LESOTHO LAND COVER DATABASE - 2020

A new methodology for national land cover production based on sentinel-2

William Ouellette
Earth Observation Data Scientist, FAO

Food and Agriculture Organization of the United Nations

KHOTSO PULA NALA

PROJECT OVERVIEW

Objectives:

- Generate an **updated National Land Cover** product for 2020 using a new cost-effective methodology based on Sentinel-2 satellite data
- **Involve the national stakeholders** in its production and its use through participatory approaches and capacity development, to guarantee optimal usability
- **Collect Field Data** to validate the approach and ensure highest product quality

Impact:

- Support accurate reporting and decision-making for enhancing food security monitoring in Lesotho

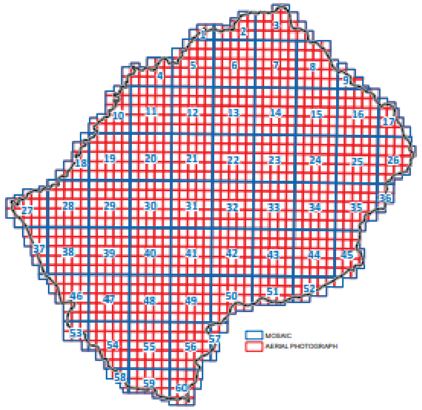
Food and Agriculture Organization of the United Nations

2015 LESOTHO LAND COVER DATABASE

Prepared in the framework of the FAO Emergency Program: **"Building Lesotho Resilience through the Upscale of Climate Smart Agriculture and Functional DRR Land Resources Information"**

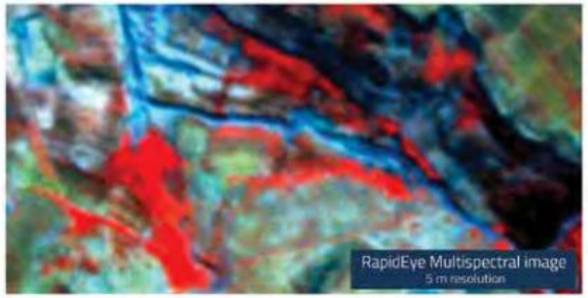
Developed in Collaboration between FAO, Bureau of Statistics (BOS), Ministry of Agriculture and Food Security (MAFS), Ministry of Forestry, Range and Soil Conservation (MFRSC)

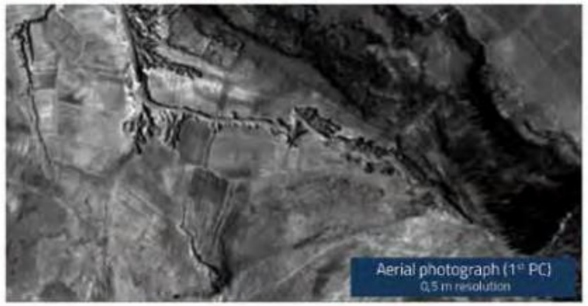
Ortho-photos from 2014 provided by BOS were essential in the production process

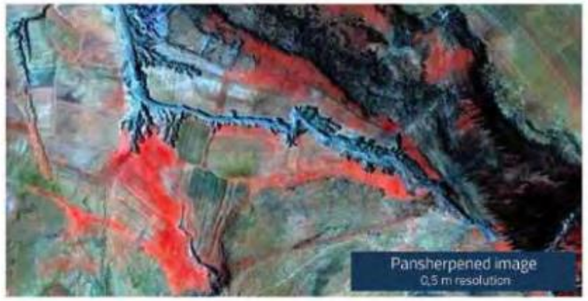


2015 LCDB METHODOLOGY

1. **Pan-sharpening** of Commercial 5m satellite imagery using Orthophotos



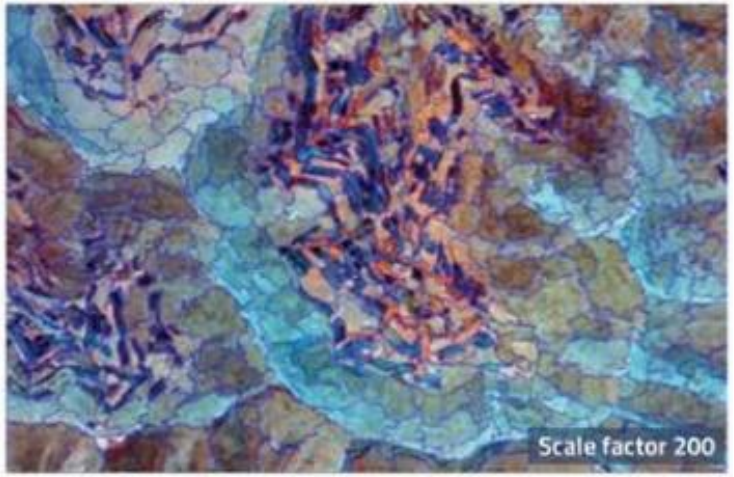


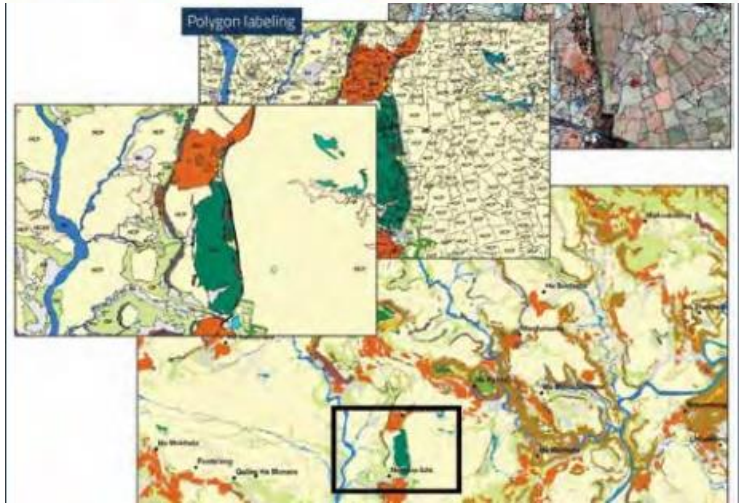


2. Custom **Land Cover Legend** was defined

Land Cover name	Code	Air photos	Pansharp	RapidEye & NDVI	Topo maps	Existing datasets	Ground truth
Urban	UA1						
Urban commercial Industrial	UA2						
Rural settlements Plan areas	RH1						
Rural settlements Sloping areas	RH2						
Rainfed agriculture Plain areas	HCP						
Rainfed agriculture Sloping regions	HCSM						
Rainfed agriculture Sheet erosion	HCER						
irrigated agriculture	HCIR						
Rainfed agriculture Rainfed orchards	HCT						
Trees, Needleaved (closed)	TNL1						
Trees, Needleaved (open)	TNL2						
Trees, Broadleaved (closed)	TBL1						
Trees, Broadleaved (open)	TBL2						
Trees, Undifferentiated (closed)	TM1						
Trees, Undifferentiated (closed)	TM2						
Trees (sparse)	TS						
Large waterbody	WB1						
Small waterbody	WB2						
Wetland (perennial or seasonal)	WET						
River bank	RB						
Shrubland (closed)	SH1						
Shrubland (open)	SH2						
Grassland	GR						
Grassland - Degraded	GRD						
Bare Rock	BR						
Bare area	BA						
Boulders & loose rocks	BLR						
Gullies	GU						
Mines & Quarries	MQ						

3. **Image Segmentation** and **photo-interpretation**





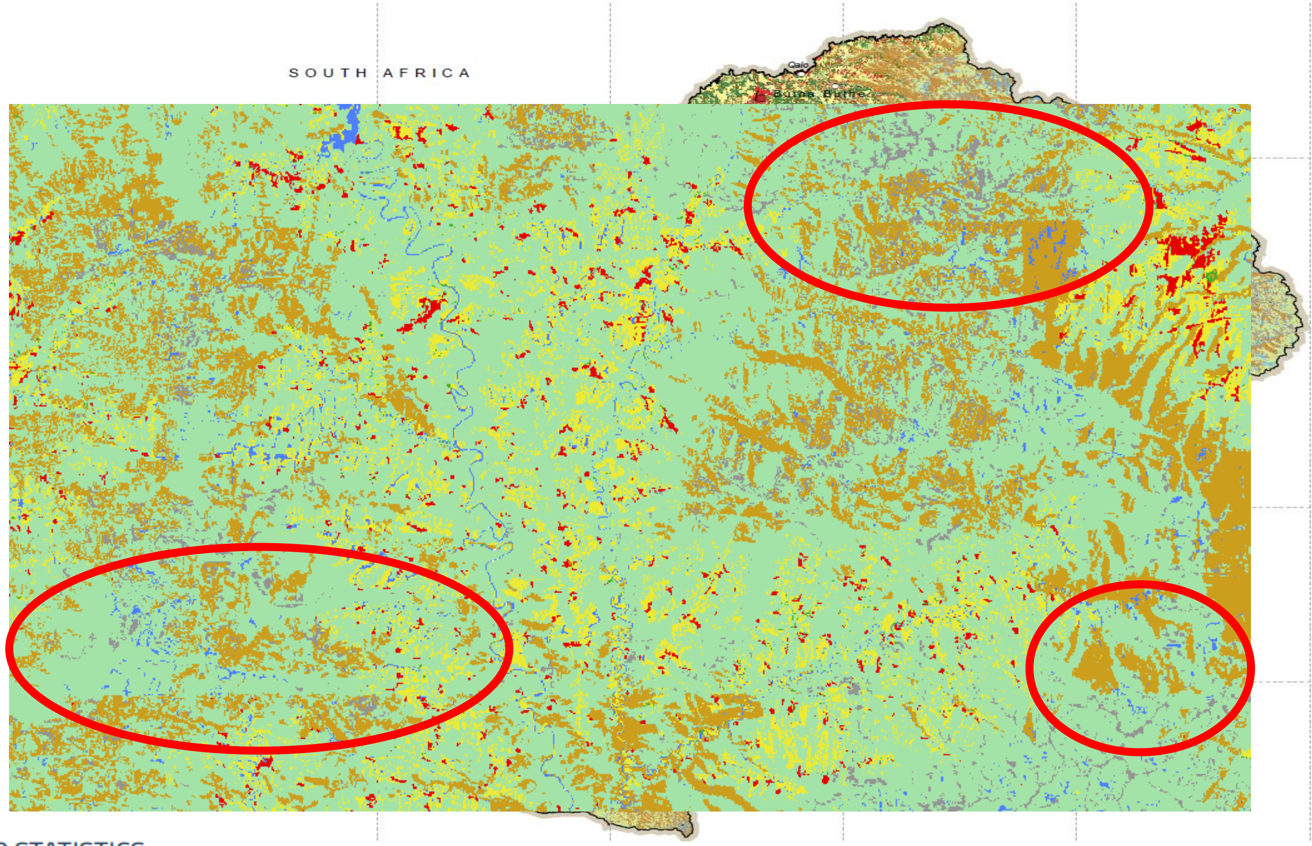
Food and Agriculture Organization of the United Nations

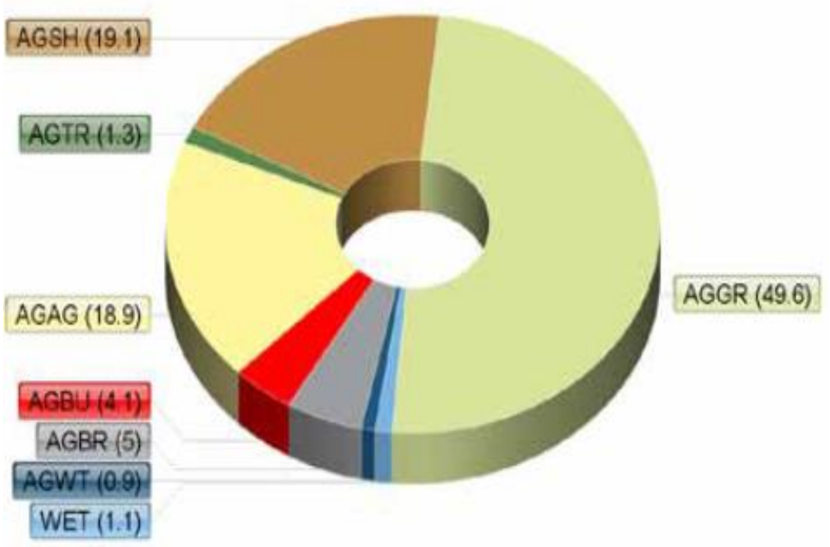
2015 LCDB RESULTS

The 2015 LCDB baseline is impressive and meticulous work which produce a continuous vector land cover output and associated zonal statistics (per catchment and per administrative areas)

