ST_LUCAS: Easy Access System for Harmonized LUCAS Dataset

Tomáš Bouček, Martin Landa, Lukáš Brodský, Ondřej Pešek, Lena Halounová

Department of Geomatics

Faculty of Civil Engineering

Czech Technical University in Prague

Content

- LUCAS dataset introduction

- ST_LUCAS system

- Conclusions

LUCAS dataset introduction

- **Land Use and Coverage Area Frame Survey**
- Managed by Eurostat
- In-situ surveys every three years since 2006 (2022 in processing)
- Sampling density given by 2x2 km grid
- Over 1.3 million points examined

Year	EU countries	Points
2006	11	168 402
2009	23	234 623
2012	27	270 272
2015	28	339 696
2018	28	337 854
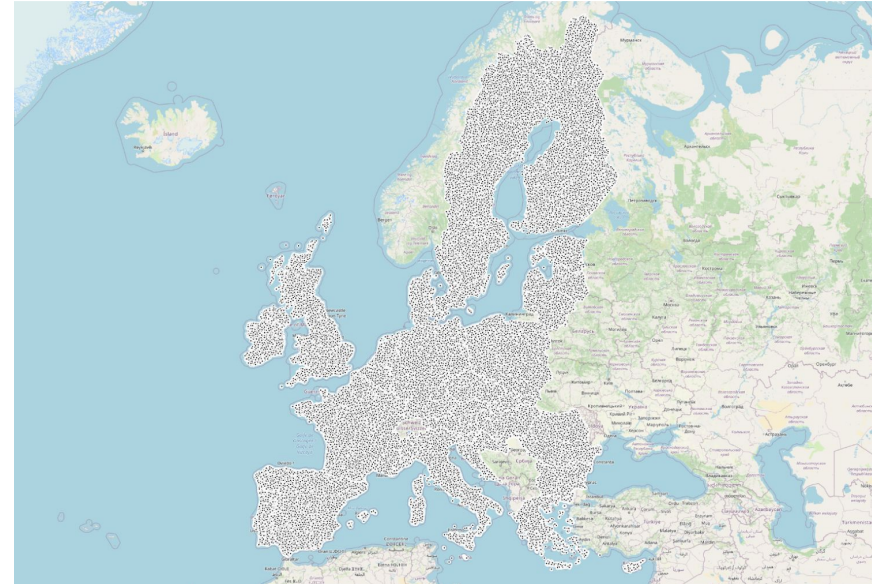
LUCAS dataset introduction

- Examine land cover (76 classes) and land use (41 classes)
- Structural elements in the landscape, agro-environmental information
- Collect photos (facing + 4 directions)
- Take a 500-gram topsoil sample at one out of 10 points
- In 2018, Copernicus, INSPIRE and EUNIS attributes were added

LUCAS dataset introduction
- Land Cover & Land Use

Land cover			
A00	ARTIFICIAL LAND	A10	Roofed built-up areas
		A20	Artificial non-built up areas
		A30	Other artificial areas
B00	CROPLAND	B10	Cereals
		B20	Root crops
		B30	Non-permanent industrial crops
		B40	Dry pulses, vegetables and flowers
		B50	Fodder crops
		B70	Permanent crops: fruit trees
		B80	Other permanent crops
C00	WOODLAND	C10	Broadleaved woodland
		C20	Coniferous woodland
		C30	Mixed woodland
D00	SHRUBLAND	D10	Shrubland with sparse tree cover
		D20	Shrubland without tree cover
E00	GRASSLAND	E10	Grassland with sparse tree/shrub cover
		E20	Grassland without tree/shrub cover
		E30	Spontaneously re-vegetated surfaces
F00	BARE LAND AND LICHENS/MOSS	F10	Rocks and stones
		F20	Sand
		F30	Lichens and moss
		F40	Other bare soil
G00	WATER AREAS	G10	Inland water bodies
		G20	Inland running water
		G30	Transitional water bodies
		G40	Sea and ocean
		G50	Glaciers, permanent snow
H00	WETLANDS	H10	Inland wetlands
		H20	Coastal wetlands

