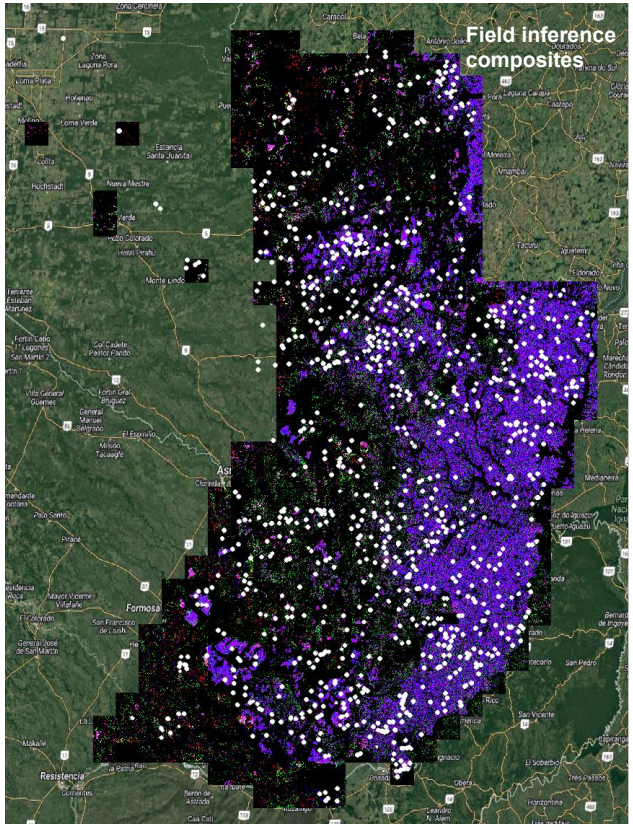
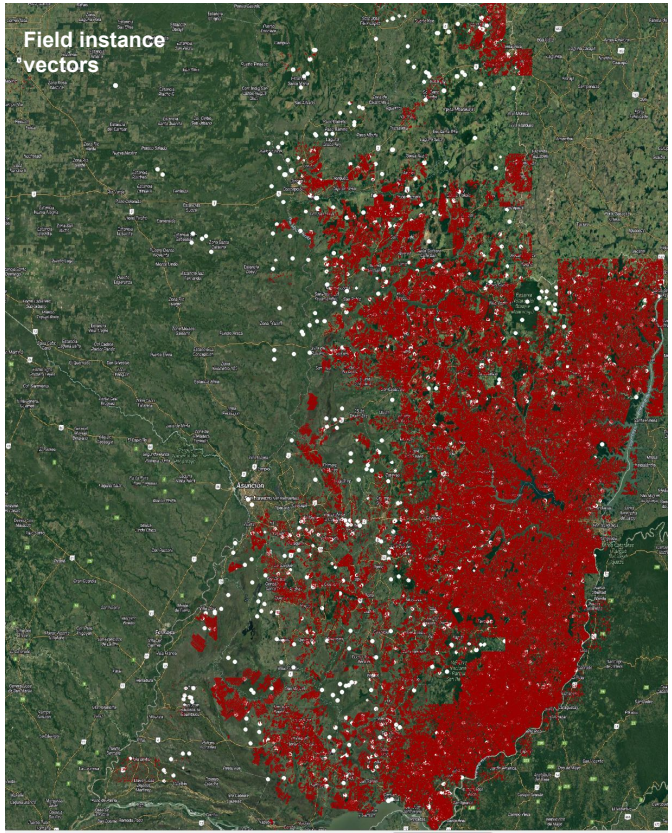
Segmentation updates