In spite of the product quality, two main issues are identified in the product:

- The manual labelling process induced inconsistencies and bias in the output, partly due to the fact that mono-temporal imagery was used as photo-interpretation layer
- The land cover classes defined are not "machine-learning optimized" ("open" classes cannot effectively be detected using machine learning)





AGGREGATED LAND COVER STATISTICS

DISTRICTS	Built-up	Agricultural Land	Trees	Shrubland	Grassland	Wetland	Water Bodies & Rivers	Barrenland	TOTAL LAND
	AGBU	AGAG	AGTR	AGSH	AGGR	WET	AGWT	AGBR	
	UA1, UA2, RH1, RH2	HCP, HCSM, HCER, HCIR, HCT	TNL1, TNL2, TBL1, TBL2, TM1, TM2, TS	SH1, SH2	GR, GRD	WET	WB1, WB2, RB	BR, BA, GU, BLR, MQ	
Butha-Buthe	5,787	22,792	2,235	46,913	86,064	2,979	1,082	10,933	**178,785**
Leribe	17,087	83,711	8,574	44,201	111,033	1,531	4,291	12,132	**282,559**
Berea	15,086	75,212	7,741	21,376	67,683	339	1,162	9,008	**197,606**
Maseru	29,073	89,401	7,115	55,775	195,842	2,204	4,023	15,714	**399,146**
Mafeteng	18,370	95,534	2,410	16,887	63,981	384	2,218	14,847	**214,641**
Mohale's Hoek	14,042	71,381	2,952	79,862	173,386	2,430	4,190	20,861	**369,102**
Quthing	10,942	34,997	3,935	84,388	146,572	2,157	2,508	11,675	**297,174**
Qacha's Nek	4,557	24,015	1,796	43,028	118,774	3,068	1,802	15,728	**212,768**
Mokhotlong	3,644	32,857	595	106,887	235,973	12,826	2,631	22,132	**417,544**
Thaba-Tseka	7,504	48,129	1,051	85,012	316,745	4,662	4,335	18,550	**485,989**
TOTAL (ha)	**126,091**	**578,039**	**38,404**	**584,328**	**1,516,051**	**32,580**	**28,241**	**151,581**	**3,055,314**
TOTAL (%)	**4.1**	**18.9**	**1.3**	**19.1**	**49.6**	**1.1**	**0.9**	**5.0**	**100**

METHODOLOGY CHANGE IMPACT (1)

Characteristic	LCDB 2015	LCDB 2020
Sensor	RapidEye/Orthophotos	Sentinel-2
Spatial Resolution	5-0.5 m	10 m
Temporal Resolution	Single acquisition of each sensor	5-days revisit Generation of 6 2-months temporal composite from Sep 2019-Aug 2020
Classification Approach	Manual labelling	Random Forest Pixel-based Classifier
Constraints	No prior data available	No Field Survey possible

LCS3code	Area Statistics			
	10m Area (km^2)	1.5/2m Area (km^2)	Diff (km^2)	Diff (%)
UA1	198,239	198,198	0,0406	**0,020%**
UA2	23,377	23,373	0,0044	**0,019%**
RH1	782,582	782,556	0,0264	**0,003%**
RH2	256,792	256,781	0,0109	**0,004%**
HCP	3534,227	3534,290	-0,0637	**-0,002%**
HCSM	2067,193	2067,079	0,1142	**0,006%**
HCER	172,197	172,430	-0,2330	**-0,135%**
HCIR	1,242	1,244	-0,0018	**-0,143%**
HCT	5,520	5,520	0,0003	**0,006%**
TNL1	13,733	13,739	-0,0066	**-0,048%**
TNL2	2,927	2,927	0,0000	**-0,001%**
TBL1	26,902	26,899	0,0032	**0,012%**
TBL2	1,734	1,728	0,0062	**0,359%**
TM1	278,929	278,986	-0,0567	**-0,020%**
TM2	31,092	31,092	0,0004	**0,001%**
TS	28,666	28,665	0,0015	**0,005%**
WB1	46,446	46,501	-0,0550	**-0,118%**
WB2	4,795	4,795	0,0005	**0,010%**
WET	325,848	325,795	0,0532	**0,016%**
RB	230,591	231,115	-0,5242	**-0,227%**
SH1	1492,352	1492,489	-0,1374	**-0,009%**
SH2	4350,596	4350,796	-0,2002	**-0,005%**
GR	11758,549	11758,605	-0,0555	**0,000%**
GRD	3401,492	3401,906	-0,4143	**-0,012%**
BR	444,318	444,375	-0,0566	**-0,013%**
BA	824,884	824,975	-0,0909	**-0,011%**
BLR	66,888	66,879	0,0091	**0,014%**
GU	170,636	170,554	0,0827	**0,048%**
MQ	9,027	9,025	0,0017	**0,018%**
TOTAL	30551,775	30553,316	2,2510	**0,007%**

LCS3code	Object Size Distribution	
	Area < 0.25 Ha (as % total)	Area < 0.36 Ha (as % total)
UA1	0,14%	**0,19%**
UA2	1,58%	**2,70%**
RH1	0,62%	**1,04%**
RH2	1,86%	**2,99%**
HCP	0,26%	**0,45%**
HCSM	2,10%	**3,49%**
HCER	0,23%	**0,47%**
HCIR	0,46%	**0,46%**
HCT	2,29%	**4,03%**
TNL1	3,99%	**5,88%**
TNL2	2,53%	**4,10%**
TBL1	9,85%	**14,21%**
TBL2	8,43%	**13,45%**
TM1	16,85%	**23,32%**
TM2	7,62%	**12,54%**
TS	2,91%	**5,74%**
WB1	0,02%	**0,03%**
WB2	13,18%	**18,08%**
WET	1,83%	**3,26%**
RB	5,42%	**8,04%**
SH1	1,75%	**2,93%**
SH2	0,46%	**0,82%**
GR	0,62%	**0,96%**
GRD	0,36%	**0,61%**
BR	4,43%	**6,94%**
BA	3,79%	**6,03%**
BLR	1,82%	**3,19%**
GU	6,77%	**10,27%**
MQ	0,05%	**0,13%**
TOTAL	1,08%	**1,72%**

Sensitivity analysis of changing spatial resolution showed that the uncertainty linked to resampling to lower resolution was negligible (<2% overall area change)

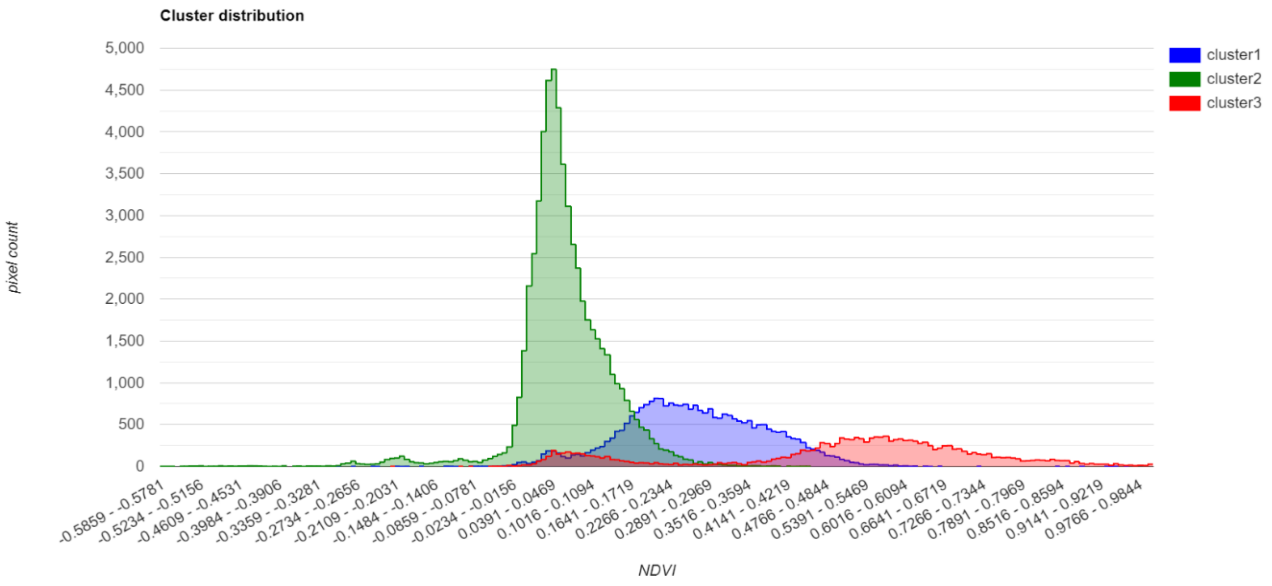
- This analysis took the potential area loss/gain of each land cover class due to resampling, as well as a Minimum Mapping Unit of 0.36 Ha (which is realistic when working with Sentinel-2)
- The spectral consistency analysis of labelled classes showed

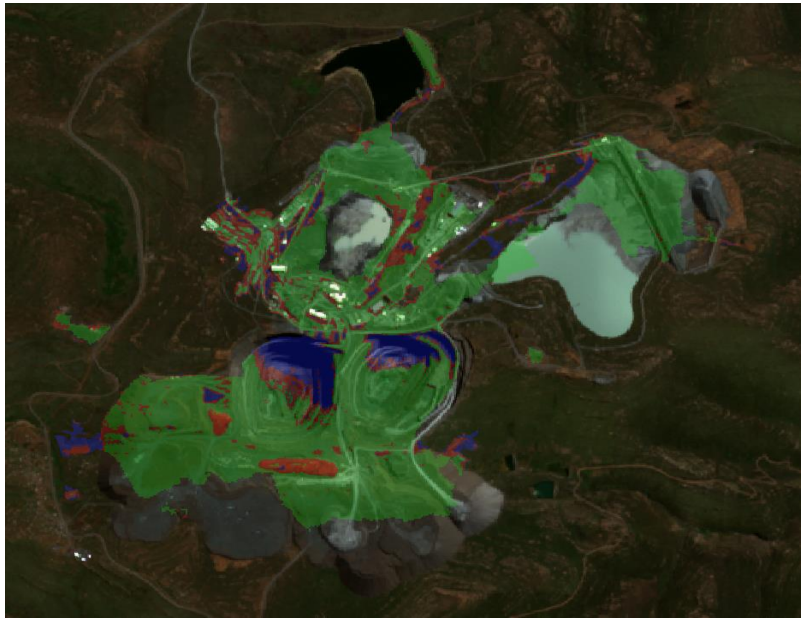
METHODOLOGY CHANGE IMPACT (2)

The spectral consistency analysis of labelled classes highlighted the multi-modality of each land cover class as per LCDB 2015 labels

- While it is expected that cropland classes would be multi-modal (different crop types), most classes are expected not to be, yet they were
- Multi-modal classes which should not be run the risk of overlapping with other classes, and decreasing overall inter-class separability

The labels cannot be used as such for a straight-forward 2020 LCDB classification





