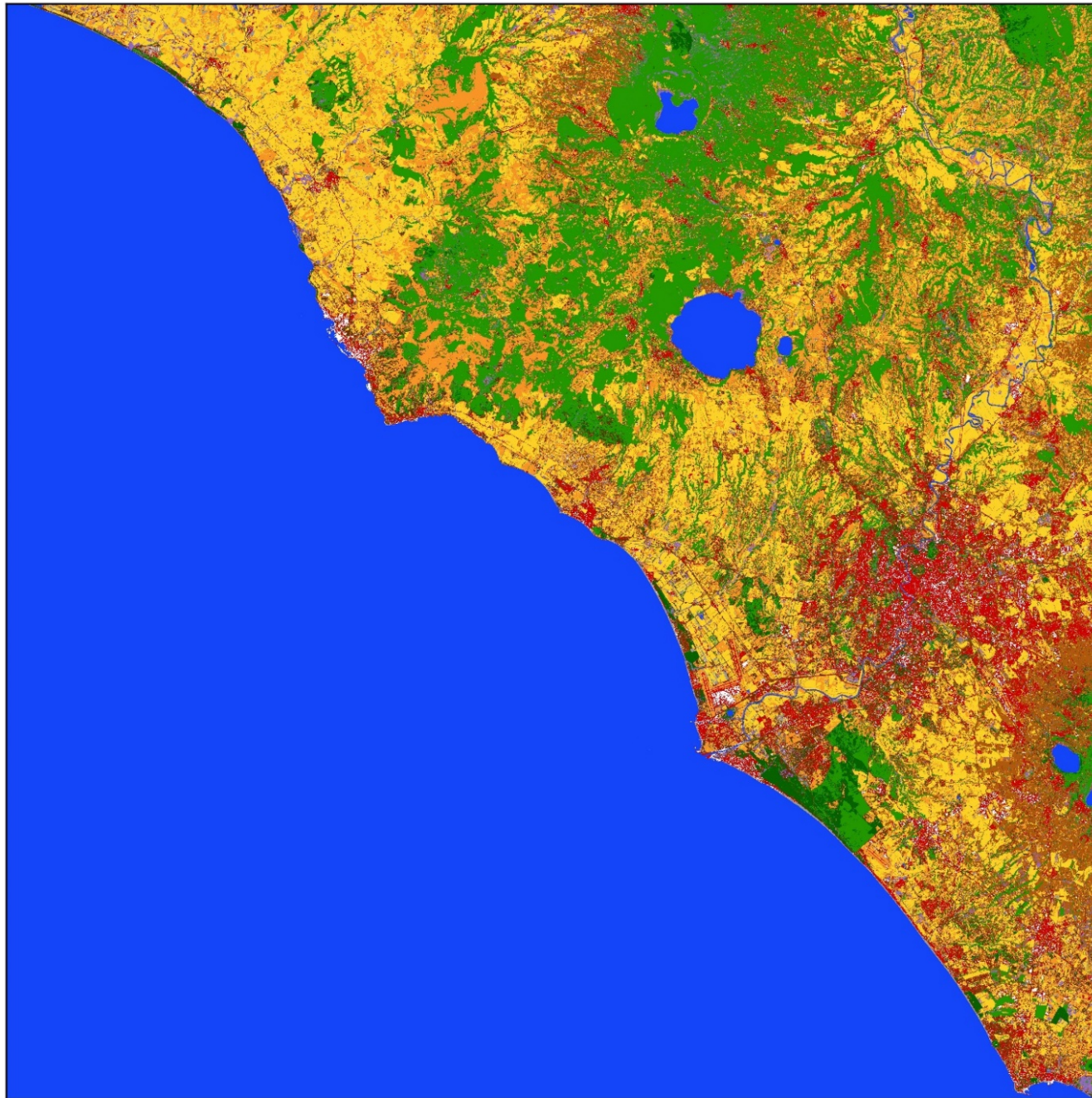


Tile 33TTG - Italy

Classification of Sentinel-2 data from year 2017

## Legend

| | |
|---|---|
| | Clouds |
| ■ | Artificial surfaces and constructions |
| ■ | Cultivated areas |
| ■ | Vineyards |
| ■ | Broadleaf tree cover |
| ■ | Coniferous tree cover |
| ■ | Herbaceous vegetation |
| ■ | Moors and Heathland |
| ■ | Sclerophyllous vegetation |
| ■ | Marshes |
| ■ | Peatbogs |
| ■ | Natural material surfaces |
| ■ | Permanent snow covered surfaces |
| ■ | Water bodies |

