**ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE**

FAKULTA STAVEBNÍ

STUDIJNÍ PROGRAM GEODÉZIE A KARTOGRAFIE

OBOR GEOMATIKA



DIPLOMOVÁ PRÁCE

NÁVRH ZÁSUVNÉHO MODULU QGIS PRO STAŽENÍ A ANALÝZU DATOVÉ SADY LUCAS

Vedoucí práce: Ing. Martin Landa, Ph.D.

Katedra geomatiky

|  |  |
| --- | --- |
| 2021 | Bc. Jaroslav ZEMAN |

Obsah obrázku text

Popis byl vytvořen automaticky

**Prohlášení**

Prohlašuji, že jsem diplomovou práci na téma „Návrh zásuvného modulu QGIS pro stažení a analýzu datové sady LUCAS“ vypracoval samostatně a veškeré použité zdroje jsem uvedl v kapitole „Použité zdroje“.

|  |  |
| --- | --- |
| V Praze dne | Jaroslav Zeman |
| . . . . . . . . . . . . . . . . . . . | . . . . . . . . . . . . . . . . . . . |

**Poděkování**

**Anotace**

Tato diplomová práce se zabývá tvorbou zásuvného modulu pro program QGIS. Samotné implementaci předcházela harmonizace datové sady LUCAS, překlad do nomenklatury CORINE a zveřejnění na mapovém serveru.

Zásuvný modul poskytuje možnost na základě zvolených parametrů stáhnout pomocí WFS služby body z datové sady LUCAS. Pro vygenerování dotazu a samotné stažení dat je zásuvným modulem využívána knihovna pyeumap. Po stažení dat jsou body zobrazeny v mapovém okně a je vygenerována jejich statistika.

**Klíčová slova**

QGIS, zásuvný modul, LUCAS, PostGIS, harmonizace, Python

**Annotation**

Překlad anotace…..

**Key words**

QGIS, plugin, LUCAS, PostGIS, harmonization, Python

**Obsah**

[Úvod 9](#_Toc70583734)

[1 LUCAS 10](#_Toc70583735)

[1.1 LUCAS 10](#_Toc70583736)

[1.2 Sběr dat 11](#_Toc70583737)

[1.3 Kontrola kvality 12](#_Toc70583738)

[1.4 Klasifikace krajiny 12](#_Toc70583739)

[1.4.1 Land cover 13](#_Toc70583740)

[1.4.2 Land use 14](#_Toc70583741)

[1.5 Rešerše 14](#_Toc70583742)

[2 Harmonizace 15](#_Toc70583743)

[2.1 Přejmenování atributu 16](#_Toc70583744)

[2.1.1 SQL dotazy 17](#_Toc70583745)

[2.2 Změna hodnot 17](#_Toc70583746)

[2.2.1 SQL dotazy 17](#_Toc70583747)

[2.3 Převod na intervaly 18](#_Toc70583748)

[2.3.1 SQL dotazy 18](#_Toc70583749)

[2.4 Nevalidní hodnoty 18](#_Toc70583750)

[2.4.1 SQL dotazy 19](#_Toc70583751)

[2.5 Překlad LC a LU nomenklatury do nomenklatury CORINE 19](#_Toc70583752)

[2.6 Změna datových typů 19](#_Toc70583753)

[2.6.1 SQL dotaz 20](#_Toc70583754)

[2.7 Nové atributy 21](#_Toc70583755)

[2.8 Zveřejnění dat 21](#_Toc70583756)

[3 Použité technologie 21](#_Toc70583757)

[3.1 QGIS 21](#_Toc70583758)

[3.2 Python 22](#_Toc70583759)

[3.3 PostgreSQL 22](#_Toc70583760)

[3.4 GeoServer 22](#_Toc70583761)

[4 Postup implementace 23](#_Toc70583762)

[5 Modul lucas 24](#_Toc70583763)

[5.1 Sestavení dotazu 24](#_Toc70583764)

[5.2 Vytvoření výstupů 25](#_Toc70583765)

[5.3 Agregace land cover tříd 26](#_Toc70583766)

[5.4 Jupyter Notebook 26](#_Toc70583767)

[6 Zásuvný modul 27](#_Toc70583768)

[6.1 Funkcionalita 28](#_Toc70583769)

[6.1.1 Download 28](#_Toc70583770)

[6.1.2 Class aggregation 30](#_Toc70583771)

[Závěr 32](#_Toc70583772)

[Použité zdroje 33](#_Toc70583773)

[Seznam obrázků 34](#_Toc70583774)

[Seznam tabulek 35](#_Toc70583775)

[Seznam zkratek 36](#_Toc70583776)

[Seznam příloh 37](#_Toc70583777)

# Úvod

Tato diplomová práce se zabývá tvorbou nástroje, který umožní snadný přístup k bodům datové sady LUCAS.

Práce vznikla v rámci mezinárodního projektu nazvaného „Geo-harmonizer: EU - wide automated mapping system for harmonization of Open Data based on FOSS4G and Machine Learning“. Prací na tomto projektu se zabývají odborníci z pěti zemí Evropské unie, a to Německa, Holandska, Chorvatska, Rumunska a České republiky. Českou republiku zastupuje ČVUT. Cílem víš zmíněného projektu je harmonizace prostorových dat napříč členskými státy Evropské unie a jejich následná publikace. Snahou je zpřístupnit data nejen odborníkům, ale i široké veřejnosti v takové formě, aby byla snadno použitelná. [1]

V první části diplomové práce je představena datová sada LUCAS. Je zde popsáno, jaká data obsahuje a jaká je jejich využitelnost. Podstatné je znát historii sady LUCAS a její vývoj. Poprvé byla data pro LUCAS sbírána v roce 2006 na území jedenácti států Evropské unie. Od té doby probíhá měření každé tři roky. Nejaktuálnější data jsou tedy v současné době z roku 2018. Nejvíce jsou využívány atributy nazvané land cover a land use, které především určují charakter krajiny v okolí daného bodu.

Druhá část popisuje způsoby, jak byla data před publikací upravována a harmonizována. Datová sada LUCAS se neustále vyvíjí. S tím souvisí změny, ke kterým docházelo při každém novém sběru dat. Jelikož od roku 2006 doposud (2018) bylo provedeno pět měření, je možné identifikovat značné množství změn. Aby bylo možné data z jednotlivých let porovnávat, bylo třeba je sjednotit.

Pro rozšíření způsobů využití dat byla pomocí atributů land cover a land use vytvořena překladová tabulka do nomenklatury CORINE, která popisuje charakter krajiny jiným způsobem.

Další část popisuje softwarové nástroje, které byly při práci využity. Nejvýznamnější z nich byl programovací jazyk Python a jeho rozšíření pro program QGIS zvaný PyQGIS, pomocí kterého byl vytvářen samotný zásuvný modul a modul lucas. Pro práci s datovou sadou LUCAS a její úpravu byl použit jazyk PostGIS a jeho rozšíření PostgreSQL. V neposlední řadě je zde popsán samotný program QGIS, pro který byl zásuvný modul vytvářen a mapový server, na kterém jsou data zveřejněna.

Čtvrtá část obsahuje obecný popis implementace modulu lucas a samotného zásuvného modulu.

V páté kapitole je představena funkcionalita modulu lucas. Dále je zde zmíněn Jupyter Notebook, jenž byl vytvořen jako názorná ukázka možností tohoto modulu.

V závěrečné části je vysvětlena tvorba zásuvného modulu. Zároveň jsou zde zmíněny možnosti, jak zásuvný modul používat, tak aby uživatel obdržel požadovaná data.