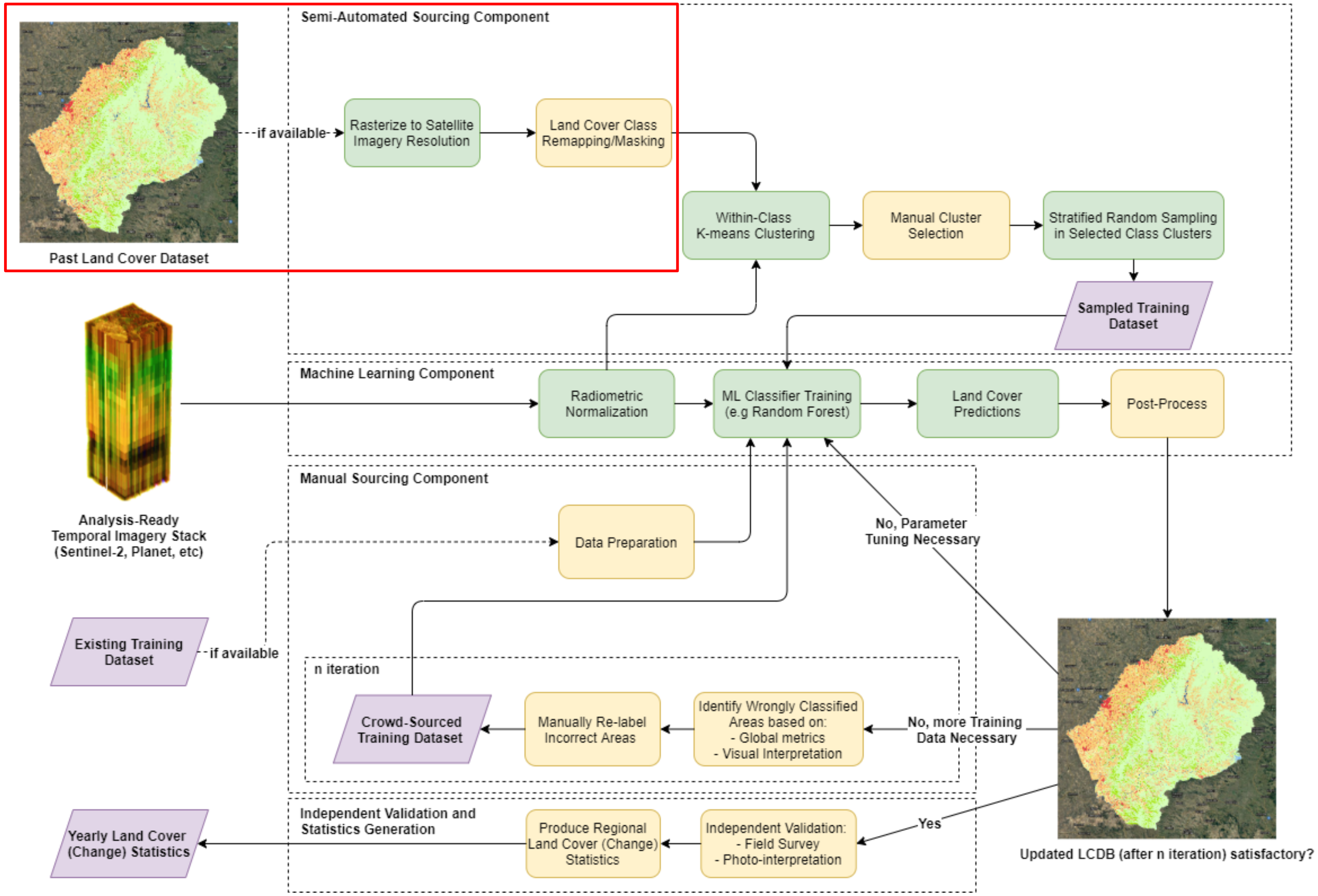
PROCESSING WORKFLOW

This Processing Workflow can be applied to any new Area/Country of Interest

If prior Land Cover dataset available, can be leveraged to semi-automatically source pixel training labels still relevant in the new year of land cover production

- The same method could be used to source objects rather than pixels for OBIA approach

Leaves room for iterating based on global metrics and visual inspection of results, either through parameter tuning (software fix) or through manual capturing of training data (data fix)



REDUCED LAND COVER NOMENCLATURE

Classes Removed from classification due to heterogeneous/fuzzy nature:

LCS3Name	LCS3Code	pixel count (10m)	Surface Area (km²)	Percentage total area
Rainfed Agriculture - Rainfed Orchards	HCT	55203	5,5203	**0,02%**
Trees, Needleleaved (open)	TNL2	29273	2,9273	**0,01%**
Trees, Broadleaved (open)	TBL2	17339	1,7339	**0,01%**
Trees, Undifferentiated (closed)	TM1	2789227	278,9227	**0,92%**
Trees, undifferentiated (open)	TM2	310921	31,0921	**0,10%**
Trees (sparse)	TS	286691	28,6691	**0,09%**
Shrubland (open)	SH2	43508744	4350,8744	**14,33%**
Total		57391958	5739,1958	**18,91%**

Pixel-based classifiers cannot adequately handle heterogeneous land cover classes (that contain a mixture of multiple land cover classes)

- E.g. Open shrubland can be a mixture of anywhere between 10-90% of grassland and shrubland
- Even object-based methods have performed poorly to classify fuzzy land cover classes

The following classes were merged due to their overlapping class definitions, once again to minimize fuzziness between classes:

- Bare Rock (BR), Bare Area (BA) and Boulders and Rocks (BLR) as **Bare Surfaces**
- Plain (HCP), slopes (HCSM) merged as **Rainfed Croplands**
- Urban (UA1) and Industrial (UA2) settlements **as Urban**
- Small (WB1) and Big (WB2) Water Bodies merged as single **Water** class

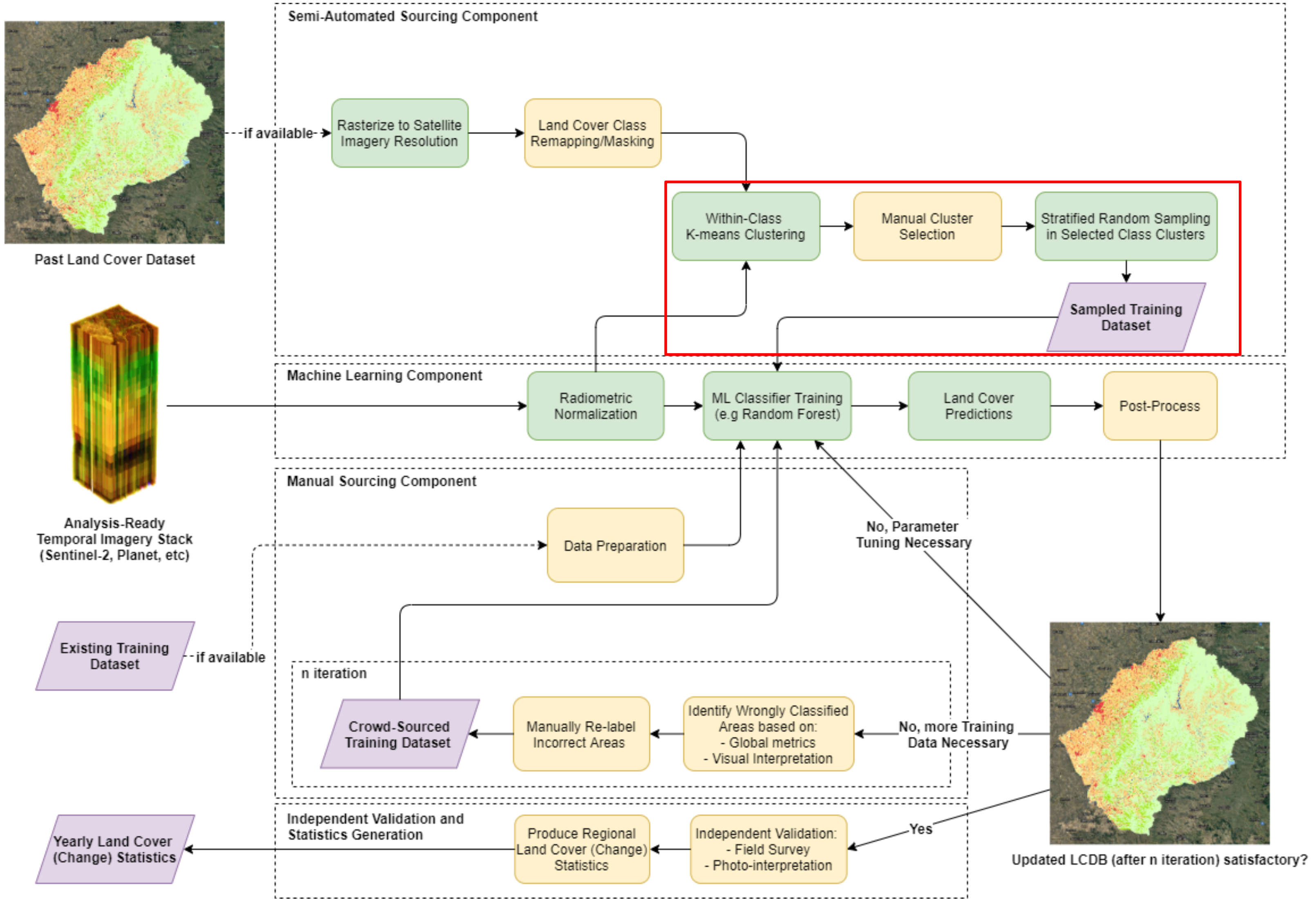
PROCESSING WORKFLOW

This Processing Workflow can be applied to any new Area/Country of Interest

If prior Land Cover dataset available, can be leveraged to semi-automatically source pixel training labels still relevant in the new year of land cover production

- The same method could be used to source objects rather than pixels for OBIA approach

Leaves room for iterating based on global metrics and visual inspection of results, either through parameter tuning (software fix) or through manual capturing of training data (data fix)



SEMI-AUTOMATED SOURCING COMPONENT

Adapted from the methodology from Paris et al., 2019, 2020 for unsupervised updating of land-cover maps using multispectral satellite time series

Within-Class Clustering (K-means or Gaussian Mixture Model), using agro-ecological zones as stratification layer, to extract cluster/distribution most representative of the given land cover class

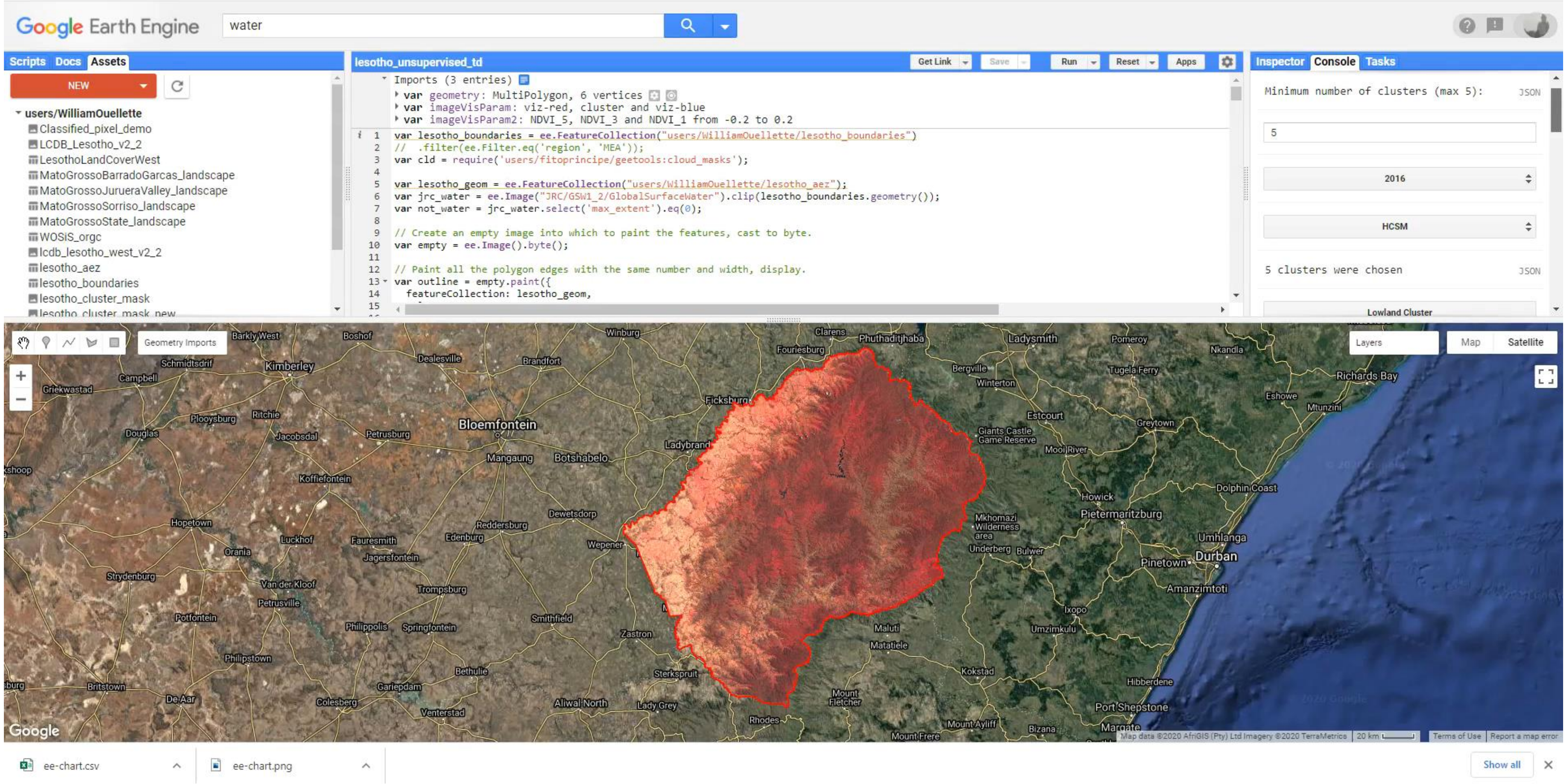
- Clustering is not done at object level like in Paris et al. 2020, but at class level with AEZ stratification for computational efficiency

Advantage of using GMM is that a score is produced per sample, which can be further used to perform "smart" sub-sampling Instead of purely random sampling

(a) (b) (c) (d)

Fig. 2: Qualitative example of polygon k-means clustering result: (a) original polygon associated to the "crop" label, (b) dominant land-cover class detected $\mathbf{C}_1$, (c) first minor class detected $\mathbf{C}_2$ (road), and (d) second minor class detected $\mathbf{C}_3$ (river).