Lauren Sharwood
1/25/2024



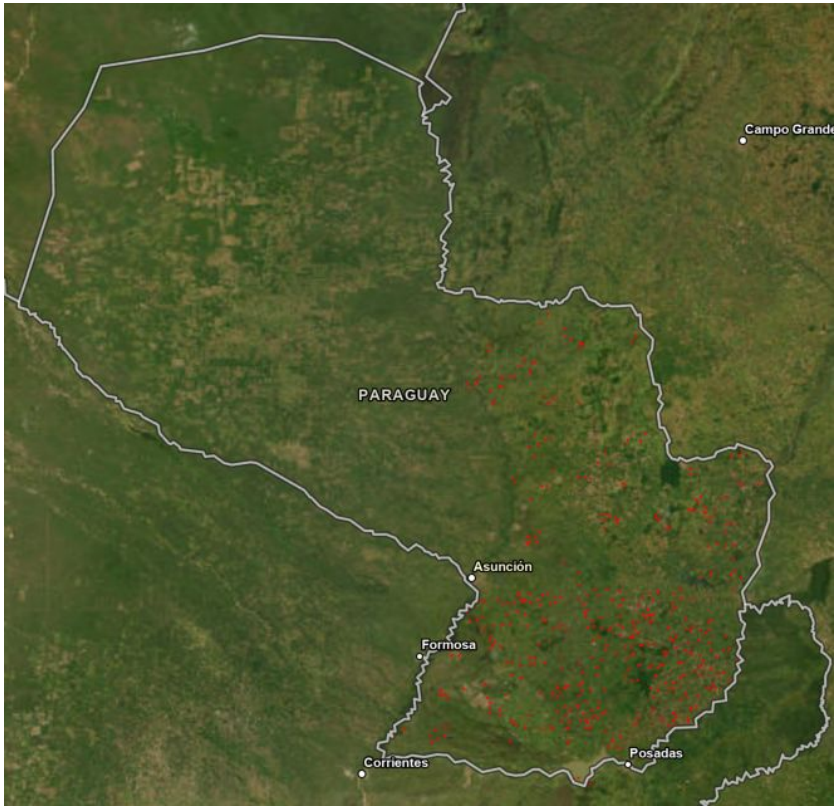
Paraguay Crop Digitization

Goal:

Create instances of crop fields across Paraguay that will be used to help with the crop-type classification.

Task:

Digitize polygons around individual agriculture fields across Paraguay to use as training for cultionet, a model that creates semantic segmentations (crop vs. not crop) of cultivated land using a neural network, UNet 3+.



Types of fields to digitize and label

Code	Class Name	Description
0	NON-CROP	If there aren't any crops (or code's 1, 2, 3, 4, 5) in the entire chip. Draw polygon encompassing chip, with a few meters buffered around.
1	**MONO-CROP**	**Single fields > 30m x 30m. If you can clearly see a road or a tree-line break between two fields growing the same plot in Planet imagery, draw separate fields as long as each individual field is > 30m x 30m**
2	**MIXED-CROPS**	**Single smallholder fields < 30m x 30m and inter-/mixed-cropping. Draw polygon around all contiguous inter/mixed cropping rows. Separate property means separate field as long as the polygon per property is > 30m x 30m.**
3	**YERBA-MATE**	**Medium-height crop vegetation**
4	**ORCHARD**	**To distinguish yerba mate.**
5	MANAGED GRASSLAND	To distinguish crops. Managed/planted grass ONLY; no wetlands or natural grasslands. Shape will likely look like soy or sugar crops (rectangular fields, very green).
99	UNSURE - CHECK	To FLAG if you're really unsure, someone will go back to check

ArcGIS

PROJECT SETUP:

- Load in imagery: Map tab > Basemap > World Imagery
- Load in training POLYGONS, and CHIPS to digitize from G:/My Drive/PyCropSeg_2023/training_2021
- LABEL field digitization polygons by code and chips by region (right-click layer > label to turn label on. Right-click > label properties > $feature.code OR $feature.region)
- For each chip region (labeled), search for it's planet composite in G:/My Drive/PyCropSeg_2023/planet_NDVI_composites
- Create a group (highlight layers > right-click > Group) if multiple images are loaded at once, when you zoom to a new location, turn image group off. This tells Arc it doesn't have to worry about loading and displaying the image while it jumps to the new location.