# LUCAS

## LUCAS

LUCAS je zkratkou Land Use and Coverage Area frame Survey (průzkum využití půdy a krajinného pokrytí). Některé státy vytvářejí vlastní produkty mapující svá území, avšak tyto produkty nejsou vzájemně kompatibilní, z tohoto důvodu není možné získat na jejich základě dobrou představu o stavu a vývoji krajiny v celé Evropské unii. Od roku 2006 vytváří Eurostat databázi bodů nazvanou LUCAS, popisující charakter krajiny na území Evropské unie. Díky této databázi je možné snadno detailně analyzovat stav a změny v krajině napříč státy.

Vznik databáze LUCAS potvrdily Evropský parlament a Rada Evropy dne 22.5. 2000 rozhodnutím 1445/2000/ES „On the application of aerial-survey and remote-snsing techniques to the agricultural statistics for 1999 to 2003“ (O použití technik leteckého průzkumu a dálkového průzkumu Země pro zemědělské statistiky v letech 1999 až 2003). První roky probíhaly přípravy a testování. Počínaje rokem 2006 je průzkum prováděn každé 3 roky. Dosud bylo provedeno pět měření, a to v letech 2006, 2009, 2012, 2015 a 2018. Na obrázku (Obr. 1) jsou znázorněny body měřené v roce 2018.

Obsah obrázku mapa

Popis byl vytvořen automatickyPři terénním šetření se zjišťuje nejen současný krajinný pokryv a využití území, ale také informace o životním prostředí. Pořizují se fotografie a u vybraných bodů se odebírají vzorky půdy, které jsou následně analyzovány v laboratoři. Postupně se rozšiřuje nejen počet sledovaných atributů (charakteristiky popisující území), ale také počet států, na jejichž území je měření prováděno. Tabulka (Tab. 1) zobrazuje počet států, na jejichž území bylo měřeno v jednotlivých letech a zároveň počty měřených bodů.

Obr. : Rozložení LUCAS bodů měřených v roce 2018

|  |  |  |
| --- | --- | --- |
| Rok | Počet států EU | Počet bodů |
| 2006 | 11 | 168402 |
| 2009 | 23 | 234623 |
| 2012 | 27 | 270272 |
| 2015 | 28 | 339696 |
| 2018 | 28 | 337854 |

Tab. : Rozsah LUCAS

Data LUCAS se využívají v mnoha rozdílných oborech nejen pro ochranu přírody, lesní a vodní hospodářství, městské plánování, zemědělství, ale například i pro monitorování biologické rozmanitosti.

[2], [3]

<https://ec.europa.eu/eurostat/web/lucas/data/primary-data> (data)

<https://ec.europa.eu/eurostat/web/lucas/overview>

## Sběr dat

Měřené body jsou rozmístěny po celém zájmovém území (členské státy Evropské unie) 2 km od sebe v pravidelné mřížce. Body je možné rozdělit do dvou skupin podle toho, zda je možné se k nim bez problému dostat, v tom případě jsou interpretovány v terénu (in situ). Druhou skupinu tvoří body, k nimž není přístup možný, případně by byl příliš náročný, tyto body jsou fotointerpretovány v kanceláři. V roce 2018 bylo necelých 240000 bodů navštíveno v terénu a pouze necelých 100000 bodů interpretováno v kanceláři. Z těchto čísel vyplývá, že data popisující pouze necelou třetinu celkového počtu bodů nebyla určena terénním šetřením.

Obsah obrázku stůl

Popis byl vytvořen automatickyPrůzkum v terénu provádí vyškolený pracovník, který s pomocí podrobného popisu jednotlivých sledovaných vlastností charakterizuje daný bod a jeho okolí. Údaje jsou zaznamenávány do dotazníku, který pro většinu atributů obsahuje výčet možností. Jednu z částí dotazníku, zabývající se krajinným pokryvem a využitím, znázorňuje obrázek (Obr. 1). Tímto způsobem se předchází vyplňování nevalidních informací. Je snaha o to, aby všichni pracovníci podílející se na sběru dat používaly stejné metody a výsledek byl tedy objektivní.

Obr. : Ukázka dotaníku LUCAS [4]

Informace o bodech, které není možné navštívit fyzicky jsou získávány fotointerpretací. Pro tu jsou využívány ortofoto snímky, případně fotografie z terénu.

Každý bod je podrobně popsán velkým množstvím vlastností. Mezi zjišťované charakteristiky patří například poloha bodu, datum měření, krajinný pokryv, využití území, výška stromů, sklonitost terénu a mnohé další. Je pořizována fotografie samotného bodu i jeho okolí. V některých případech je navíc odebírán vzorek půdy, který je následně podroben rozboru.

[2], [3]

## Kontrola kvality

Data LUCAS jsou kontrolována různými způsoby ve třech fázích. Nejprve je automatizovaně zjišťována úplnost a konzistence dat. Následně jsou veškeré body kontrolovány vizuálně. Ve třetí fázi jsou body kontrolovány jednotlivě nezávislým kontrolorem kvality. Tímto způsobem je ověřována přesnost údajů u následujících bodů:

* 36 % bodů z celého datasetu
* Prvních 20 % bodů měřených novým pracovníkem

Víše zmíněné kontroly odhalí určitou chybu, případně chybějící informaci u zhruba 7 % z celkového počtu bodů. Nedostatky jsou oznámeny terénnímu pracovníkovi, který je následně opraví.

[2], [3]

## Klasifikace krajiny

Jak již bylo zmíněno LUCAS využívá pro popis krajiny dva hlavní parametry, kterými jsou krajinné pokrytí (land cover – LC) a využití půdy (land use – LU). Pro tyto parametry byly vytvořeny nomenklatury, pomocí kterých je možno přesně definovat charakter daného místa. Jelikož v okolí každého bodu se může vyskytovat vícero krajinných typů, případně může být využíváno vícero způsoby, určuje pracovník v terénu hlavní (LC1/LU1) a vedlejší třídu (LC2/LU2). Zároveň stanovuje procentuální zastoupení těchto tříd na daném bodě a v jeho nejbližším okolí. Podrobný popis toho, co která třída obsahuje je dostupný na webových stránkách Eurostatu.

### Land cover

Krajinné pokrytí určuje fyzický a biologický pokryv zemského povrchu, včetně uměle vytvořených ploch, zemědělských oblastí, lesů, přirozených a částečně přirozených oblastí, mokřadů, vodních těles. LUCAS v současné době používá 8 kategorií, které se dále dělí na 29 tříd a 76 podtříd. Hlavní kategorie jsou označeny velkými písmeny A-H. Pro podrobnější dělení jsou používány kódy složené z písmene označujícího kategorii a čísla. Obrázek (Obr.:) vyobrazuje první a druhou úroveň dělení.

Obsah obrázku stůl

Popis byl vytvořen automaticky[5]

Obr. : Land cover třídy [2]

### Land use

Využití půdy popisuje, jakým způsobem je krajina využívána. Tímto způsobem můžeme určit, zda je dané území užíváno pro bydlení, průmysl, zemědělství, dopravu a další účely. LUCAS v současné době rozlišuje 4 hlavní LU kategorie, které se dělí do 16 tříd a ty následně do 40 podtříd. Kódy jednotlivých tříd jsou vytvořeny kombinací písmene „U“ a číselného označení. Obrázek (Obr.: ) znázorňuje rozdělení LU tříd první a druhé úrovně.

[5]

Obsah obrázku stůl

Popis byl vytvořen automaticky

Obr. : Land use třídy [2]

<https://geoportal.cuzk.cz/(S(sgmfbk3yfcqbsovpwd1aet1x))/Default.aspx?mode=TextMeta&text=INSPIRE_K_Pokryv&side=INSPIRE_dSady&head_tab=sekce-04-gp&menu=421>

<https://ec.europa.eu/eurostat/statistics-explained/index.php/LUCAS_-_Land_use_and_land_cover_survey#The_LUCAS_survey>

## Rešerše

Přestože jsou data z datasetu LUCAS bodů hojně využívána v mnoha různých oborech, nebyl v minulosti velký zájem o jejich harmonizaci a jejich snadnější zpřístupnění většímu množství uživatelů.