SEMI-AUTOMATED TRAINING DATA SOURCING WITH GEE



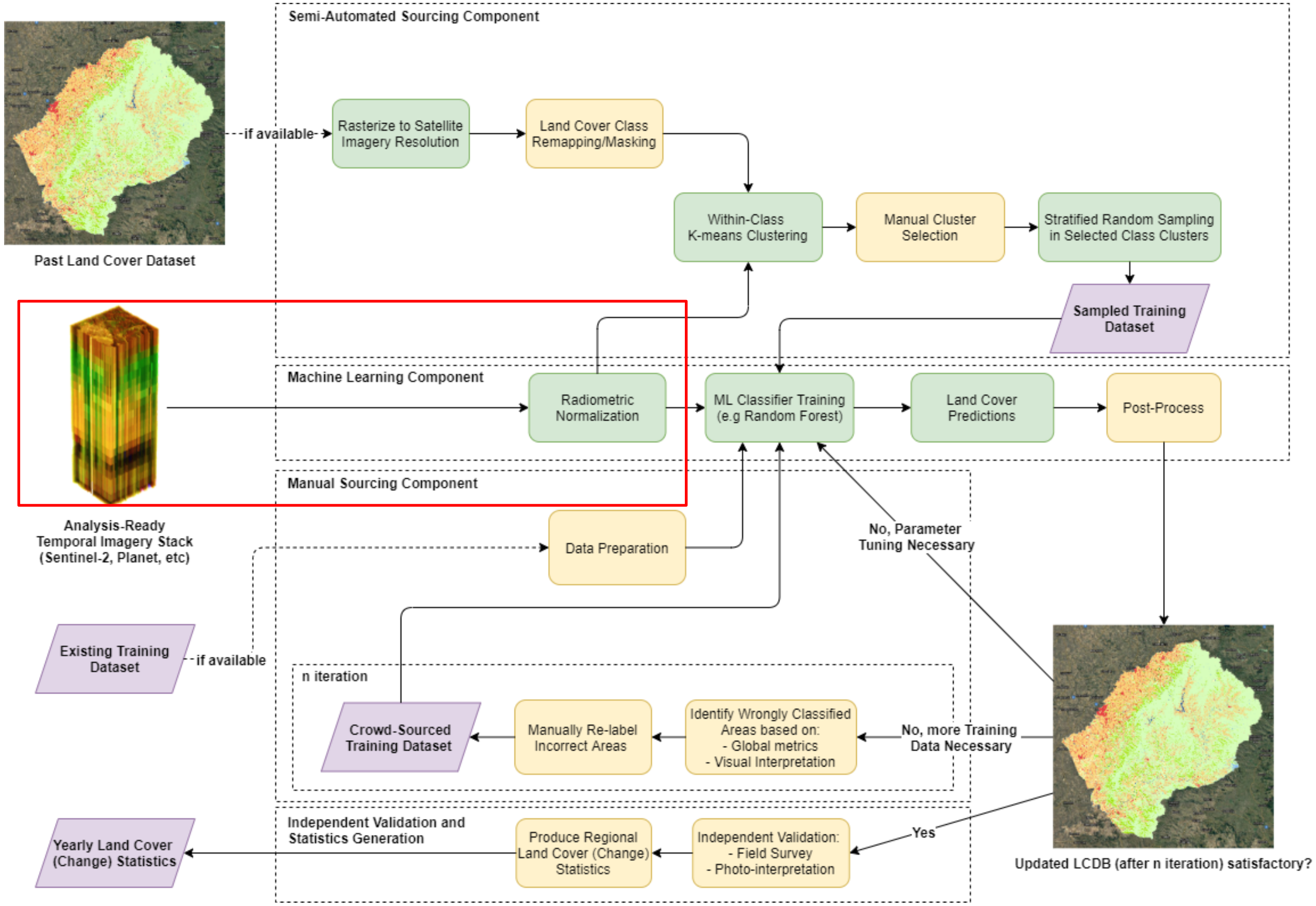
PROCESSING WORKFLOW

This Processing Workflow can be applied to any new Area/Country of Interest

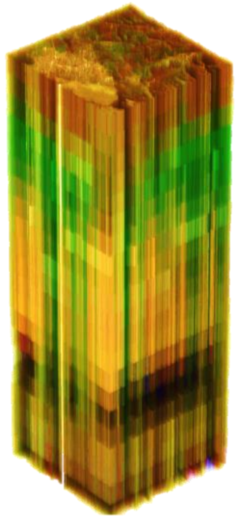
If prior Land Cover dataset available, can be leveraged to semi-automatically source pixel training labels still relevant in the new year of land cover production

- The same method could be used to source objects rather than pixels for OBIA approach

Leaves room for iterating based on global metrics and visual inspection of results, either through parameter tuning (software fix) or through manual capturing of training data (data fix)

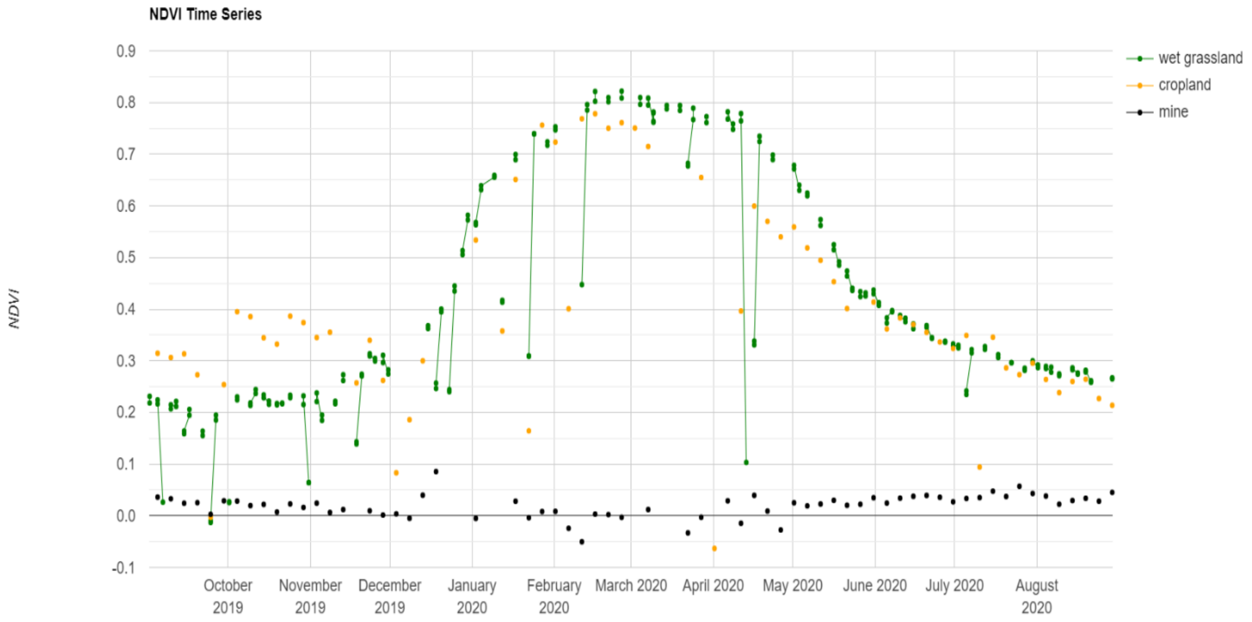


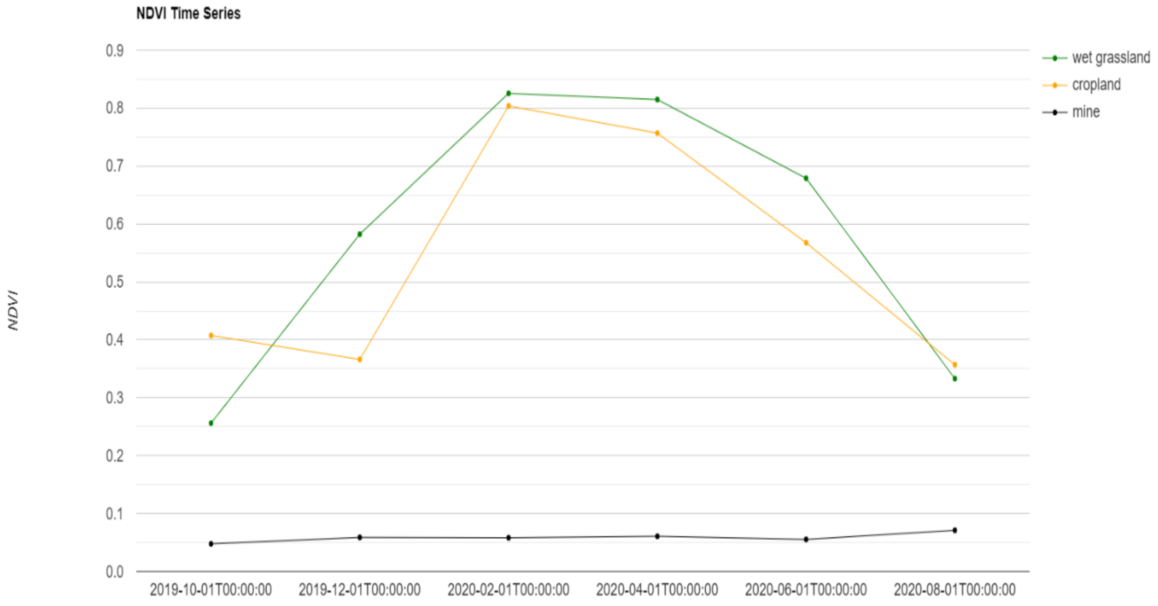
INPUT FEATURES



Input Features Generation: 6 * 2-months, radiometrically normalized, cloud-masked, Sentinel-2 **Max-NDVI** temporal composites

- All 10 and 20m bands + NDVI + GLCM Correlation and Constrast of 10m bands
- **Goals:** reduce data size, keep only cloud-free observations at key phenological stages of the year