Land use			
U100	PRIMARY SECTOR	U110	Agriculture
		U120	Forestry
		U130	Aquaculture and fishing
		U140	Mining and quarrying
		U150	Other primary production
U 200	SECONDARY SECTOR	U210	Energy production
		U220	Industry and manufacturing
U300	TERTIARY SECTOR, TRANSPORT, UTILITIES & RESIDENTIAL	U310	Transport, communication networks, storage, protection works
		U320	Water and waste treatment
		U330	Construction
		U340	Commerce, financial, professional and information services
		U350	Community services
		U360	Recreation, leisure, sport
		U361	Residential
U400	UNUSED AND ABANDONED AREAS	U410	Abandoned areas
		U420	Semi-natural and natural areas not in use

LUCAS dataset introduction - user disadvantages

Year	Number of attributes
2006	20
2009	44
2012	46
2015	59
2018	97

- Attribute changes across the years
  - 5 removed
  - 77 added
  - 24 renamed
  - 30 affected by different coding - example: LC1
    - C21 (2006) - Other broadleaved tree land
    - C21 (2018) - Spruce dominated coniferous woodland

  → **Harmonization** needed to work effectively across years

- Official distribution as plain CSV files
  - country by country or entire dataset

  → Preprocessing needed for GIS analysis, etc.

ST_LUCAS system

- Provides harmonized space-time aggregated LUCAS dataset
- Fully automated harmonization process
  - Configurable & Extensible
- LUCAS data provided through OGC web service
- Python API for geospatial developers and scientists
- QGIS plugin for wider audience
- Analytical methods for nomenclature translation and class aggregation

ST_LUCAS
   - system architecture

Persistent data storage (P1, P2)

Automation of the harmonization process & space-time aggregation (A1)

Standardized (OGC) web service (A2)

Client Python API (C1) & QGIS plugin (C3)
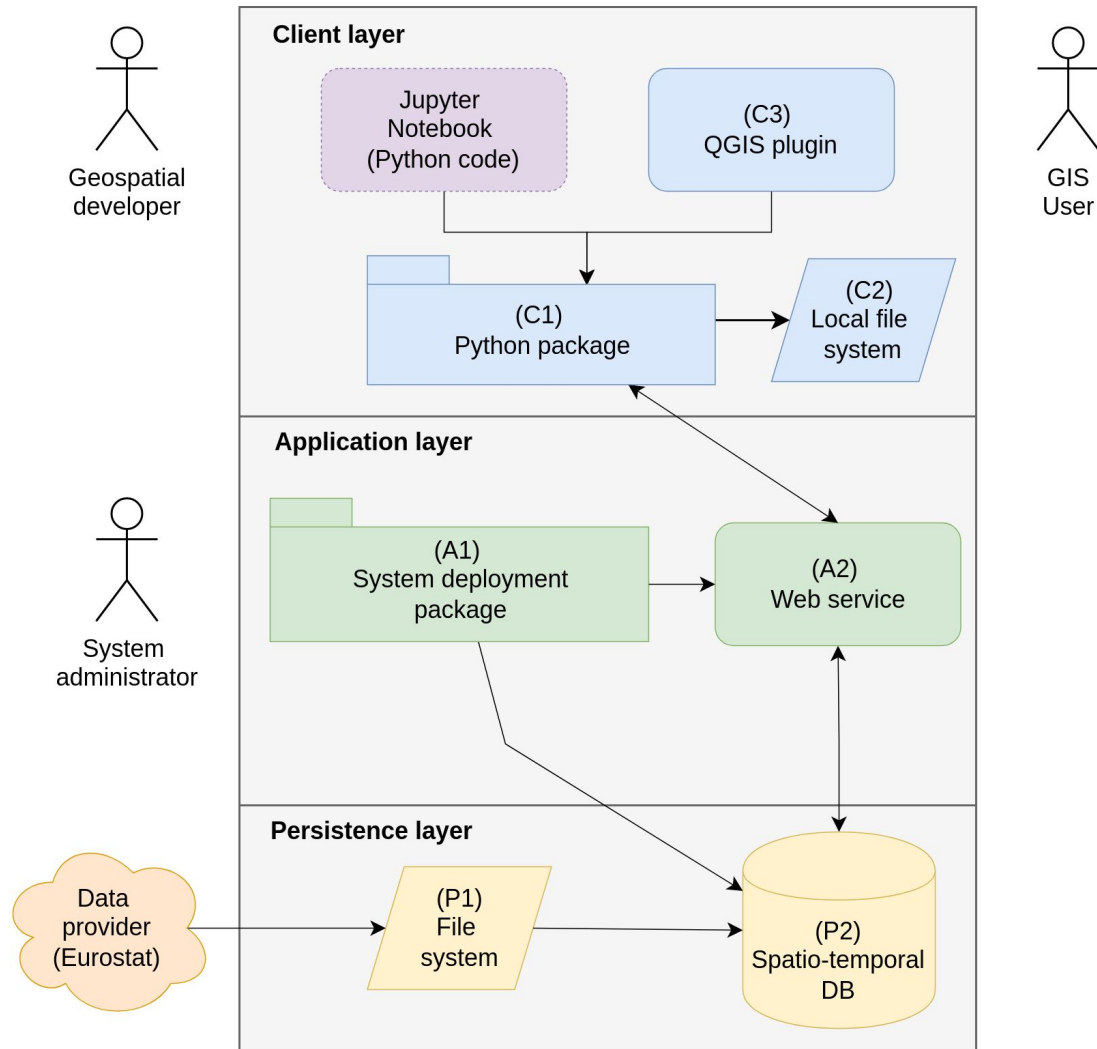
ST_LUCAS - request example