V současné době jsou data odpovídající jednotlivým rokům dostupná na webových stránkách tvůrce tohoto datasetu, tedy Evropského statistického úřadu (Eurostat) <https://ec.europa.eu/eurostat/web/lucas/data/primary-data>. Tato data mají dva významné nedostatky. Prvním z nich je skutečnost, že data nejsou harmonizována, což ztěžuje možnost jejich porovnávání napříč roky. Druhým nedostatkem je forma, ve které jsou dostupná. Data je možné stáhnout pouze ve formátu *csv*. Vzhledem k tomu, že se jedná o prostorová data, je nutné je pro efektivní práci s nimi převést do jiného formátu.

Víše popsané problémy částečně řešil tým, jehož součástí byli zaměstnanci Joint Research Centre, Eurostatu a GOPA Luxembourg. Způsob, jak s daty pracovali je popsán v článku Harmonised LUCAS in-situ land cover and use database for field surveys from 2006 to 2018 in the European Union ([3]). Hlavním cílem jejich práce byla harmonizace LUCAS. Výsledná data je možné snadno porovnávat napříč roky. Byly vytvořeny nové atributy, kterými jsou například rok měření daného bodu, počet opakovaných navštívení bodu a geometrie popisující polohu bodu. Hodnoty atributů, které jsou v původní databázi vyjádřeny kódem byly převedeny na textový řetězec. Výhodou této změny je, že z hodnoty atributu je na první pohled zřejmé vlastnost, kterou popisuje, není tedy nutná tabulka interpretující jednotlivé kódy. Nevýhodou uložení hodnot atributů pomocí textových řetězců je obtížnější práce s nimi. Data je možné stáhnou pomocí FTP (File Transfer Protocol). Dostupná jsou na adrese <https://jeodpp.jrc.ec.europa.eu/ftp/jrc-opendata/LUCAS/LUCAS_harmonised/>.

Informace popisující body jsou uložené v souborech *csv*. Pouze geometrie se základní identifikací jednotlivých bodů jsou uloženy ve formátu *shapefile*. Pro plnohodnotné využívání veškerých informací je uživatel nucen data předzpracovat.

[3]

# Harmonizace

Přestože je databáze LUCAS vytvářena a udržována organizací Eurostat, tedy statistickým úřadem Evropské unie, vyskytuje se v ní mnoho nekonzistencí napříč roky. Jedním z důvodů, proč tomu tak je, je postupný vývoj a rozšiřování počtu měřených parametrů území. Tabulka (Tab. 4) zobrazuje jejich počet v jednotlivých letech.

|  |  |
| --- | --- |
| Rok | Počet atributů |
| 2006 | 20 |
| 2009 | 44 |
| 2012 | 46 |
| 2015 | 59 |
| 2018 | 97 |

Tab. : Počet LUCAS atributů

Vznik nových atributů v průběhu let nebyl jediný typ změny. Dalšími změnami, které v mnoha případech znepřehlednily databázi jsou: zánik atributu (přestal být měřen), změna názvu atributu, změna hodnot, které daný atribut nabývá, a změna datového typu. Problémem, který se v databázi Lucas vyskytuje v nemalé míře je, že některé atributy nabývají neplatných hodnot. Neplatnou hodnotou rozumíme takovou hodnotu, která není zmíněna v dokumentaci. Výše zmíněné nekonzistence a chyby byly identifikovány, a ty které bylo možné opravit, byly opraveny. Byla snaha veškeré změny vztahovat k v současnosti nejaktuálnější verzi, tedy atributům z roku 2018.

Jako vzor pro harmonizace dat sloužil v mnoha případech článek Harmonised LUCAS in-situ land cover and use database for field surveys from 2006 to 2018 in the European Union ([3]) popsaný v kapitole 1.5, jenž se touto problematikou zabýval. V následujících podkapitolách jsou popsány jednotlivé případy, ve kterých se přístupy k danému problému, využité pro tuto práci, s článkem shodují a také ty, ve kterých byl postup odlišný.

Před samotnou harmonizací bylo nutné data ve formátu *csv*, stažená z Eurostatu importovat do databázového sftware. Následně pomocí SQL dotazů bylo možné provést podrobnou analýzu veškerých atributů v jednotlivých letech. Na základě této analýzy byly definovány XXX typy úprav, které byly implementovány do dat. Příklady jednotlivých změň, které databázi upravovaly jsou popsány v následujících podkapitolách.

## Přejmenování atributu

V případě, že se lišil pouze název atributu, byl změněn do podoby, kterou měl v roce 2018. Takováto změna byla provedena například u atributu *protected\_area*, který byl přejmenován na *special\_status*. Název *protected\_area* byl používán v letech 2012 a 2015. V roce 2018 se již atribut s tímto názvem nevyskytuje, avšak nahradil ho atribut *special\_status* obsahující totožné hodnoty.

Atributy byly přejmenovávány stejným způsobem, jaký je popsán v článku [3], až na atribut *soil\_crop* vyskytující se v letech 2009, 2012 a 2015. Tento atribut byl přejmenován na *crop\_residues*, což souvisí s vlastností, kterou atribut popisuje. V letech 2009, 2012 a 2015 představuje atribut procentuální zastoupení zbytků plodin na poli po sklizni. V roce 2018 je pomocí atributu *crop\_residues* rozlišováno pouze zda se na poli nějaké posklizňové zbytky plodin nachází. Jelikož dle mého názoru není zásadní podíl zmíněných zbytků, byly atributy spojeny do jednoho. Intervaly procentuálního zastoupení používané do roku 2015 byly předeny podle následujících pravidel.

Pro automatické vygenerování SQL dotazů sloužících k přejmenování atributů byl pomocí jazyka Python vytvořen skript *Attribute\_rename.py.* Vstupem pro tento skript je *csv* soubor obsahující změny názvů atributů. Obsah tohoto *csv* souboru, a tedy i změny názvů atributů zobrazuje tabulka (Tab.: ).

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Rok** | **Původní název** | **Nový**  **název** | **Rok** | **Původní**  **název** | **Nový**  **název** |
| 2006 | surv\_date | survey\_date | 2012 | photo\_s | photo\_south |
| 2006 | x\_laea | th\_lat | 2012 | photo\_p | photo\_point |
| 2006 | y\_laea | th\_long | 2012 | tree\_height\_srv | tree\_height\_survey |
| 2009 | area\_size | parcel\_area\_ha | 2012 | tree\_height\_mat | tree\_height\_maturity |
| 2009 | lc1\_pct | lc1\_perc | 2012 | soil\_plough | lndmng\_plough |
| 2009 | lc2\_pct | lc2\_perc | 2012 | gps\_alt | gps\_altitude |
| 2009 | lc1\_species | lc1\_spec | 2012 | protected\_area | special\_status |
| 2009 | lc2\_species | lc2\_spec | 2012 | soil\_stones | soil\_stones\_perc |
| 2009 | land\_mngt | grazing | 2015 | area\_size | parcel\_area\_ha |
| 2009 | obs\_dir | obs\_direct | 2015 | lc1\_pct | lc1\_perc |
| 2009 | photo\_e | photo\_east | 2015 | lu1\_pct | lu1\_perc |
| 2009 | photo\_w | photo\_west | 2015 | lu2\_pct | lu2\_perc |
| 2009 | photo\_n | photo\_north | 2015 | lc2\_pct | lc2\_perc |
| 2009 | photo\_s | photo\_south | 2015 | lc1\_species | lc1\_spec |
| 2009 | photo\_p | photo\_point | 2015 | lc2\_species | lc2\_spec |
| 2009 | tree\_height\_srv | tree\_height\_survey | 2015 | land\_mngt | grazing |
| 2009 | soil\_plough | lndmng\_plough | 2015 | obs\_dir | obs\_direct |
| 2009 | gps\_alt | gps\_altitude | 2015 | photo\_e | photo\_east |
| 2009 | soil\_stones | soil\_stones\_perc | 2015 | photo\_w | photo\_west |
| 2012 | area\_size | parcel\_area\_ha | 2015 | photo\_n | photo\_north |
| 2012 | lc1\_pct | lc1\_perc | 2015 | photo\_s | photo\_south |
| 2012 | lc2\_pct | lc2\_perc | 2015 | photo\_p | photo\_point |
| 2012 | lc1\_species | lc1\_spec | 2015 | protected\_area | special\_status |
| 2012 | lc2\_species | lc2\_spec | 2015 | soil\_plough | lndmng\_plough |
| 2012 | land\_mngt | grazing | 2015 | pi\_extension | office\_pi |
| 2012 | obs\_dir | obs\_direct | 2015 | trees\_height\_maturity | tree\_height\_maturity |
| 2012 | photo\_e | photo\_east | 2015 | trees\_height\_survey | tree\_height\_survey |
| 2012 | photo\_w | photo\_west | 2015 | soil\_stones | soil\_stones\_perc |
| 2012 | photo\_n | photo\_north |  |  |  |