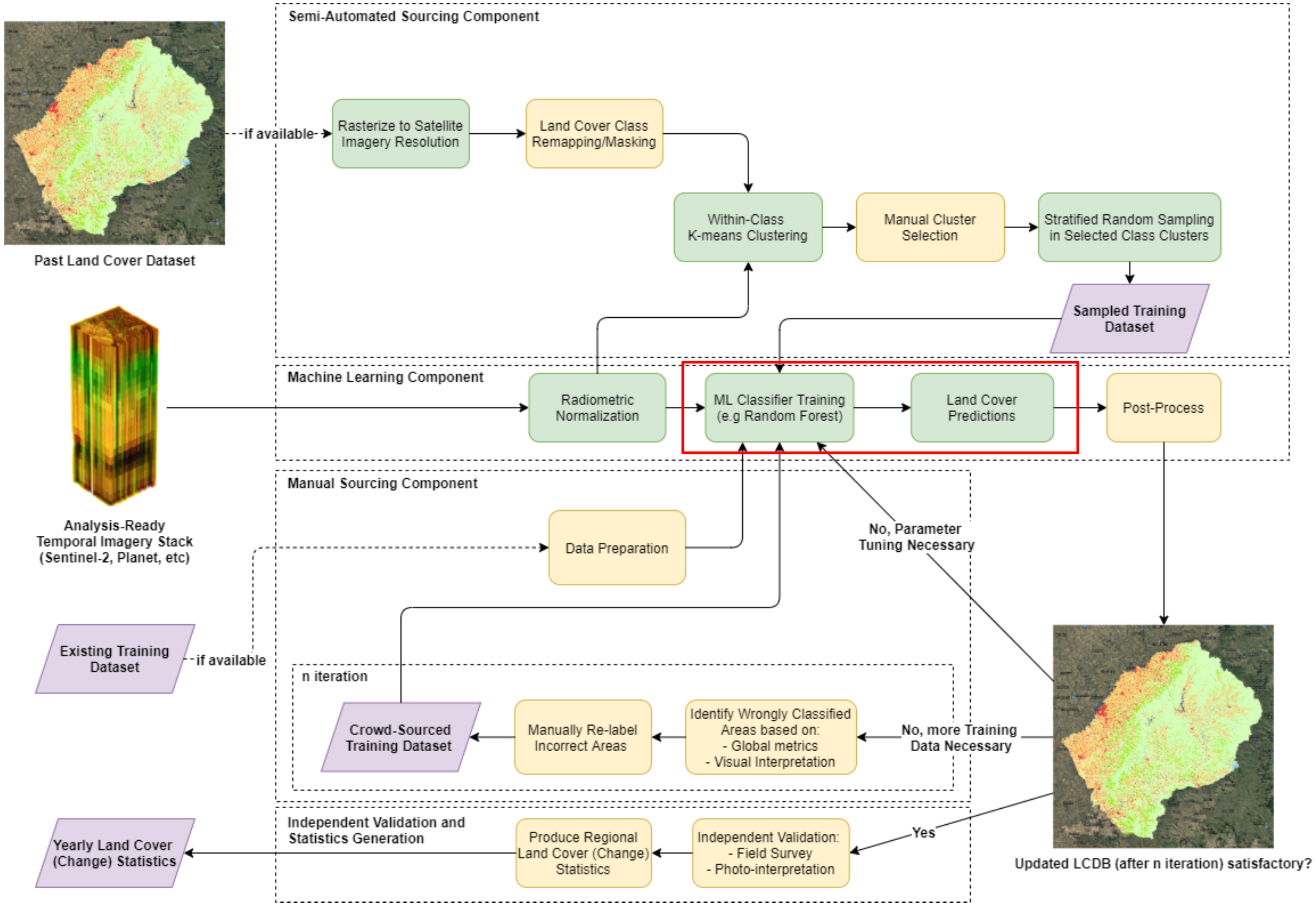
PROCESSING WORKFLOW

This Processing Workflow can be applied to any new Area/Country of Interest

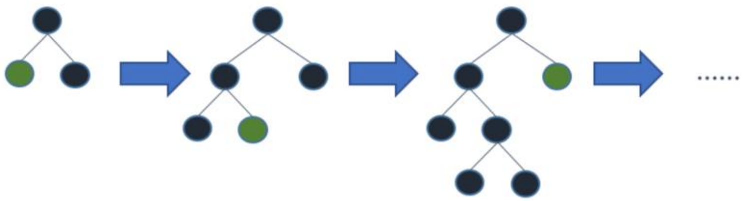
If prior Land Cover dataset available, can be leveraged to semi-automatically source pixel training labels still relevant in the new year of land cover production

- The same method could be used to source objects rather than pixels for OBIA approach

Leaves room for iterating based on global metrics and visual inspection of results, either through parameter tuning (software fix) or through manual capturing of training data (data fix)

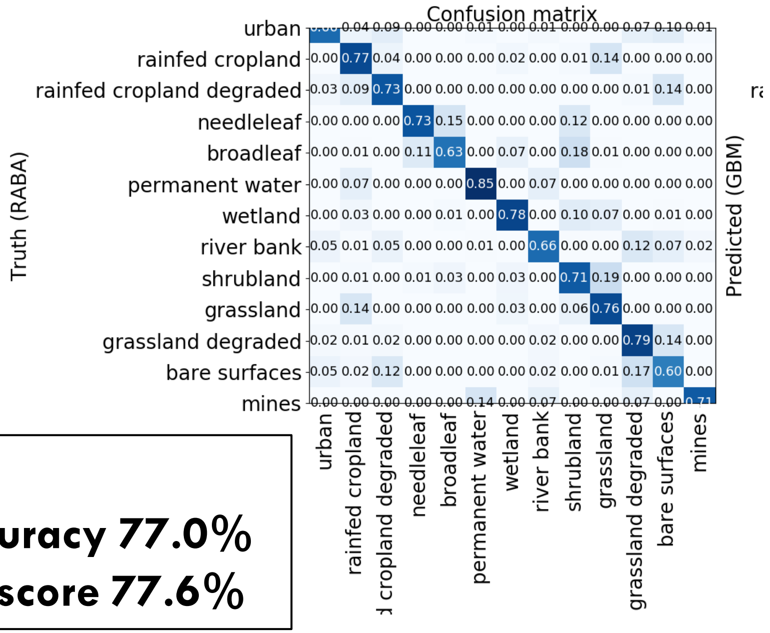


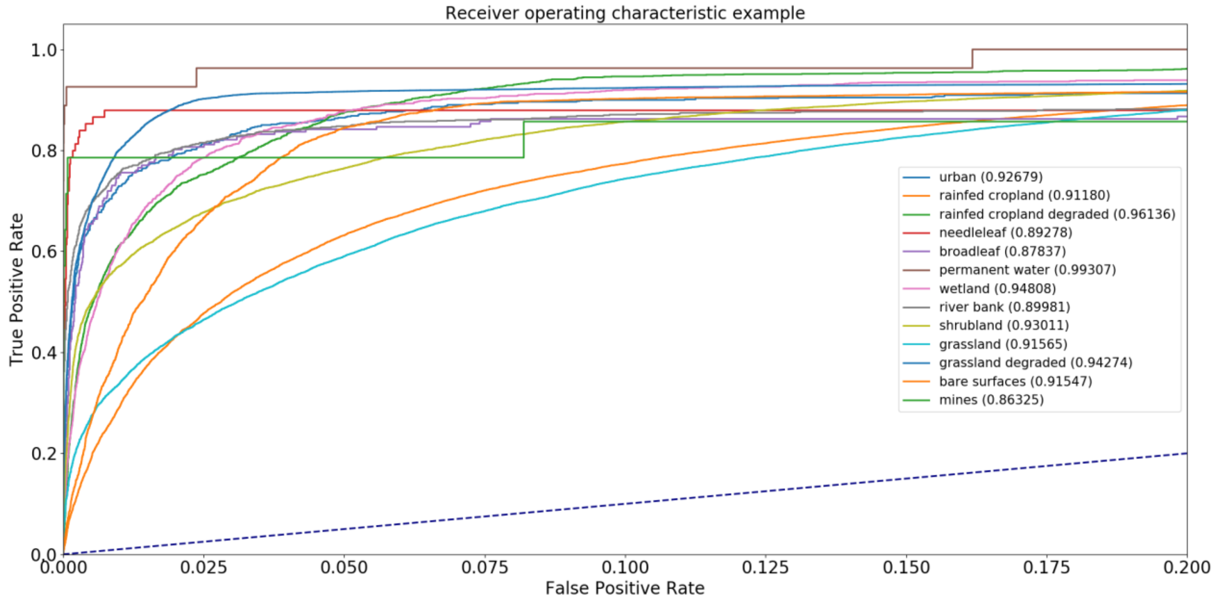
MACHINE LEARNING MODEL: RANDOM FOREST



Pixel-based Random Forest Ensemble Implementation in LightGBM with parameters:

- 200 trees due to large number of predictors (ensures all are used)
- L2 regularization with 5-fold cross-validation to avoid overfitting
- Over- and under-sampling to ensure no class overpowers the training data set by > 20% of total data





Test Results:
Classification accuracy 77.0%
Classification F1-score 77.6%

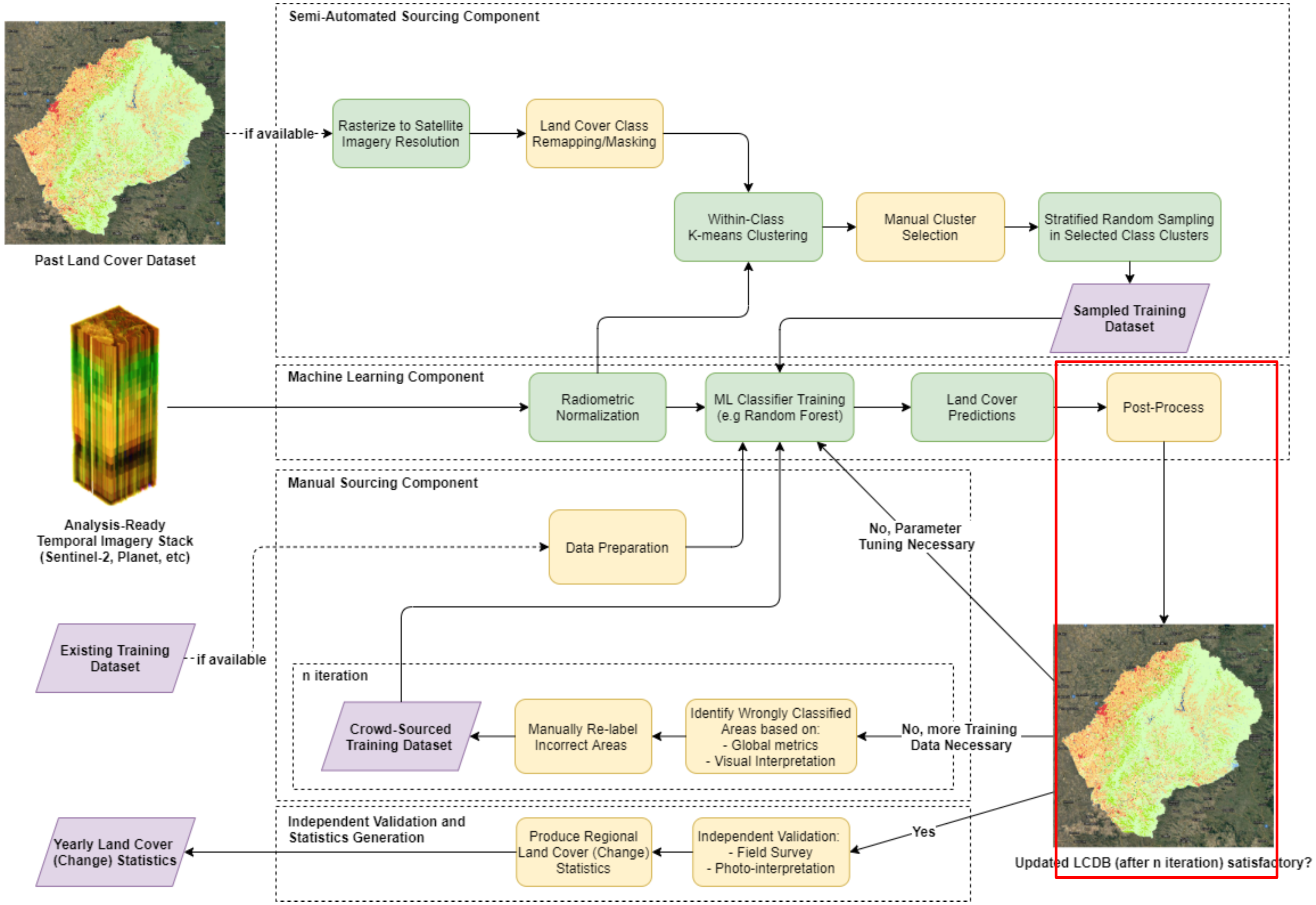
PROCESSING WORKFLOW

This Processing Workflow can be applied to any new Area/Country of Interest

If prior Land Cover dataset available, can be leveraged to semi-automatically source pixel training labels still relevant in the new year of land cover production

- The same method could be used to source objects rather than pixels for OBIA approach

Leaves room for iterating based on global metrics and visual inspection of results, either through parameter tuning (software fix) or through manual capturing of training data (data fix)



POST-PROCESSING

Sieving of 25 connected pixels (0,25 Ha)

Majority filter with disk radius of 1 pixel (10m)

Rainfed cropland **confidence >65%** in **Mountain** Agro-Ecological Zone

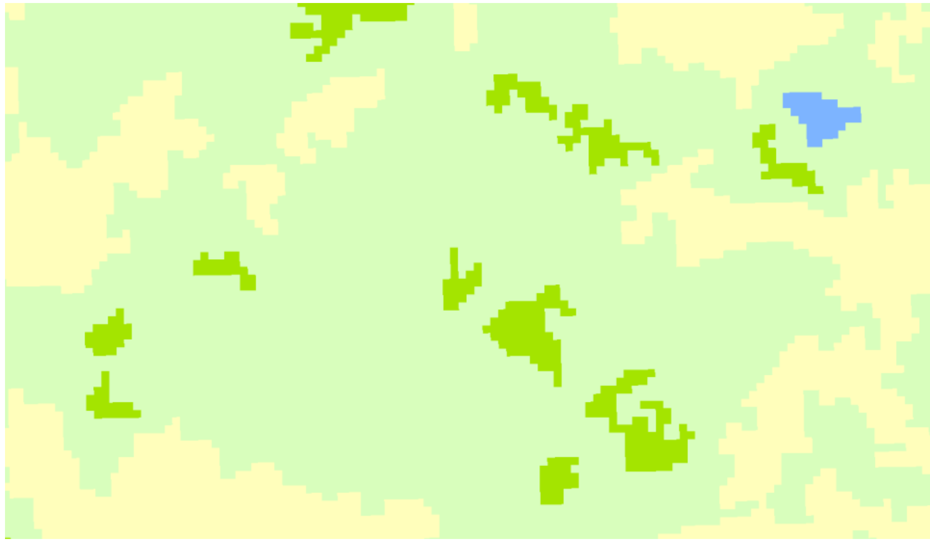
- Model over-estimated rainfed cropland extent in that AEZ