request = LucasRequest()

request.countries = ['RO']

request.years = [2015, 2018]

request.group = 'LC_LU'

lucasio = LucasIO()

lucasio.download(request)

print(lucasio.count())

Number of retrieved points: 33 445
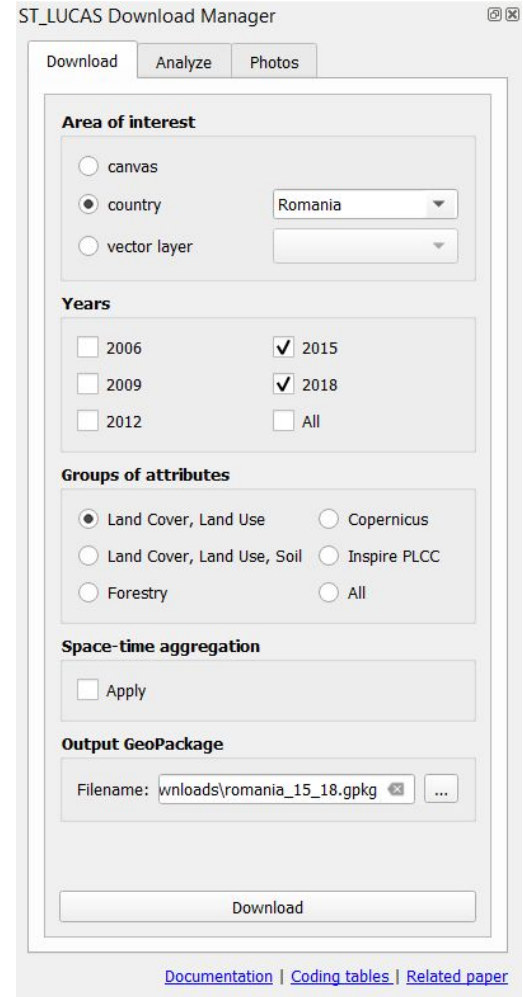
ST_LUCAS QGIS Plugin
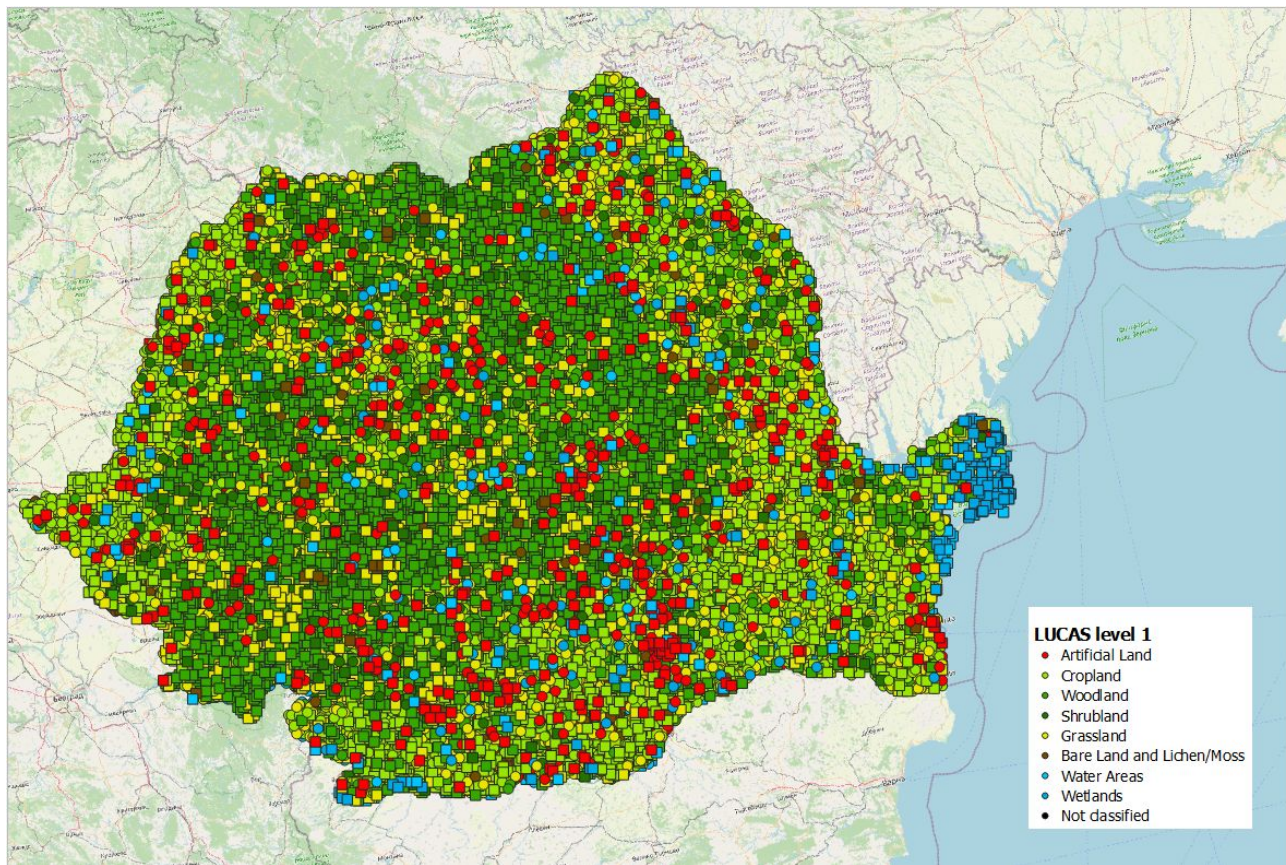
ST_LUCAS - space-time aggregation

request = LucasRequest()

request.countries = ['RO']

request.years = [2015, 2018]