Tab. : Změny názvů atributů

### SQL dotazy

ALTER TABLE lucas.lucas2012 RENAME COLUMN protected\_area TO special\_status;

ALTER TABLE lucas.lucas2015 RENAME COLUMN protected\_area TO special\_status;

## Změna hodnot

Pokud byly hodnoty atributu odlišné v jednotlivých letech a zároveň nabývaly platných hodnot, byly tyto hodnoty změněny tak, aby byly konzistentní. Tato změna byla realizována například u lc\_lu\_special\_remark a dokumentuje ji tabulka (Tab. 5). Z tabulky je zřejmé, že rozdílné kódové hodnoty popisují stejnou vlastnost. Například v roce 2015 je dočasně zaplavené území označeno hodnotou 10, avšak v roce 2018 je tato skutečnost zaznamenána hodnotou 9.

|  |  |  |  |
| --- | --- | --- | --- |
| Původní kód  (2015) | Popis  (2015) | Nový kód  (2018) | Popis  (2018) |
| 1 | Tilled and/or sowed | 2 | Tilled/sowed |
| 2 | Harvested field | 1 | Harvested field |
| 7 | No remark | 10 | No remark |
| 8 | Not relevant | 88 | Not relevant |
| 9 | Temporary dry | 8 | Temporary dry |
| 10 | Temporary flooded | 9 | Temporary flooded |

Tab. : Změna hodnot atributu lc\_lu\_special\_remark

### SQL dotazy

UPDATE lucas.lucas2015 SET lc\_lu\_special\_remark = '2' WHERE lc\_lu\_special\_remark = '1';

UPDATE lucas.lucas2015 SET lc\_lu\_special\_remark = '1' WHERE lc\_lu\_special\_remark = '2';

UPDATE lucas.lucas2015 SET lc\_lu\_special\_remark = '10' WHERE lc\_lu\_special\_remark = '7';

UPDATE lucas.lucas2015 SET lc\_lu\_special\_remark = '88' WHERE lc\_lu\_special\_remark = '8';

UPDATE lucas.lucas2015 SET lc\_lu\_special\_remark = '8' WHERE lc\_lu\_special\_remark = '9';

UPDATE lucas.lucas2015 SET lc\_lu\_special\_remark = '9' WHERE lc\_lu\_special\_remark = '10';

## Převod na intervaly

V tomto případě byly způsob úpravy podobný tomu, který byl popsán v kapitole 2.2. Změna se týkala například atributů zaznamenávajících procentuální zastoupení jednotlivých LC (Land cover) a LU (Land use) tříd. Na rozdíl od příkladu z kapitoly 2.2 nebylo možné použít jako referenční atribut ten z roku 2018, a to z následujícího důvodu. V roce 2018 bylo procentuální zastoupení určováno v absolutních procentech, nikoliv v procentuálních intervalech označených kódem, jak tomu bylo v předchozích letech. Jednotlivé hodnoty tedy bylo nutné agregovat do tříd tak, aby byla data konzistentní. Výhodou této agregace je zpřehlednění a lepší využitelnost dat. Tabulka (Tab. 4) popisuje použité intervaly převzaté z roku 2009 pro atribut lc1\_perc.

|  |  |
| --- | --- |
| Kód | Hodnota *[%]* |
| 0 | 0 |
| 1 | 1 ≤ LC < 10 |
| 2 | 10 ≤ LC < 25 |
| 3 | 25 ≤ LC < 50 |
| 4 | 50 ≤ LC < 75 |
| 5 | 75 ≤ LC ≤ 100 |

Tab. : Hodnoty atributu lc1\_perc

### SQL dotazy

UPDATE lucas.lucas2018 SET lc1\_perc = '1' where lc1\_perc::int >= 1 and lc1\_perc::int < 10;

UPDATE lucas.lucas2018 SET lc1\_perc = '2' where lc1\_perc::int >= 10 and lc1\_perc::int < 25;

UPDATE lucas.lucas2018 SET lc1\_perc = '3' where lc1\_perc::int >= 25 and lc1\_perc::int < 50;

UPDATE lucas.lucas2018 SET lc1\_perc = '4' where lc1\_perc::int >= 50 and lc1\_perc::int < 75;

UPDATE lucas.lucas2018 SET lc1\_perc = '5' where lc1\_perc::int >= 75 and lc1\_perc::int =< 100;

## Nevalidní hodnoty

Nejobtížnější korekcí byla úprava hodnot atributů, které byly nevalidní, tedy se vyskytovaly v databázi, přestože by v ní dle dokumentace být neměly. V mnoha případech byla touto nevalidní hodnotou 0 nebo prázdný řetězec. Každá neplatná hodnota byla detailně zkoumána. Především byl porovnáván počet bodů touto hodnotou označený s počty bodů jednotlivých hodnot stejného atributu v ostatních letech. V případě, že bylo zřejmé, jakou hodnotu by měly zmíněné body nabývat, byla tato hodnota změněna, v opačném případě byla hodnota nahrazena na kódem -1, reprezentujícím nevalidní hodnotu. Obě víše zmíněné eventuality byly využity například u atributu wm\_delivery. Tabulka (Tab. 7) zobrazuje počty bodů v jednotlivých letech, kterým odpovídá daná hodnota atributu. Nevalidní hodnoty se vyskytují v roce 2009 a jsou označeny červeně. Z tabulky je zřejmé, že nevalidní hodnota 0 odpovídá hodnotě 8. U dalších nevalidních hodnot, kterými jsou 5, 6, 10 a 12 již není možné jednoznačně určit co představují. Tyto kódy byly proto změněny na -1.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Kód | Počet bodů  (2009) | Počet bodů  (2012) | Počet bodů  (2015) | Počet bodů  (2018) |
| 0 | 227586 | - | - | - |
| 1 | 420 | 607 | 853 | 545 |
| 2 | 1167 | 1220 | 2142 | 804 |
| 3 | 51 | 4671 | 5251 | 4290 |
| 4 | 4015 | 2370 | 15398 | 2435 |
| 5 | 16 | - | - | - |
| 6 | 8 | - | - | - |
| 8 | 1353 | 261404 | 316052 | 329780 |
| 10 | 2 | - | - | - |
| 12 | 5 | - | - | - |

Tab. : Počet bodů atributu wm\_delivery v letech

### SQL dotazy

UPDATE lucas.lucas2009 SET wm\_delivery = '8' WHERE wm\_delivery = '0';

UPDATE lucas.lucas2009 SET wm\_delivery = '-1' WHERE wm\_delivery IN ('5', '6', '10', '12');

## Překlad LC a LU nomenklatury do nomenklatury CORINE

Pro vytvoření reklasifikační tabulky byl využit dokument [6], ve kterém jsou popsány jednotlivé kombinace LC a LU tříd a způsob, jak tyto kombinace převést do nomenklatury CORINE. V závislosti na tomto dokumentu vznikly tři překladové tabulky, a to jedna pro každou ze tří CORINE úrovní. V jednotlivých tabulkách představují řádky CORINE třídy, sloupce reprezentují kombinace LC a LU tříd. Pokud je překlad mezi zvolenými třídami možný, je průsečík daného sloupce a řádku označen hodnotou 1, v opačném případě se zde vyskytuje hodnota 0.