Removal of water and wetland class occurrence on **steep slopes** (>50°)

Harmonized rainfed cropland class with **OSM farmland tag**

Reintroduction of following classes from 2015, assuming they will remain in 2020, and because they are narrow features difficult to detect with Sentinel-2 (10m):

- **Gullies (GU)**
- **River banks (RB)**
- **Urban areas (UA1, UA2, RH1, RH2)**



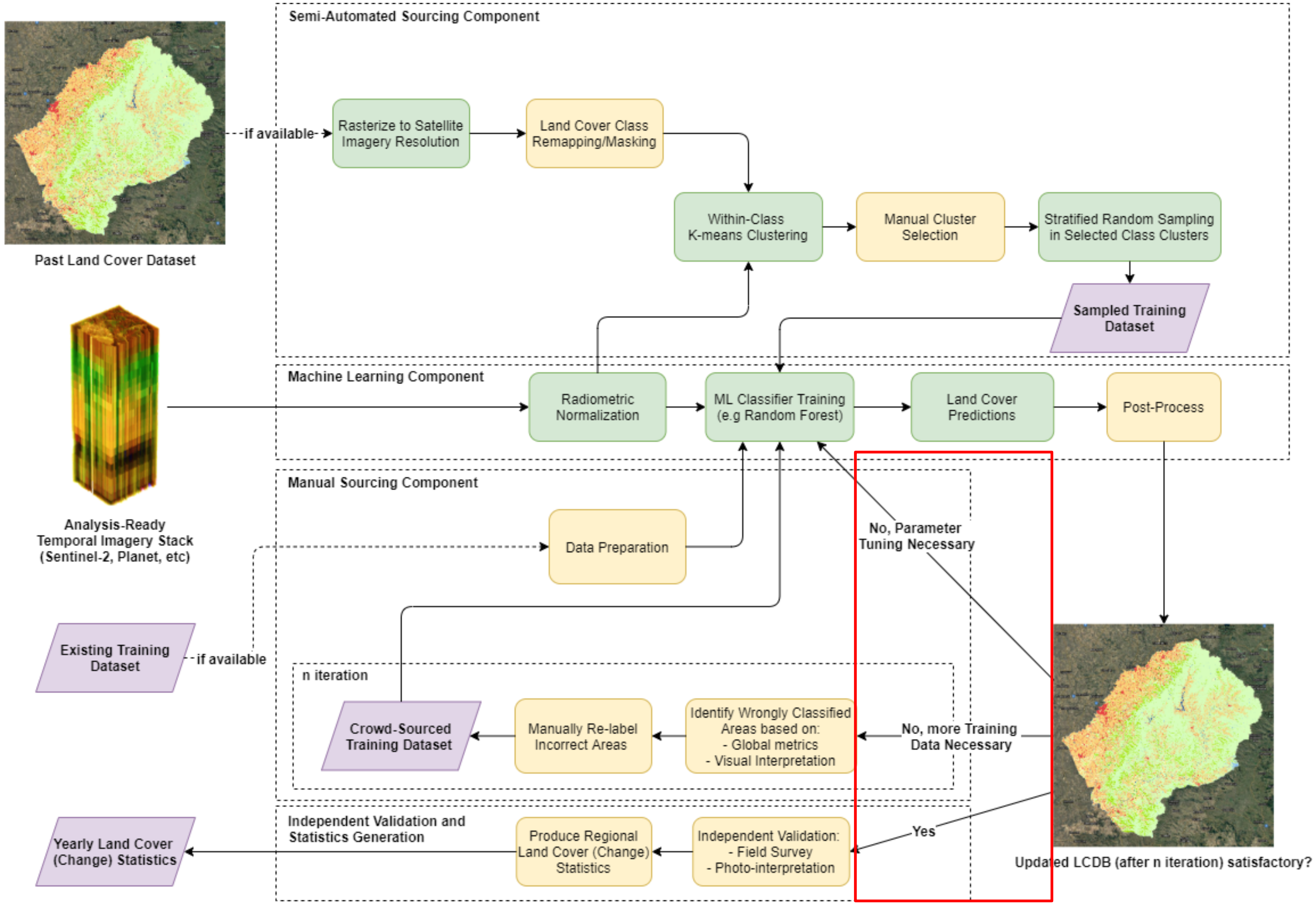
PROCESSING WORKFLOW

This Processing Workflow can be applied to any new Area/Country of Interest

If prior Land Cover dataset available, can be leveraged to semi-automatically source pixel training labels still relevant in the new year of land cover production

- The same method could be used to source objects rather than pixels for OBIA approach

Leaves room for iterating based on global metrics and visual inspection of results, either through parameter tuning (software fix) or through manual capturing of training data (data fix)



LCDB PRODUCTION: AN ITERATIVE PROCESS

Requires expert knowledge to iterate either by **parameter tuning** (software tuning) or by deciding to **add additional training data** (data tuning)

This is typically done through:

- Interpretation of test global metrics (training data distribution, ROC-AUC curve, confusion matrix)
- Visual interpretation of results and associated class confidences

As many iterations can be gone through until output is **visually consistent/satisfactory**

- Semantic alignment with historical LCDB is an important metric to keep an eye out for, but in the case of Lesotho, the methodology difference between LCDB 2015 (manual at 0.5/2m resolution) and LCDB 2020 (machine-learning at 10m resolution) makes their alignment difficult
- An independent validation still needs to be carried out on the deemed "final output" (see slide 24)

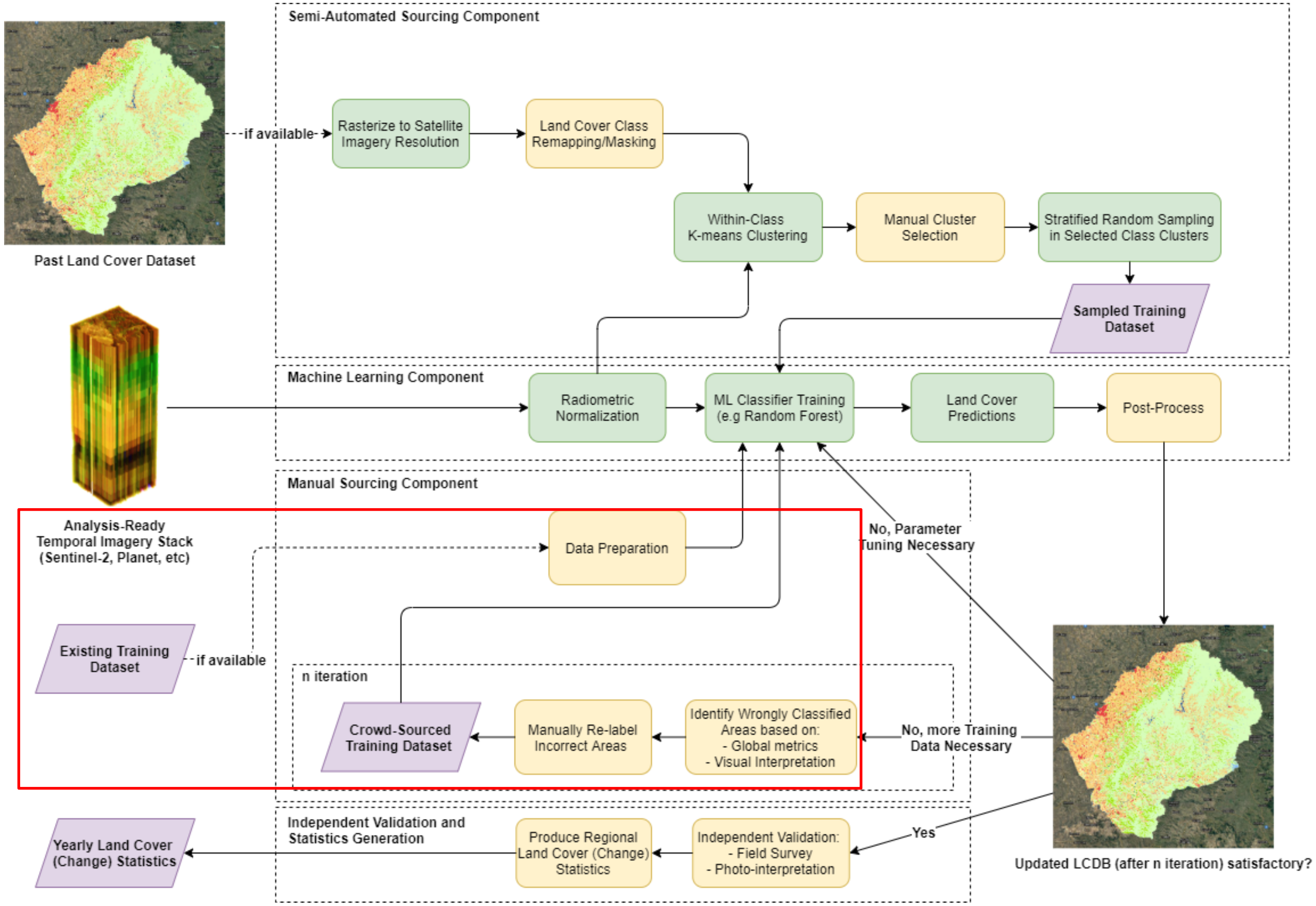
PROCESSING WORKFLOW

This Processing Workflow can be applied to any new Area/Country of Interest

If prior Land Cover dataset available, can be leveraged to semi-automatically source pixel training labels still relevant in the new year of land cover production

- The same method could be used to source objects rather than pixels for OBIA approach

Leaves room for iterating based on global metrics and visual inspection of results, either through parameter tuning (software fix) or through manual capturing of training data (data fix)



MANUAL SOURCING COMPONENT

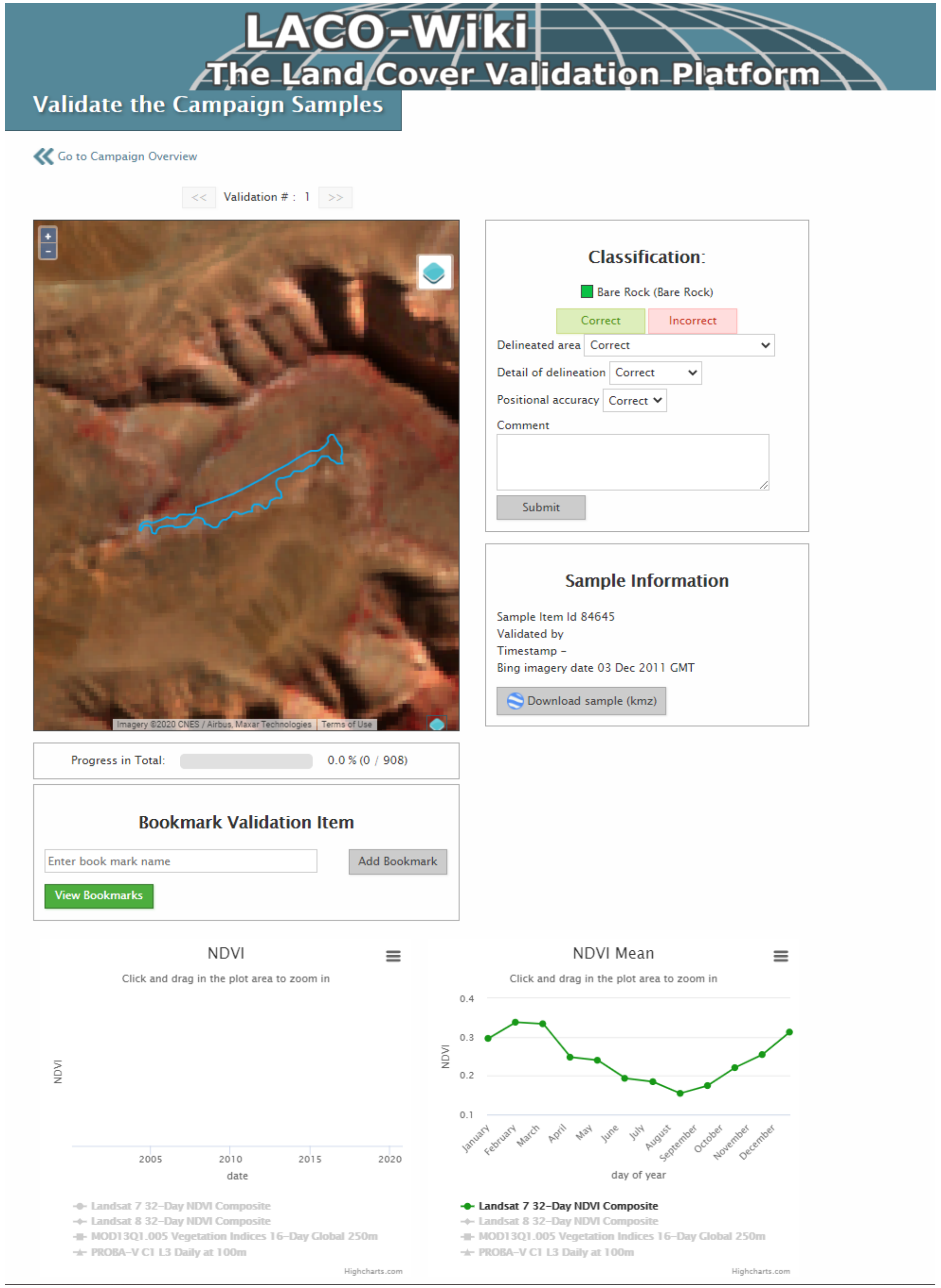
LACO-WIKI is an open-source Web/Smartphone-based crowd-sourcing tools to capture land cover data