request.group = 'LC_LU'
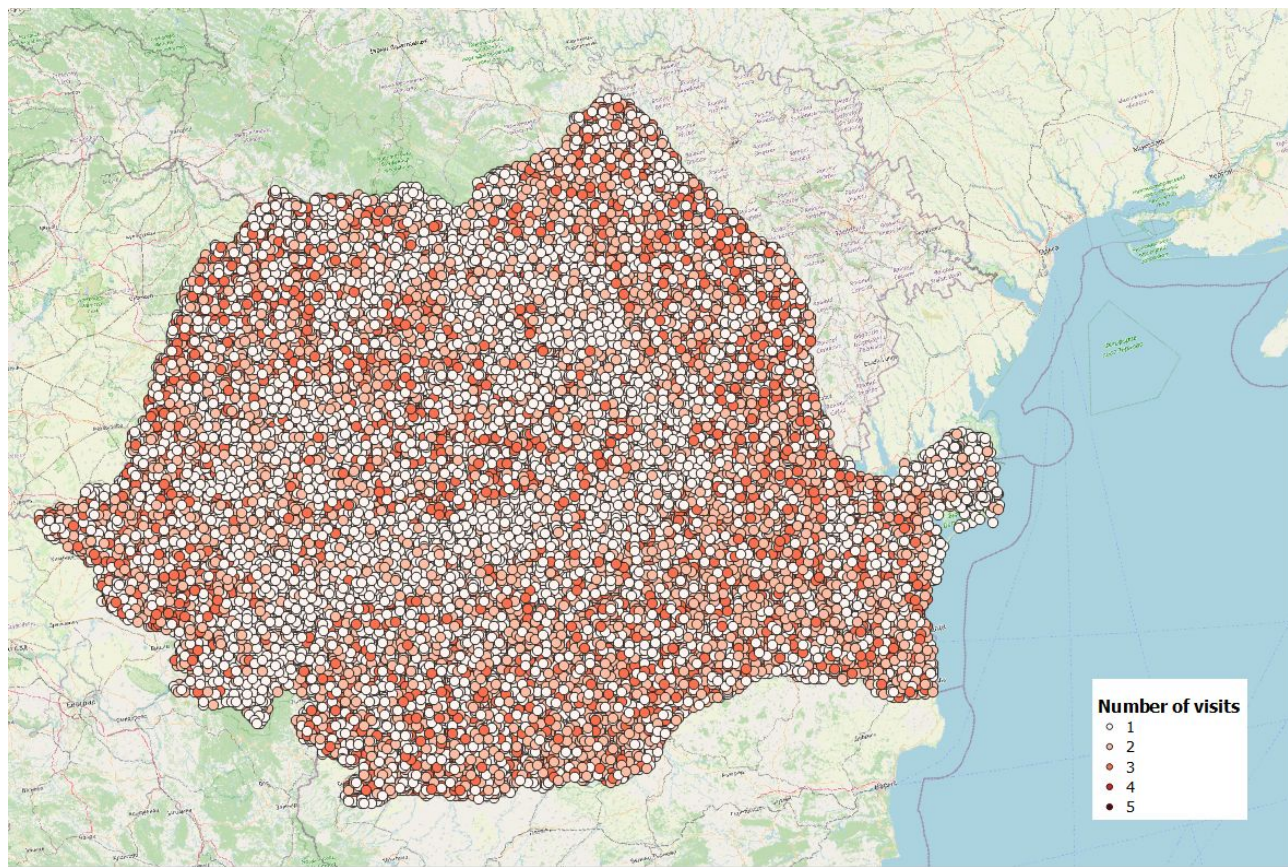
request.st_aggregated = True

lucasio = LucasIO()

lucasio.download(request)

print(lucasio.count())