Následně byl v databázi LUCAS bodů vytvořen nový atribut představující nomenklaturu CORINE a pro naplnění tohoto atributu hodnotami byly aplikovány reklasifikační tabulky.

[6]

## Změna datových typů

Víše popsané úpravy byly aplikovány na datech odpovídajících jednotlivým rokům. Pro další změny byla veškerá data vložena do jedné tabulky. Následně bylo nutné nastavit atributům odpovídající datové typy. Byl vytvořen skript *Attribute\_retype.py* pomocí jazyka Python, jehož vstupem byl *csv* soubor definující atributům jejich datové typy. Skript automaticky generuje SQL dotazy, které mění atributům datové typy. Tabulka (Tab.: ) obsahuje názvy atributů spolu s příslušnými datovými typy. Jelikož pře úpravou byl datový typ všech atributů textový řetězec (string), obsahuje tabulka pouze atributy, jejichž datový typ měl být jiný.

|  |  |  |  |
| --- | --- | --- | --- |
| Název atributu | Datový typ | Název atributu | Datový typ |
| OFFICE\_PI | integer | CPRNC\_LC1S | integer |
| GPS\_PROJ | integer | CPRNC\_LC1W | integer |
| GPS\_ALTITUDE | integer | CPRN\_LC1N\_BRDTH | integer |
| GPS\_EW | integer | CPRN\_LC1E\_BRDTH | integer |
| OBS\_DIRECT | integer | CPRN\_LC1S\_BRDTH | integer |
| OBS\_TYPE | integer | CPRN\_LC1W\_BRDTH | integer |
| LC1\_PERC | integer | CPRN\_URBAN | integer |
| LC2\_PERC | integer | CPRN\_IMPERVIOUS\_PERC | integer |
| LU1\_PERC | integer | INSPIRE\_PLCC1 | integer |
| LU2\_PERC | integer | INSPIRE\_PLCC2 | integer |
| PARCEL\_AREA\_HA | integer | INSPIRE\_PLCC3 | integer |
| TREE\_HEIGHT\_SURVEY | integer | INSPIRE\_PLCC4 | integer |
| TREE\_HEIGHT\_MATURITY | integer | INSPIRE\_PLCC5 | integer |
| LNDMNG\_PLOUGH | integer | INSPIRE\_PLCC6 | integer |
| GRAZING | integer | INSPIRE\_PLCC7 | integer |
| SPECIAL\_STATUS | integer | INSPIRE\_PLCC8 | integer |
| LC\_LU\_SPECIAL\_REMARK | integer | EUNIS\_COMPLEX | integer |
| SOIL\_TAKEN | integer | GRASSLAND\_SAMPLE | integer |
| SOIL\_STONES\_PERC | integer | GRASS\_CANDO | integer |
| PHOTO\_POINT | integer | WM | integer |
| PHOTO\_NORTH | integer | WM\_SOURCE | integer |
| PHOTO\_EAST | integer | WM\_TYPE | integer |
| PHOTO\_SOUTH | integer | WM\_DELIVERY | integer |
| PHOTO\_WEST | integer | EROSION\_CANDO | integer |
| EX\_ANTE | integer | BIO\_SAMPLE | integer |
| CAR\_LATITUDE | float | SOIL\_BIO\_TAKEN | integer |
| CAR\_LONGITUDE | float | BULK0\_10\_SAMPLE | integer |
| CAR\_EW | integer | SOIL\_BLK\_0\_10\_TAKEN | integer |
| GPS\_PREC | integer | BULK10\_20\_SAMPLE | integer |
| OBS\_DIST | integer | SOIL\_BLK\_10\_20\_TAKEN | integer |
| OBS\_RADIUS | integer | BULK20\_30\_SAMPLE | integer |
| FEATURE\_WIDTH | integer | SOIL\_BLK\_20\_30\_TAKEN | integer |
| LM\_PLOUGH\_SLOPE | integer | STANDARD\_SAMPLE | integer |
| LM\_PLOUGH\_DIRECT | integer | SOIL\_STD\_TAKEN | integer |
| LM\_STONE\_WALLS | integer | ORGANIC\_SAMPLE | integer |
| LM\_GRASS\_MARGINS | integer | SOIL\_ORG\_DEPTH\_CANDO | integer |
| CPRN\_CANDO | integer | TH\_EW | integer |
| CPRN\_LC1N | integer | CROP\_RESIDUES | integer |
| CPRNC\_LC1E | integer |  |  |

Tab. : Změny datových typů atributů

### SQL dotaz

ALTER TABLE LUCAS\_table ALTER COLUMN office\_pi TYPE INTEGER USING office\_pi::INTEGER

## Nové atributy

Poslední úpravou datové sady LUCAS bylo vytvoření nových atributů. Tyto nové atributy data zpřehledňují a usnadňují práci s nimi.

## Zveřejnění dat

Po veškerých úpravách, spočívajících v úpravě dat tak, aby byla konzistentní napříč roky, opravě chyb, překladu do nomenklatury CORINE, změně datových typů a doplnění nových atributů byla data zveřejněna na mapovém serveru. Pro zpřístupnění dat byl využit GeoServer. Vznikla tak WFS (web feature service) služba, pomocí které je možné k datům přistupovat.

# Použité technologie

## QGIS

QGIS je open source geografický informační systém. Je uveřejněn pod licencí GNU GPL, což mimo jiné zaručuje, že veškerá díla odvozená od tohoto programu musí být publikována pod stejnou licencí, tudíž je zaručen jejich open source charakter. QGIS je možné používat na všech nejrozšířenějších operačních systémech, kterými jsou například MS Windows, GNU/Linux a Mac OS X. Silnou stránkou Programu QGIS jsou zásuvné moduly („pluginy“). Plugin je doplněk programu, který rozšiřuje funkčnost základního programu. Pluginy nevytváří pouze QGIS Development Team, ale také jeho uživatelé.

Tento program začal vytvářet Gary Sherman v roce 2002 pod názvem Quantum GIS. Verze 1.0 vyšla na začátku roku 2009. V roce 2013 byl spolu s vydáním verze 2.0 změněn jeho název na současný QGIS.

V současnosti jsou k dispozici tři verze, a to jedna dlouhodobá - stabilní a dvě krátkodobé, které se rychle vyvíjí a mění.

Obr. : Logo QGIS

<https://www.qgis.org/en/site/getinvolved/styleguide.html> OBRÁZEK

training.gismentors.eu/qgis-zacatecnik/ruzne/qgis\_pluginy.html

## Python

Python je objektově orientovaný programovací jazyk, který vytvořil v roce 1991 Guido van Rossum. V současné době patří mezi nejpoužívanější programovací jazyky. Je možné jej používat na mnoha operačních systémech (Windows, Linux, macOS, …). Python je vytvářen jako open source pod licencí Python Software Foundation License. Zdrojové kódy jsou vystaveny na stránce <http://www.python.org/>. Velkou výhodou jazyka je velké množství modulů, které rozšiřují jeho použitelnost.