- Many alternatives available, but LACO-WIKI has the simplest interface while still offering fitness-for-purpose

Our experiment of sourcing additional training data through LACO-WIKI **did not improve** the results with respect to the semi-automated sourcing approach

- One limitation may have been that polygons from LCDB 2015 were provided to the validators, without the possibility of editing them
- Collecting accurate training data for a full land cover classification with 15 classes is a non-trivial crowd-sourcing exercise
- The quality of the pixel labels sourced from the semi-automated sourcing approach are of high representative quality already

Capturing labels in a GIS software environment is also a viable alternative which offers possibility to draw/edit polygons in areas where land cover is wrongly classified



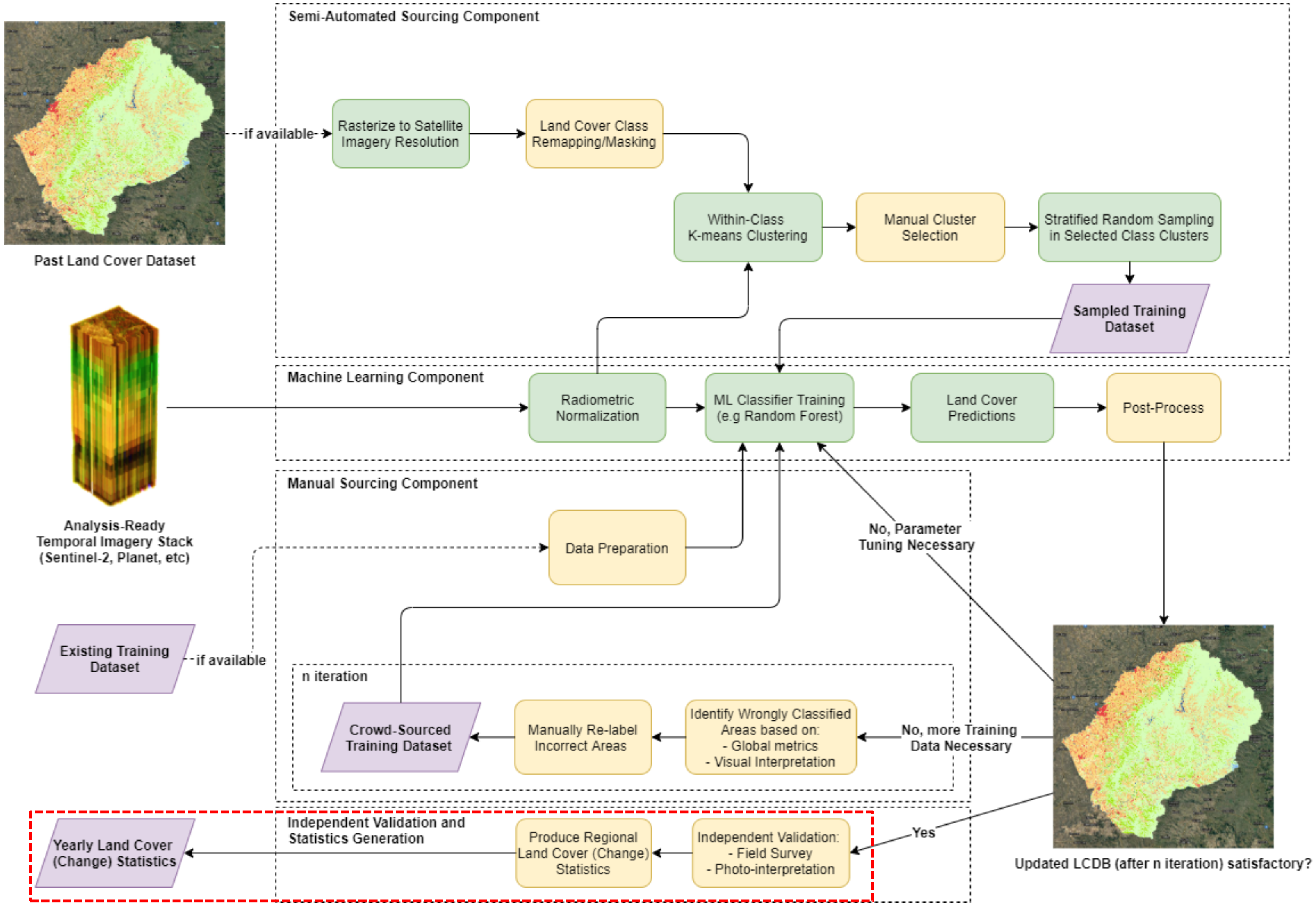
PROCESSING WORKFLOW

This Processing Workflow can be applied to any new Area/Country of Interest

If prior Land Cover dataset available, can be leveraged to semi-automatically source pixel training labels still relevant in the new year of land cover production

- The same method could be used to source objects rather than pixels for OBIA approach

Leaves room for iterating based on global metrics and visual inspection of results, either through parameter tuning (software fix) or through manual capturing of training data (data fix)



THE MISSING COMPONENT: INDEPENDENT VALIDATION

The model was trained and tested with a sub-set of the labels extracted from the semi-automated sourcing approach (see slide 15)

However, an independent validation is still necessary, with an uncorrelated dataset, to get a proper **user-centric accuracy assessment** of the product

- Stratified random sampling of 1300 pixels across all land cover classes and AEZ uploaded to a new LACO-WIKI campaign to generate first user-centric accuracy figures:

https://laco-wiki.net/c/lcdb2020_ind_val

- Once the COVID situation will allow for a field survey, this should be carried out additionally to get a more reliable sense of the LCDB 2020 quality, because certain classes are difficult to reliably identify through photo-interpretation

The **survey protocol** for the independent validation will be drafted as part of final report and **sampling locations** will be provided for execution

PRELIMINARY RESULTS

Settlement classes carried over from LCDB 2015 + new detections

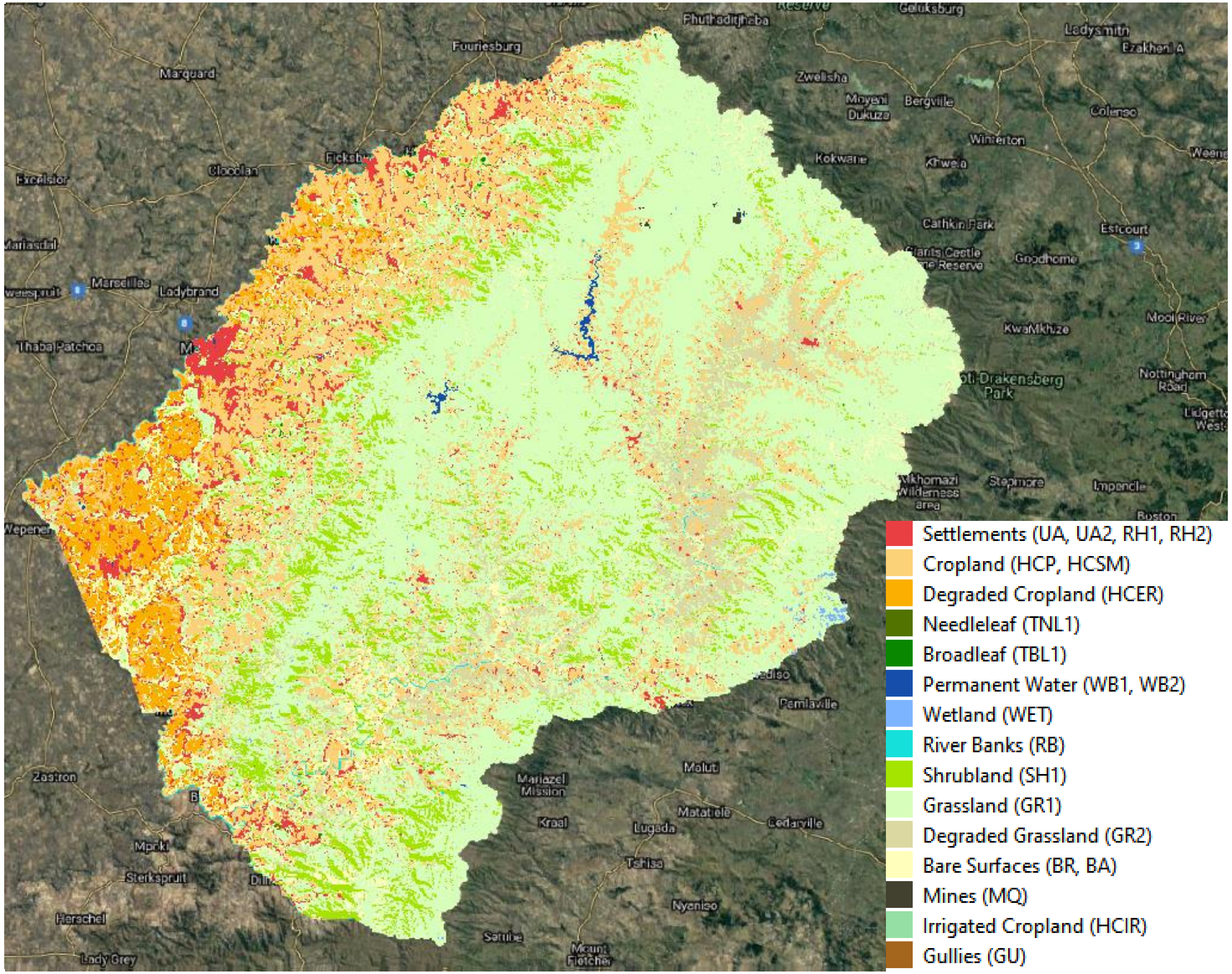
- Split between Urban and Rural Settlements could be made if necessary

Plain and **Slope/Mountain** cropland were merged because classifier could not effectively tell them apart

Bare Area and **Bare Rock** were merged because of confusion in reference 2015 data between the two

All **"open"** classes left out

LCDB 2020 – Without "Open" Classes



Raster values

1	Settlements (UA, UA2, RH1, RH2)
2	Cropland (HCP, HCSM)
3	Degraded Cropland (HCER)
4	Needleleaf (TNL1)
5	Broadleaf (TBL1)
6	Permanent Water (WB1, WB2)
7	Wetland (WET)
8	River Banks (RB)
9	Shrubland (SH1)
10	Grassland (GR1)
11	Degraded Grassland (GR2)
12	Bare Surfaces (BR, BA)
13	Mines (MQ)
14	Irrigated Cropland (HCIR)
15	Gullies (GU)

ACCESS DATA

Single band raster:

https://cf2.cloudferro.com:8080/swift/v1/AUTH_3b25838791bc4272a2d905ab2107fd13/fao-cropl
c/pred_mosaic_32735_raw_D2020-09-09T14-21-00_B1_gullies.tif

Three-band raster (with embedded colors):

https://cf2.cloudferro.com:8080/swift/v1/AUTH_3b25838791bc4272a2d905ab2107fd13/fao-cropl
c/pred_mosaic_32735_raw_D2020-09-09T14-21-00_B1_gullies_color_cog.tif