10   0   10   20 km

# S2GLC project is finished ........

Next steps:

❑ SCERIN – selection of training points   auto. vs manual selection.

❑ Europe2017 LC map will be served by CreoDIAS (?)

❑ Permanent LC service will be established by CreoDIAS (?)

❑ Validation – some tiles will be checked by voluntaries (RS friends).

❑ Land Cover Action Group of EuroGEOSS is using S2GLC map for calculation of SDG indicators.

❑ EC will use our map ?
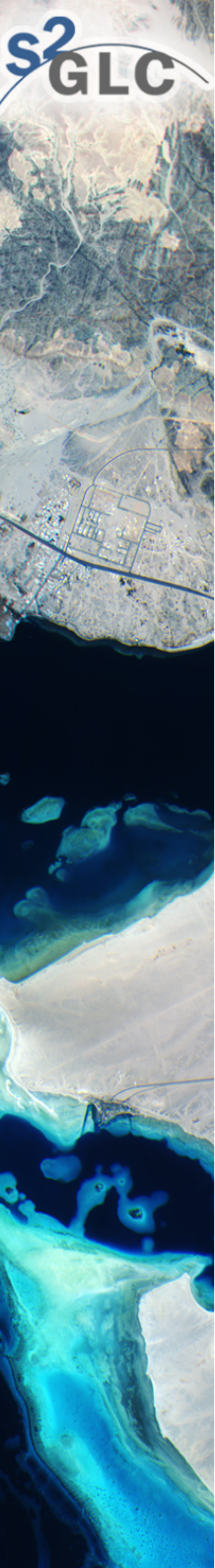
# S2GLC – Sentinel-2 Global Land Cover
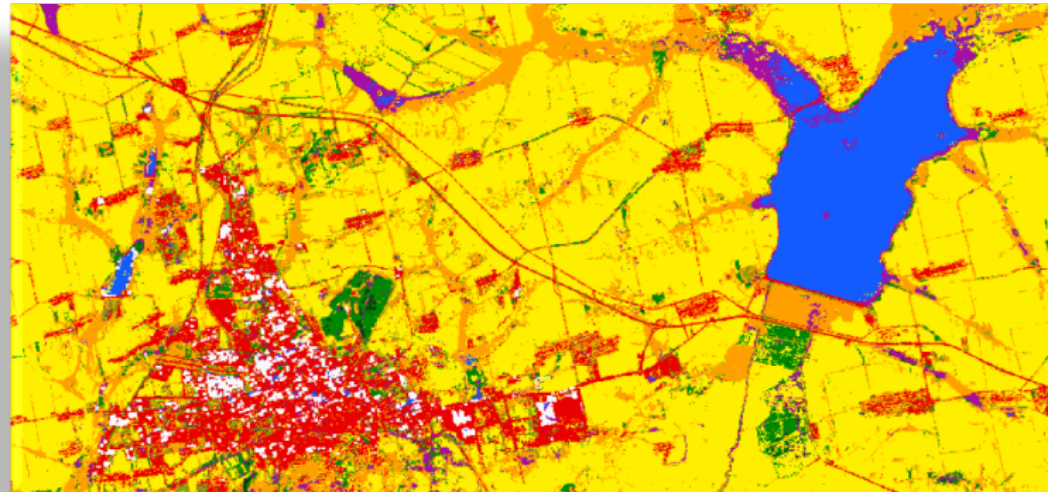
## Stanisław Lewiński, Ewa Gromny

*stlewinski@cbk.waw.pl, egromny@cbk.waw.pl*

## Earth Observation Department
## Space Research Centre of the Polish Academy of Sciences