Obr. : Logo Python

<https://www.python.org/community/logos/> OBRÁZEK

## PostgreSQL

PostgreSQL je objektově-relační databázový systém, který rozšiřuje jazyk SQL. Vývoj tohoto systému začal v roce 1986 na Kalifornské univerzitě. Je vyvíjen pro všechny hlavní operační systémy, mezi které patří Windows, Linux, macOS. Systém je vyvíjen jako open source pod licencí PostgreSQL License. PostgreSQL obsahuje doplňky, mezi které patří například PostGIS, umožňující efektivní práci s prostorovými daty.

Obr. : Logo PostgreSQL

<https://cs.wikipedia.org/wiki/PostgreSQL> OBRÁZEK

https://www.postgresql.org/about/

## GeoServer

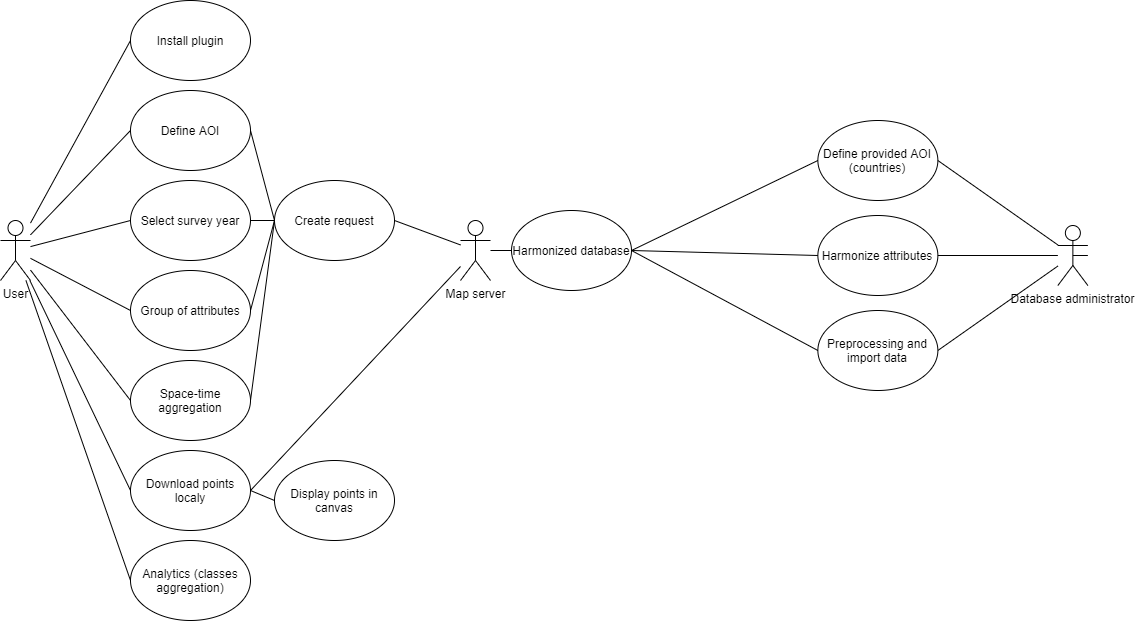
GeoServer je open source mapový server vytvořený pomocí programovacího jazyka Java. Umožňuje především publikaci prostorových dat. GeoServer podporuje standardy vyvíjené mezinárodní organizací OGC (Open Geospatial Consortium). Mezi tyto standardy patří WFS (Web Feature Service) pro práci s vektorovými daty. Pro práci s rastry jsou využívány služby WMS (Web Map Service), WMTS (Web Map Tile Service) a WCS (Web Coverage Service). GeoServer je možné rozšiřovat, například o možnost využívání WPS (Web Processing Service).

Po publikaci na GeoServeru je možné k datům přistupovat pomocí webového rozhraní.

Obr. : Logo GeoServer

http://geoserver.org/about/

# Postup implementace

V této kapitole je popsán průběh implementace. Práci je možné rozdělit do dvou částí, kterými jsou tvorba modulu lucas a vývoj samotného zásuvného modulu, jenž zmíněnou knihovnu využívá.

Obr. : Diagram popisující průběh prací

Obrázek (Obr. 9) zobrazuje diagram obsahující tři aktéry, kterými jsou administrátor databáze Lucas bodů, mapový server a uživatel zásuvného modulu. Jsou zde znázorněny procesy, které musí zmínění aktéři provést tak aby se požadované body dostaly ve správné formě z databáze k uživateli. Úkolem administrátora databáze je především úprava dat, která byla detailně popsána v kapitole 2 Harmonizace. Mapový server, v tomto případě GeoServer, zabezpečuje publikaci dat pomocí WFS služby. WFS služba umožňuje snadný přístup pomocí webového rozhraní k vektorovým datům. Třetím aktérem znázorněným v diagramu je samotný uživatel zásuvného modulu. Ten musí nejprve nainstalovat zásuvný modul do QGIS a následně v něm definovat požadované vlastnosti, podle kterých se vyberou odpovídající body. Následně jsou data uložena do zvoleného adresáře a uživatel s nimi může dále pracovat v prostředí QGIS a to pomocí zásuvného modulu, případně pomocí mnoha funkcí, které program QGIS poskytuje.

# Modul lucas

Modul lucas byl vytvářen v rámci projektu GeoHarmonizer jako jedna ze součástí knihovny pyeumap. Knihovna pyeumap slouží primárně pro snazší zpřístupnění prostorových dat pokrývajících území evropské unie týkajících se životního prostředí a implementaci funkcí umožňující práci s těmito daty.

Modul lucas umožňuje uživateli přístup k datům LUCAS a jejich následné zpracování. Funkcionalitu je možné rozdělit na tři hlavní části. První z nich zabezpečuje sestavení dotazu určeného pro GeoServer na základě požadavků uživatele. Druhá část požadavek zpracuje a odešle na GeoServer. Obdržená data jsou následně uložena na zvolené místo na disku. Třetí část umožňuje zpracování získaných dat. Zpracování dat představuje možnost agregace land cover tříd na základě předem definovaných pravidel.

## Sestavení dotazu

Sestavení dotazu zabezpečuje soubor *request.py*. Tento soubor obsahuje jednu třídu, a to LucasRequest jejíž proměnné definují parametry výsledného dotazu. Těmito proměnnými jsou:

* bbox – Určuje oblast, ve které se mají body nacházet pomocí obdélníku definovaného souřadnicemi vrcholů.
* propertyname, literal, operator, logical – Jsou čtyři parametry, které umožňují výběr na základě atributového dotazu. Zadávají se pomocí nich následující vlastnosti:
  + propertyname – Název atributu na základě kterého má být atributový dotaz vykonán.
  + literar – Hodnota zvoleného atributu.
  + operator – Funkce, která má být použita pro vykonání atributového dotazu. Např.: PropertyIsEqualTo (vybere body u kterých daný atribut nabývá zvolené hodnoty), PropertyIsEqualTo (vybere body u kterých daný atribut nenabývá zvolené hodnoty), PropertyIsGreaterThan (vybere body u kterých daný atribut nabývá vyšší hodnoty, než je hodnota zvolená) atd.
  + logical – Logický operátor spojující atributové dotazy. Určuje, zda mají podmínky platit zároveň (*And*), případně zda postačí platnost jedné z podmínek (*Or*).
* polygoninter – Pomocí textového řetězce ve formátu *gml* obsahujícího lomové body polygonu určuje území, ve kterém se mají body nacházet.
* years – Je parametr, pomocí kterého je možný výběr na základě let, ve kterých byly body měřeny. Zadává se pomocí seznamu obsahující hodnoty 2006, 2009, 2012, 2015, 2018, případně jejich libovolnou kombinaci.
* typename – Definuje název WFS vrstvy.
* staggregation – Tento parametr může nabývat hodnot *True* (pravda), nebo *False* (nepravda) podle toho, zda mají být data časoprostorově agregována či nikoliv.
* group – Parametr group slouží k rozhodnutí, jaké atributy mají data obsahovat.