Useful Tools in ArcPro:

- **Open Web Maps Geocoder**: Add-In tab > Bing Maps-Aerial OR Google Maps-Hybrid
  a. Click a spot on the map to bring up high-resolution Bing or Google maps of that area
- **Image Swipe**: Appearance tab > Swipe
  a. Use image swipe tool to compare two overlaid images
- **Ruler**: Map tab > Measure
  a. Use measuring tool to make sure all field digitizations are > mmu
- **Time Series Profile Viewer**: Right-click on yearly Planet NDVI composite > Create Chart > Spectral Profile
  a. Within the profile viewer, drop points or draw polygons to see how given pixels' NDVI changes each month
-
-

TROUBLESHOOTING TIPS:

- **Save work often**
- With too many layers (particularly large images) loaded, Arc can get slow and buggy. When it starts to lag, it might crash soon. Remove or turn off some layers.
- Have images grouped, when you zoom to a new location, turn image group off. This tells Arc it doesn't have to worry about loading and displaying the image while it jumps to the new location.
- If an image isn't displaying…
  a. map might be over the wrong extent: Right-click layer > Zoom to layer
  b. refresh the symbology settings: Right-click layer > Remove (to temporarily remote the layer). Undo (Ctrl+Z) to add the layer back

Digitizing

Steps:

1: Zoom to chip region (right-click layer > zoom to layer)

2: In file explorer, search for Planet image by chip region then drag file into Arc

3: Digitize fields

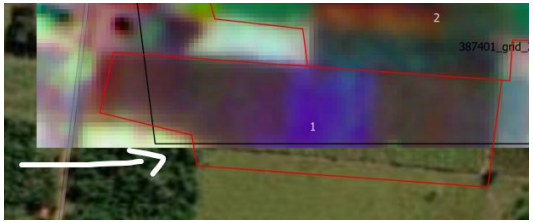
- Make new polygons under Map tab > Edit > Create > select polys layer in Create Features
- To edit existing shapes, use Reshape, Edit Vertices, Merge, Split

4: Update Attribute tables

- Add CODE to Polys attribute table
- Add time it took to complete the region in Chips shape

5: Save

Make sure all digitizations end outside the chip boundary



Should field be digitized?

- Any single digitization's shortest dimension must be at least 50m: To segment a field, cultionet needs one pixel for each boundary and one for the field itself, so at least 3 pixels across, which makes the minimum digitization size 50m x 50m

Should be digitized separately or grouped?

- look at time-series to see if NDVI/greenness ever significantly diverges for sample points in the two fields
- if you can see a road or trees in the Sentinel-2, separate the fields (if still > 50m x 50m) because cultionet should learn that that is an edge