Number of retrieved points: 28 061

ST_LUCAS - space-time aggregation

ST_LUCAS - showing LUCAS photos
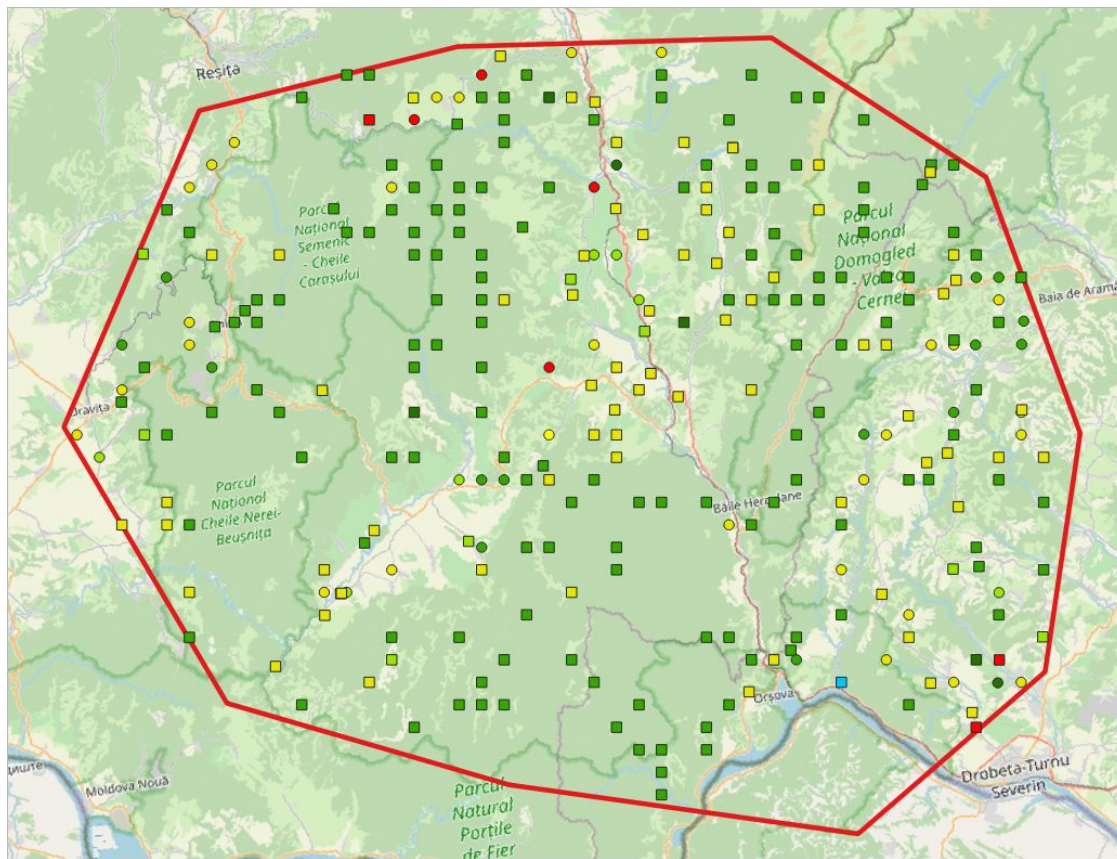
Mini use case - analysis of abandoned land
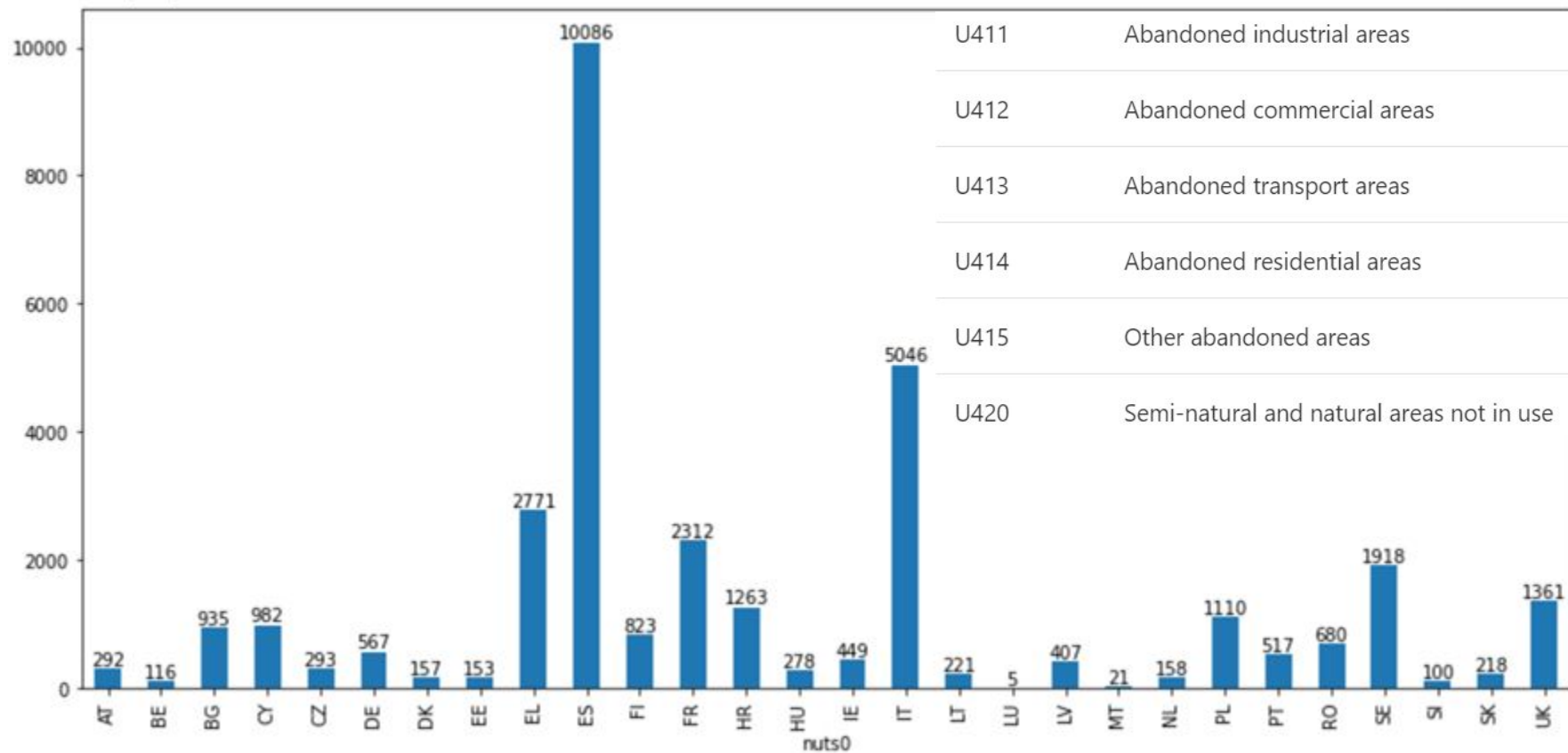
Related paper