WMTS endpoint:

https://services.sentinel-hub.com/ogc/wmts/aba8113d-0185-4d83-a654-33ec8a64f891

(WMTS layer name: **LCDB_2020_lesotho_latest**)

QGIS styling file for single band raster colors: Lesotho_landcover_style_2015_pres.qml

PRELIMINARY STATISTICS

2015 Aggregated LC Classes (Ha)	Upper Caledon	%	Middle Caledon	%	Lower Caledon	%	Makhaleng	%	Upper Senqu	%	Lower Senqu	%
Built-up	20886	8,04%	32370	12,57%	19903	10,91%	13936	4,69%	20595	1,37%	18401	3,30%
Agricultural land	91103	35,07%	110318	42,83%	92240	50,57%	75248	25,32%	138784	9,21%	70346	12,70%
Trees	10096	3,89%	12005	4,66%	3201	1,75%	2851	0,96%	4361	0,29%	5889	1,10%
Shrubland	38706	14,90%	22213	8,62%	8054	4,42%	53850	18,12%	312648	20,75%	148857	27,00%
Grassland	84670	32,59%	63645	24,71%	40561	22,24%	134308	45,20%	917318	60,89%	275549	49,90%
Wetland	508	0,20%	481	0,19%	403	0,22%	617	0,21%	25562	1,70%	5008	0,90%
Water Bodies & Rivers	1365	0,53%	2027	0,79%	2015	1,10%	2745	0,92%	15171	1,01%	4918	0,90%
Barrenland	10144,64	3,91%	10107,53	3,92%	10723,32	5,88%	10626,31	3,58%	71524,06	4,75%	21389,38	3,88%
Gullies	2303,36	0,89%	4415,47	1,71%	5315,68	2,91%	2988,69	1,01%	456,94	0,03%	1585,62	0,29%
TOTAL	259783	100%	257582	100%	182415	100%	297170	100%	1506419	100	551944	100%

2020 Aggregated LC Classes (Ha)	Upper Caledon	%	Middle Caledon	%	Lower Caledon	%	Makhaleng	%	Upper Senqu	%	Lower Senqu	%
Built-up	22028,35	8,57%	33339,98	13,21%	21328,51	12,21%	14664,9	5,02%	21474,4	1,43%	18928,39	3,4%
Agricultural land	103118,09	40,11%	125637,87	49,78%	104784,47	59,99%	89604,86	30,67%	178922,48	11,88%	85814,2	15,6%
Trees	5178,05	2,01%	2869,96	1,14%	498,54	0,29%	1762,29	0,60%	2124,18	0,14%	3177,47	0,6%
Shrubland	28646,24	11,14%	24888,39	9,86%	5503,29	3,15%	45069,06	15,43%	107961,62	7,17%	92953,81	16,9%
Grassland	85622,32	33,31%	51490,16	20,40%	24467,85	14,01%	115339,56	39,48%	1096353,55	72,80%	288415,09	52,4%
Wetland	583,15	0,23%	564,12	0,22%	139,81	0,08%	534,03	0,18%	7780,54	0,52%	3952,23	0,7%
Water Bodies & Rivers	1490	0,58%	2201,11	0,87%	2258,31	1,29%	2940,83	1,01%	16089,4	1,07%	5181,9	0,9%
Barrenland	10396,52	4,04%	11380,16	4,51%	15703,42	8,99%	22257,02	7,62%	75238,59	5,00%	51605,31	9,4%
Gullies	2717,94	1,06%	5210,41	2,06%	7730,47	4,43%	4997,75	1,71%	464,92	0,03%	1911,3	0,3%
TOTAL	257062,72	100%	252371,75	100%	174684,2	100%	292172,55	100%	1505944,76	100%	550028,4	100%

Difference Aggregated LC Classes (Ha)	Upper Caledon	%	Middle Caledon	%	Lower Caledon	%	Makhaleng	%	Upper Senqu	%	Lower Senqu	%
Built-up	1142,35	5,5%	969,98	3,0%	1425,51	7,2%	728,9	5,2%	879,4	4,3%	527,39	2,9%
Agricultural land	12015,09	13,2%	15319,87	13,9%	12544,47	13,6%	14356,86	19,1%	40138,48	28,9%	15468,2	22,0%
Trees	-4917,95	-48,7%	-9135,04	-76,1%	-2702,46	-84,4%	-1088,71	-38,2%	-2236,82	-51,3%	-2711,53	-46,0%
Shrubland	-10059,76	-26,0%	2675,39	12,0%	-2550,71	-31,7%	-8780,94	-16,3%	-204686,38	-65,5%	-55903,19	-37,6%
Grassland	952,32	1,1%	-12154,84	-19,1%	-16093,15	-39,7%	-18968,44	-14,1%	179035,55	19,5%	12866,09	4,7%
Wetland	75,15	14,8%	83,12	17,3%	-263,19	-65,3%	-82,97	-13,4%	-17781,46	-69,6%	-1055,77	-21,1%
Water Bodies & Rivers	125	9,2%	174,11	8,6%	243,31	12,1%	195,83	7,1%	918,4	6,1%	263,9	5,4%
Barrenland	251,88	2,5%	1272,63	12,6%	4980,1	46,4%	11630,71	109,5%	3714,53	5,2%	30215,93	141,3%
Gullies	414,58	18,0%	794,94	18,0%	2414,79	45,4%	2009,06	67,2%	7,98	1,7%	325,68	20,5%

DISCREPANCIES WITH LCDB 2015

Discrepancies != from misclassification errors

- Class definitions are fuzzy and landscape ecologies are complex
- Need to explore discrepancies through photo-interpretation in GIS software, but also through field survey, to understand whether the class definition are fit for purpose

Main discrepancies observed:

- “Open” shrublands mostly classified as grasslands
- “Open” tree classes classified as shrubland (small/sparse trees)
- Degraded Cropland much more extensive than 2015 → Need to analyse reason why (representative of different crop type, or real degradation?)
- Degraded Grassland classified as bare surface in many cases → Could be suggesting an erosion trend

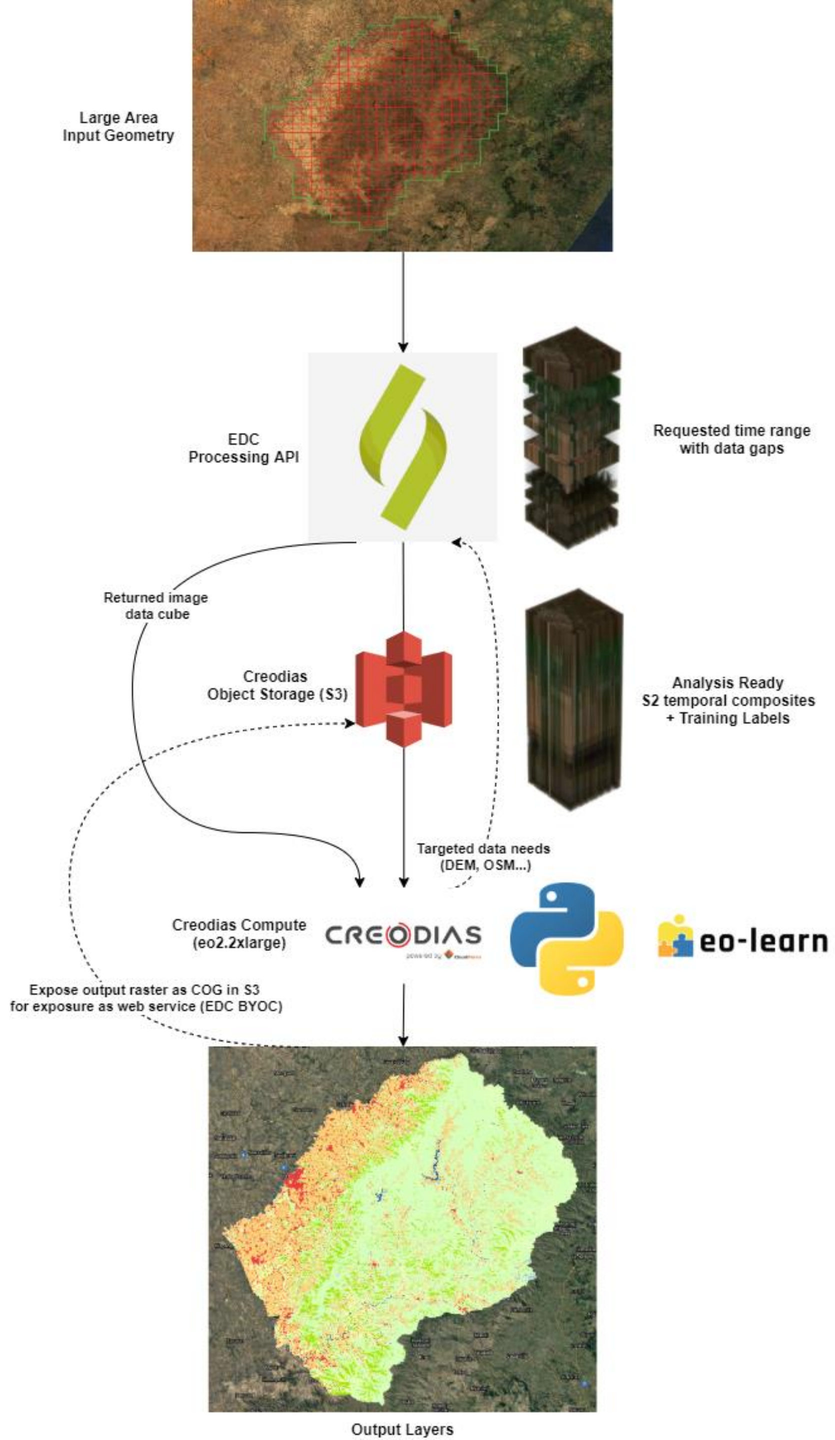
List of QGIS bookmarks showing examples of different types of discrepancies: LCDB2020_bookmarks.xml

CLOUD COMPUTING SETUP

Google Cloud Stack was used based on availability of resources:

- Google Earth Engine to generate the input features from Sentinel-2 and perform semi-automated sourcing of training data
- Input Features and training labels are exported to Google Cloud Storage
- Data is processed in Google Compute to generate LCDB 2020
- LCDB output is exposed through S3 as Web Map Tile Service (WMTS) and public download

Based on resources available, same process could be performed using an AWS/DIAS stack and solely relying on Sentinelhub (dashed arrows in diagram)



CONCLUSIONS

A novel approach for semi-automated land cover update was implemented for the LCDB 2020 of Lesotho

- The resolution drop from 0.5/2m to 10m (Sentinel-2) has limited impacts on overall predicted surface areas
- The land cover class nomenclature of LCDB 2015 required adaptation for the new LCDB 2020 methodology
- Results are promising, but require in-depth analysis of discrepancies between 2015/2020 and zonal statistics to assess usability
- An independent validation, based on photo-interpretation (https://laco-wiki.net/c/lcdb2020_ind_val) and a field survey (whenever the situation allows it) is still required to make a reliable accuracy assessment

A crowd-sourcing campaign for sourcing training data was carried out with participants from the ministry of agriculture and statistics

- The collected data didn't show an improvement to the LCDB 2020 output
- Crowd-sourcing training data is more appropriate once a first land cover classification iteration has been performed → digitalization of training in areas where land cover is wrongly classified

The cloud infrastructure costs to produce a national land cover update map are negligible (currently free under NoR sponsorship, but cheap if costs were internalized)

- This methodology could be deployed for other countries requiring a land cover update at manageable costs
- Cloud infrastructure costs could be pooled across projects to further reduce costs

REFERENCES

- Paris, C., Bruzzone, L., & Fernández-Prieto, D. (2019). A novel approach to the unsupervised update of land-cover maps by classification of time series of multispectral images. *IEEE Transactions on Geoscience and Remote Sensing*, *57*(7), 4259-4277.
- Paris, C., & Bruzzone, L. (2020). A novel approach to the unsupervised extraction of reliable training samples from thematic products. *IEEE Transactions on Geoscience and Remote Sensing*.
- Stoian, A., Poulain, V., Inglada, J., Poughon, V., & Derksen, D. (2019). Land cover maps production with high resolution satellite image time series and convolutional neural networks: Adaptations and limits for operational systems. *Remote Sensing*, *11*(17), 1986.
- Wulder, M. A., Coops, N. C., Roy, D. P., White, J. C., & Hermosilla, T. (2018). Land cover 2.0. *International Journal of Remote Sensing*, *39*(12), 4254-4284.