Třída LucasRequest má tři metody. Nejdůležitější z nich, která vytváří samotný dotaz je metoda *build*. Tato metoda nejprve volá metodu *\_chceck\_args*, která zabezpečuje kontrolu toho, zda jsou vstupní argumenty zadány správně. Následně je dotaz postupně vytvářen pomocí metod OWSLib knihovny. Jednotlivé části dotazu jsou generovány na základě toho, které argumenty byly zadány. Pro spojení dílčích částí jsou využívány metody OWSLib knihovny *And* a *Or*. Na závěr je výsledný objekt převeden do formátu *xml*, který je návratovou hodnotou metody *build*. V případě dotazu vytvořeného kombinací atributového dotazu a minimálního ohraničujícího obdélníku (bbox) není výsledek správný. Tento problém byl opraven ve výsledném *xml* nahrazením chybných částí pomocí metody *replace*.



Obr. : kombinace prostorového a atributového dotazu (xml)

Obrázek (Obr. 10) zobrazuje výsledné *xml* vygenerované metodou build v případě, že mají být vybrány body, které byly měřeny v roce 2012 nebo 2018 a nacházejí se uvnitř obdélníku definovaného souřadnicemi levého dolního rohu (4624127, 2998330) a pravého horního rohu (4650393, 3013986).

## Vytvoření výstupů

O vytvoření výstupů se stará soubor *io.py* a jediná třída kterou obsahuje LucasIO. Nejpodstatnější metodou, kterou zmíněná třída obsahuje je *download*. Metoda *download* zabezpečuje stažení dat na základě parametrů definovaných v dotazu vytvořeném třídou LucasRequest (viz kapitola 5.1). Pro stažení dat je stejně jako v případě vytvoření dotazu využívána knihovna OWSLib. Požadavek je odeslán na GeoServer. WFS vrátí data, která jsou převedena na textový řetězec ve formátu *gml*. Následně je volána metoda *\_load*, jejímž cílem je převést stažená data do formátu *gpkg* a uložit do přechodného adresáře. Jelikož při stahování dat vzniká atribut nazvaný *gml\_id*, který není relevantní a pouze znepřehledňuje atributovou tabulku, je pomocí metody *\_postprocessing* z *gpkg* souboru odstraněn.

Metoda *to\_gml* převádí *gpkg* soubor uložený v přechodném adresáři a navrací textový řetězec ve formátu *gml*.

O uložení *gpkg* souboru do zvolené místo na disku se stará metoda *to\_gpkg*. Vstupním parametrem této metody je cesta ke *gpkg* souboru, který se má vytvořit. Funkcionalita spočívá ve vytvoření kopie souboru dříve uloženého do přechodného adresáře ve zvoleném adresáři a přejmenování souboru na požadovaný název.

Další metodou je *to\_geopandas*, jejímž úkolem je pouze návrat dat ve formátu *GeoDataFrame*. Tento formát usnadňuje práci s atributovou tabulkou pomocí python.

Metody *num\_of\_features* a *is\_empty* spolu souvisí. První z nich opět pracuje s *gpkg* souborem uloženým v přechodném adresáři. Tento soubor otvírá pomocí metody *OpenEx* z knihovny GDAL. Následně určí počet prvků, které obsahuje. Zmíněný počet je návratovou hodnotou metody. Druhá metoda na základě zjištěného počtu prvků určuje, zda soubor obsahuje alespoň jeden prvek. Pokud gpkg soubor není prázdný vrací metoda hodnotu False. V opačném případě je metodou navráceno True.

## Agregace land cover tříd

Třetí částí modulu lucas je soubor *analyze.py*. Tento soubor obsahuje třídu LucasClassAggr, která slouží k agregaci land cover tříd.

Metoda *\_load\_classes* na základě vstupního *json* souboru vytváří slovník obsahující agregační pravidla. Před tím, než je slovník vytvořen probíhá kontrola formy i obsahu *json* souboru. Z *csv* souboru jsou načteny veškeré validní land cover kódy, pomocí kterých je zjišťováno, zda agregační pravidla neobsahují kódy nevalidní. Dále metoda určuje, zda se kódy neopakují, což by vedlo k nejednoznačnosti agregačních pravidel.

O samotnou agregaci se stará metoda *apply*. Tato metoda připojí gpkg soubor pomocí sqlite3, tak aby bylo možné pomocí python volat SQL dotazy pracující s atributovou tabulkou. SQL dotazy jsou využívány pro vytvoření indexu nad sloupečkem lc1\_h (hlavní land cover třída harmonizovaná), přidání nového atributu lc1\_a a zapsání agregovaných tříd na základě agregačních pravidel do tohoto atributu. Před spouštěním SQL dotazů je na základě názvu vrstvy rozhodováno, zda se jedná o časoprostorově agregovaná data či nikoliv. V případě časoprostorově agregovaných dat je index vytvořen nad všemi lc1\_h sloupečky (za každý rok jeden) a taktéž je vytvořen odpovídající počet nových atributů pro agregované třídy.

## Jupyter Notebook

V rámci projektu GeoHarmonizer byl vytvořen soubor umožňující přehlednou demonstraci funkcionality lucas modulu v prostředí internetu. Tvorbou tohoto souboru se zabýval především Ing. Martin Landa, Ph.D. Pomocí Jupyter Notebook je možné nejen psát kód a spouštět ho, ale také vytvářet uspořádanou strukturu s komentáři jednotlivých částí. Výsledný dokument obsahuje ukázky mnoha možností, které dává uživateli modul lucas popsaný v předchozích kapitolách 5.1, 5.2 a 5.3.

# Zásuvný modul

Kostra zásuvného modulu byla vytvořena pomocí QGIS pluginu Plugin Builder dostupného v oficiálním QGIS plugin repozitáři. Tento plugin vygeneroval po zadaní základních informací o zásuvném modulu jeho kostru. Mezi tyto informace patří například název modulu, jeho umístění, stručný popis funkčnosti a další. Plugin Builder je zobrazen na následujícím obrázku (Obr.:).

## Obsah obrázku text Popis byl vytvořen automaticky Funkcionalita

Obr. : Plugin Builder

Zásuvný modul je rozdělen do dvou částí, z nichž každá plní jinou funkci. První část nazvaná „Download“ zajišťuje stažení požadovaných dat. Pomocí druhé části „Class aggregation“ je možné získaná data upravovat.

### Download

Tato část zásuvného modulu umožňuje uživateli definovat parametry požadovaných dat. Je nutné zvolit zájmové území („Area of interest“). Na výběr jsou možnosti:

* canvas – pokud uživatel zvolí tuto možnost, jsou staženy body nacházející se v území, které aktuálně zobrazuje mapové okno.
* country – po označení této možnosti je nutné zvolit v „combo boxu“ země, na jejichž území se mají body nacházet. V tomto případě je využíván atribut *nuts0*, na základě kterého jsou odpovídající body vybírány.
* vector layer – tato možnost umožňuje zvolit jednu z aktivních vektorových vrstev a získat body uvnitř území pokrytého touto vrstvou

Druhým parametrem, pomocí kterého uživatel definuje vlastnosti bodů, o něž má zájem je rok, kdy byl bod měřen. Do dnešního dne byly provedeny tři měření LUCAS bodů, a to v letech 2006, 2009, 2012, 2015 a 2018. Uživatel pomocí „check boxů“ má možnost označit roky, ve kterých byly body měřeny.