Open Geospatial System for LUCAS In Situ Data Harmonization and Distribution

by ⊘ Martin Landa *,† ✉ ⓘ , ⊘ Lukáš Brodský † ✉ ⓘ , ⊘ Lena Halounová ✉ ⓘ , ⊘ Tomáš Bouček ✉ ⓘ and ⊘ Ondřej Pešek ✉ ⓘ

Department of Geomatics, Faculty of Civil Engineering, Czech Technical University in Prague, 166 29 Prague, Czech Republic

* Author to whom correspondence should be addressed.

† These authors contributed equally to this work.

Download   **Browse Figures**   **Versions Notes**

Abstract

The use of in situ references in Earth observation monitoring is a fundamental need. LUCAS (Land Use and Coverage Area frame Survey) is an activity that has performed repeated in situ surveys over Europe every three years since 2006. The dataset is unique in many aspects; however it is currently not available through a standardized interface, machine-to-machine. Moreover, the evolution of the surveys limits the performance of change analysis using the dataset. Our objective was to develop an open-source system to fill these gaps. This paper presents a developed system solution for the LUCAS in situ data harmonization and distribution. We have designed a multi-layer client-server system that may be integrated into end-to-end workflows. It provides data through an OGC (Open Geospatial Consortium) compliant interface. Moreover, a geospatial user may integrate the data through a Python API (Application Programming Interface) to ease the use in workflows with spatial, temporal, attribute, and thematic filters. Furthermore, we have implemented a QGIS plugin to retrieve the spatial and temporal subsets of the data interactively. In addition, the Python API includes methods for managing thematic information. The system provides enhanced functionality which is demonstrated in two use cases.

Conclusions

- Harmonized and space-time aggregated LUCAS dataset

    - Whole dataset: https://zenodo.org/record/7878266

    - List of attributes: https://geoforall.fsv.cvut.cz/st_lucas/tables/list_of_attributes.html

- Open source software: https://gitlab.com/geoharmonizer_inea/st_lucas

    - System deployment package:

      https://gitlab.com/geoharmonizer_inea/st_lucas/st_lucas-system-deployment

    - Python package: https://gitlab.com/geoharmonizer_inea/st_lucas/st_lucas-python-package

    - QGIS plugin: https://gitlab.com/geoharmonizer_inea/st_lucas/st_lucas-qgis-plugin

- Documentation and tutorials: https://geoforall.fsv.cvut.cz/st_lucas/

Future work

- Incorporate LUCAS 2022 survey into ST_LUCAS

- Representativeness of LUCAS points for EO applications

Thank you for your attention!

tomas.boucek@fsv.cvut.cz