V průběhu pěti měření, která byla doposud provedena narůstal počet charakteristik popisující jednotlivé body a tím pádem i počet atributů. V případě, že by si uživatel stáhl body s veškerými atributy, obsahovala by atributová tabulka mnoho desítek sloupečků a byla by značně nepřehledná. Ve většině případů zajímá uživatele pouze malé množství vlastností a ostatní pouze překážejí. Z tohoto důvody byly atributy rozčleněny do pěti skupin. Tyto skupiny jsou zobrazeny v následující tabulce spolu s počtem atributů, které obsahují.

|  |  |
| --- | --- |
| land cover + land use | 40 |
| land cover + land use + soil | 55 |
| forestry | 2 |
| copernicus | 16 |
| inspire | 8 |

Tab. : Třídy atributů

Je možné si zvolit jednu z výše uvedených skupin a výsledná atributová tabulka bude obsahovat krom povinných atributů definujících především polohu bodu pouze atributy z této skupiny.

Jelikož je velké množství bodů měřeno opakovaně, tedy v různých letech, je možné tato měření porovnávat a zjistit tak jak se dané území vyvíjí v čase. V tomto případě je vhodné, aby každý prvek v atributové tabulce odpovídal jednomu bodu a jednotlivé atributy popisovaly veškerá měření provedená na tomto bodě. Naopak pokud uživatele zajímá stav krajiny v určitém období, je výhodnější, pokud každý prvek představuje jedno měření. Atributová tabulka obsahuje méně nepotřebných atributů a je přehlednější. Parametr, který tímto způsobem definuje charakter získaných dat je časová agregace („Space-time aggregation“).

Posledním parametrem, který je možné zvolit, a tak lépe definovat vlastnosti požadovaných dat je jejich nomenklatura.

Po vybrání veškerých parametrů je možné pomocí tlačítka „Number of points“ před stažením zjistit kolik bodů odpovídá výběru. V případě že počat bodů není dostačující dají se parametry upravit. Následně je nutné definovat cestu, kam má být soubor se staženými daty uložen a pomocí tlačítka „Downoad“ spustit stahování.

### Class aggregation

Obsah obrázku text

Popis byl vytvořen automatickyDruhá část zásuvného modulu „Class aggregation“ umožňuje uživateli definovat si pravidla na základě kterých mají být land cover třídy agregovány. Pravidla je nutné zapsat do json souboru. Tento soubor musí obsahovat takzvané klíče, které reprezentují názvy nových tříd a seznamy land cover tříd třetí úroveň odpovídající jednotlivým klíčům. Na obrázku (Obr.: ) je příklad json souboru, který popisuje převod land cover tříd třetí úrovně na land cover třídy druhé úrovně.

Obr. : Soubor json definující agregaci tříd

Po vytvoření souboru s agregačními pravidly je nutné tento soubor načíst pomocí zásuvného modulu. Následně je možné jednotlivé třídy a jejich obsah v zásuvném modulu zobrazit. Forma zobrazení je zde velmi přehledná a umožňuje snadnou kontrolu. Agregační pravidla se aplikují na *gpkg* soubor, jehož umístění je definováno v předchozí části „Download“. V daném souboru se vytvoří nové atributy, jejichž počet odpovídá počtu atributů popisující lc1 (land cover 1) třídy. To znamená, že v případě časoprostorově agregovaných dat vznikne pět nových atributů, které odpovídají rokům 2006, 2009, 2012, 2015 a 2018. V případě neagregovaných dat, u kterých každý záznam v atributové tabulce představuje unikátní měření se vytvoří pouze jeden nová atribut. Do nových sloupečků atributové tabulky se zapíšou nově definované hodnoty odpovídající klíčům v *json* souboru. Zápis je proveden pomocí SQL příkazu, který může vypadat následovně:

UPDATE lucas\_points SET lc1\_a = 'A10' WHERE lc1\_h IN ('A11', 'A12', 'A13')

Tento příklad popisuje aplikaci prvního agregačního pravidla zobrazeného na obrázku (Obr.:). Řetězec lucas\_points zde definuje název tabulky, lc1\_a odpovídá názvu nově vytvořeného atributu a lc1\_h atribut obsahující harmonizované land cover třídy.

# Závěr

Závěr……………………….

# Použité zdroje

[1] About Geo-harmonizer. *Open Data Science Europe* [online]. [cit. 2021-03-28]. Dostupné z: <https://opendatascience.eu/geoharmonizer-project/>

[2] LUCAS - Land use and land cover survey. *EUROSTAT Statistics Explained* [online]. [cit. 2021-03-28]. Dostupné z: https://ec.europa.eu/eurostat/statistics-explained/index.php/LUCAS\_-\_Land\_use\_and\_land\_cover\_survey

[3] D’ANDRIMONT, R., YORDANOV, M., MARTINEZ-SANCHEZ, L. et al. Harmonised LUCAS in-situ land cover and use database for field surveys from 2006 to 2018 in the European Union. *Scientific Data* [online]. 16.10. 2020, **7**(352 (2020) [cit. 2021-03-29]. Dostupné z: doi:https://doi.org/10.1038/s41597-020-00675-z

[4] *Land Use / Cover Area Frame Survey: Technical reference document C2* [online]. 2018 [cit. 2021-03-29]. Dostupné z: <https://ec.europa.eu/eurostat/documents/205002/8072634/LUCAS2018-C2-FieldForm-GD-Template.pdf>

[5] *Land Use / Cover Area Frame Survey: Technical reference document C3* [online]. 2018 [cit. 2021-03-29]. Dostupné z: <https://ec.europa.eu/eurostat/documents/205002/8072634/LUCAS2018-C3-Classification.pdf>

[6] BUCK, O., HAUB, C., WODITSCH, S., LINDEMANN, M. D., KLEINWILLINGHÖFER, L., HAZEU, G. W., KOSZTRA, B., KLEESCHULTE, S., ARNOLD, S., & HÖLZL, M. (2015). Analysis of the LUCAS nomenclature and proposal for adaptation of the nomenclature in view of its use by the Copernicus land monitoring services. EEA/EFTAS. Contract No 3436/B2015/RO-COPERNICUS/EEA.56195. . EEA - European Environment Agency. <https://land.copernicus.eu/user-corner/technical-library/LUCAS_Copernicus_Report_v2-2.pdf>

# Seznam obrázků

[Obr. 1: Rozložení LUCAS bodů měřených v roce 2018 10](file:///D:\Skola_zaloha\DP\Text\DP.docx#_Toc67931204)

[Obr. 2: Ukázka dotaníku LUCAS [4] 12](file:///D:\Skola_zaloha\DP\Text\DP.docx#_Toc67931205)

[Obr. 3: Land cover třídy [2] 13](file:///D:\Skola_zaloha\DP\Text\DP.docx#_Toc67931206)

[Obr. 4: Land use třídy [2] 16](file:///D:\Skola_zaloha\DP\Text\DP.docx#_Toc67931207)

[Obr. 5: Logo QGIS 25](file:///D:\Skola_zaloha\DP\Text\DP.docx#_Toc67931208)

[Obr. 6: Logo Python 26](#_Toc67931209)

[Obr. 7: Logo PostgreSQL 26](file:///D:\Skola_zaloha\DP\Text\DP.docx#_Toc67931210)

[Obr. 8: Logo GeoServer 27](file:///D:\Skola_zaloha\DP\Text\DP.docx#_Toc67931211)

# Seznam tabulek

[Tab. 1: Rozsah LUCAS 11](#_Toc67931277)

[Tab. 2: Počet LUCAS atributů 19](#_Toc67931278)

[Tab. 3: Změny názvů atributů 20](#_Toc67931279)

[Tab. 4: Změna hodnot atributu lc\_lu\_special\_remark 21](#_Toc67931280)

[Tab. 5: Hodnoty atributu lc1\_perc 22](#_Toc67931281)

[Tab. 6: Počet bodů atributu wm\_delivery v letech 23](#_Toc67931282)

[Tab. 7: Změny datových typů atributů 24](#_Toc67931283)

# Seznam zkratek

# Seznam příloh

**Digitální přílohy**