Paraguay training chips

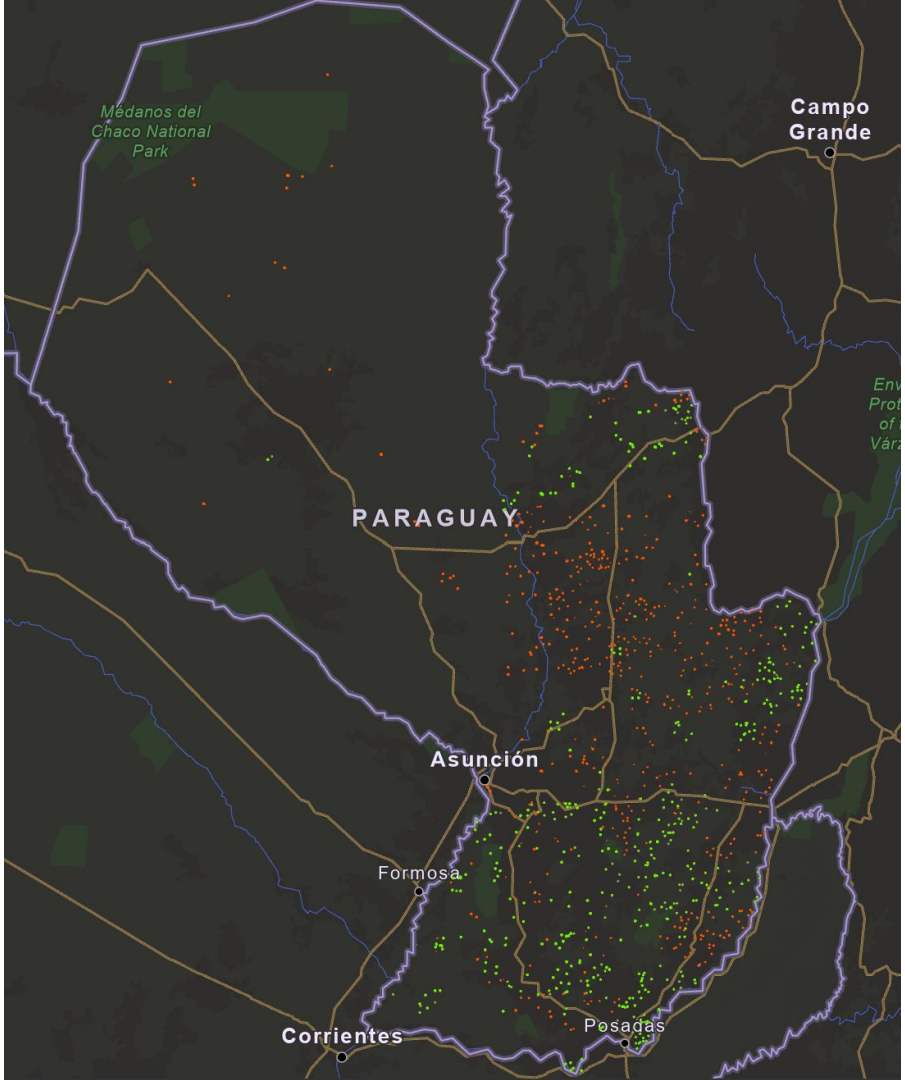
Number of chips:

- 1153 digitized (red)

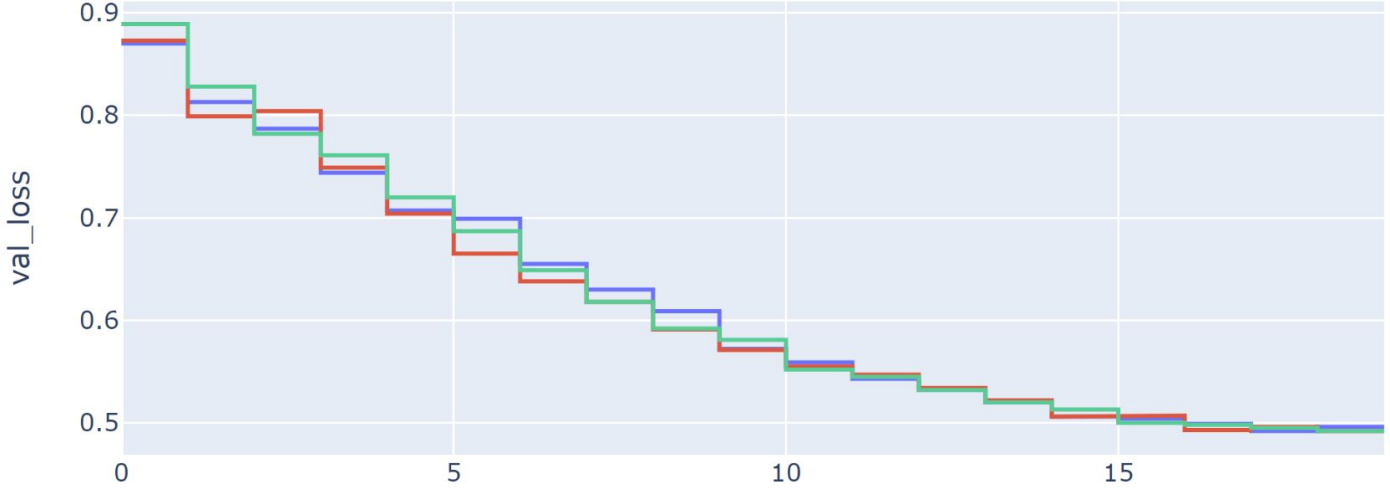
Crop-types digitized:

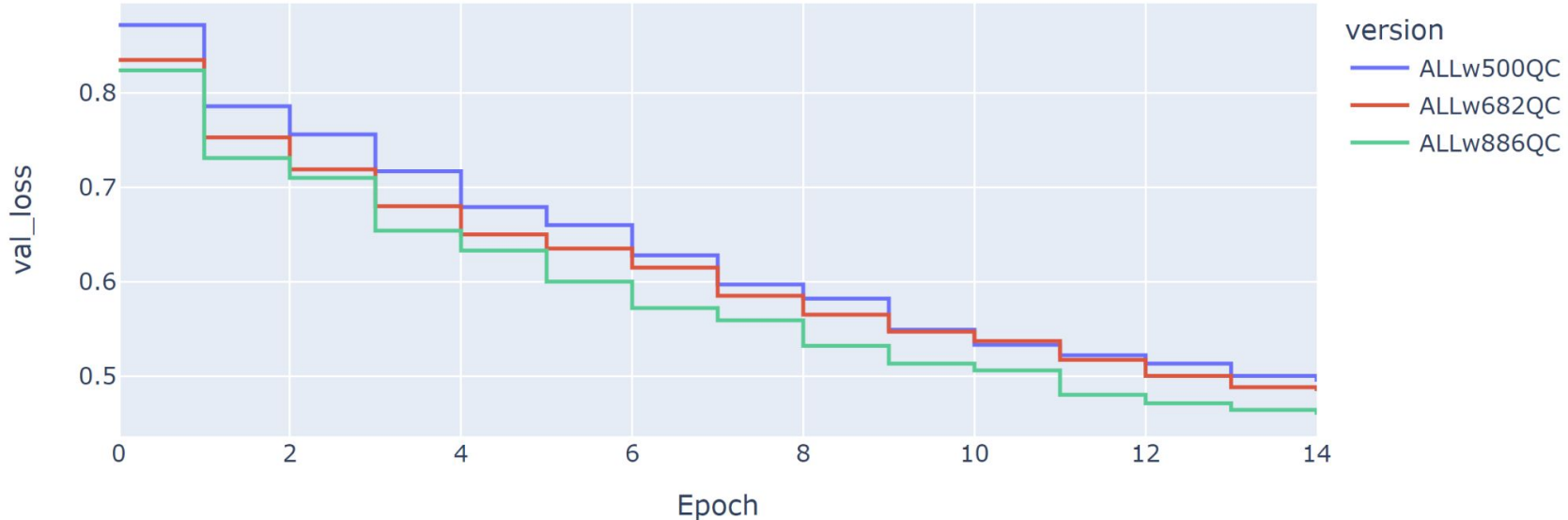
1. Mono-crop
2. Mixed-crops
3. Yerba-mate
4. Orchards

Sample chip:

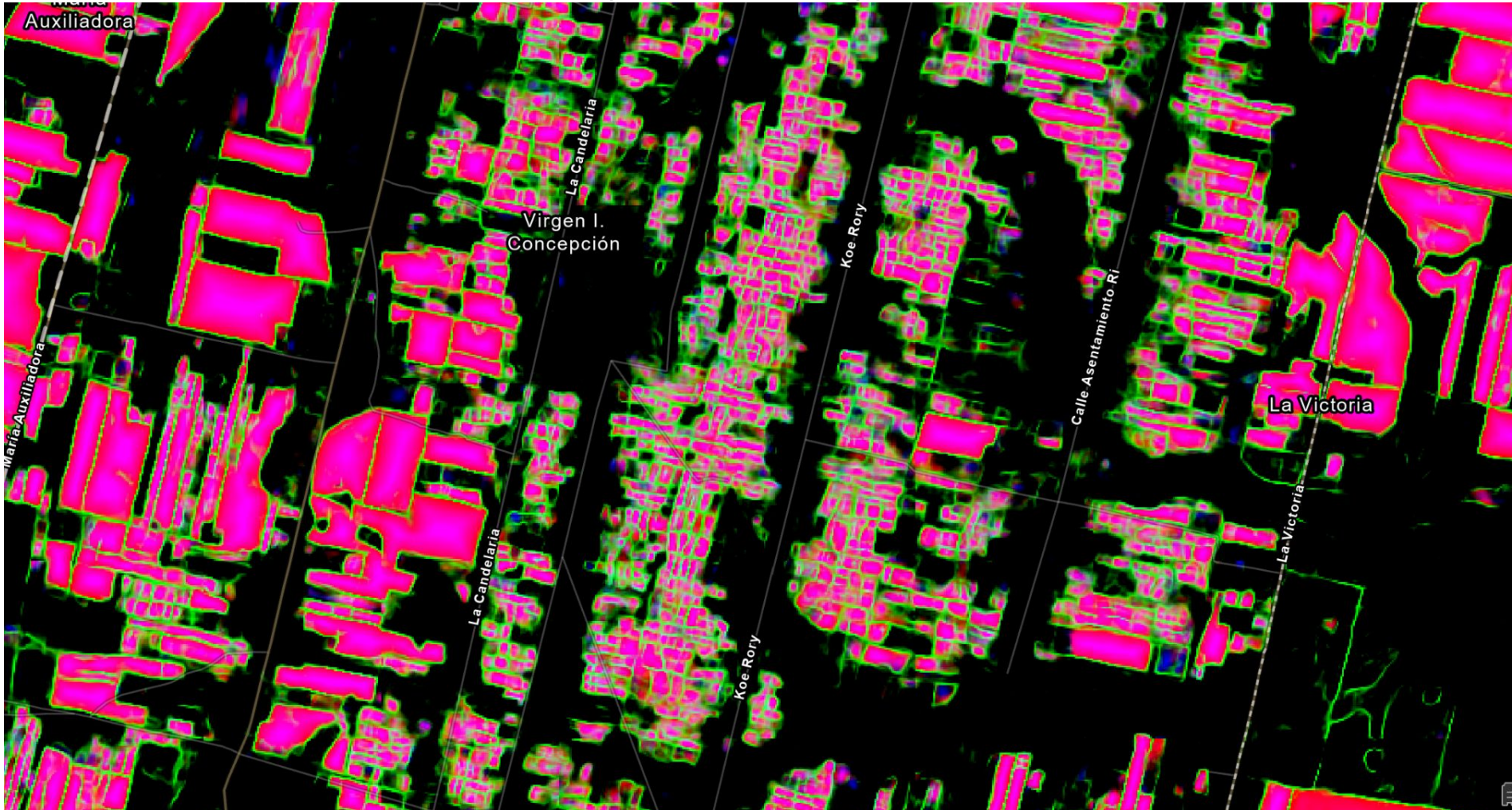


Cultionet


val_loss
0.9
0.8
0.7
0.6
0.5
0
5
10
15
version
PY_1wMate
PY_2wMed
PY_3wOrch


val_loss
0.8
0.7
0.6
0.5
0
2
4
6
8
10
12
14
Epoch
version
ALLw500QC
ALLw682QC
ALLw886QC

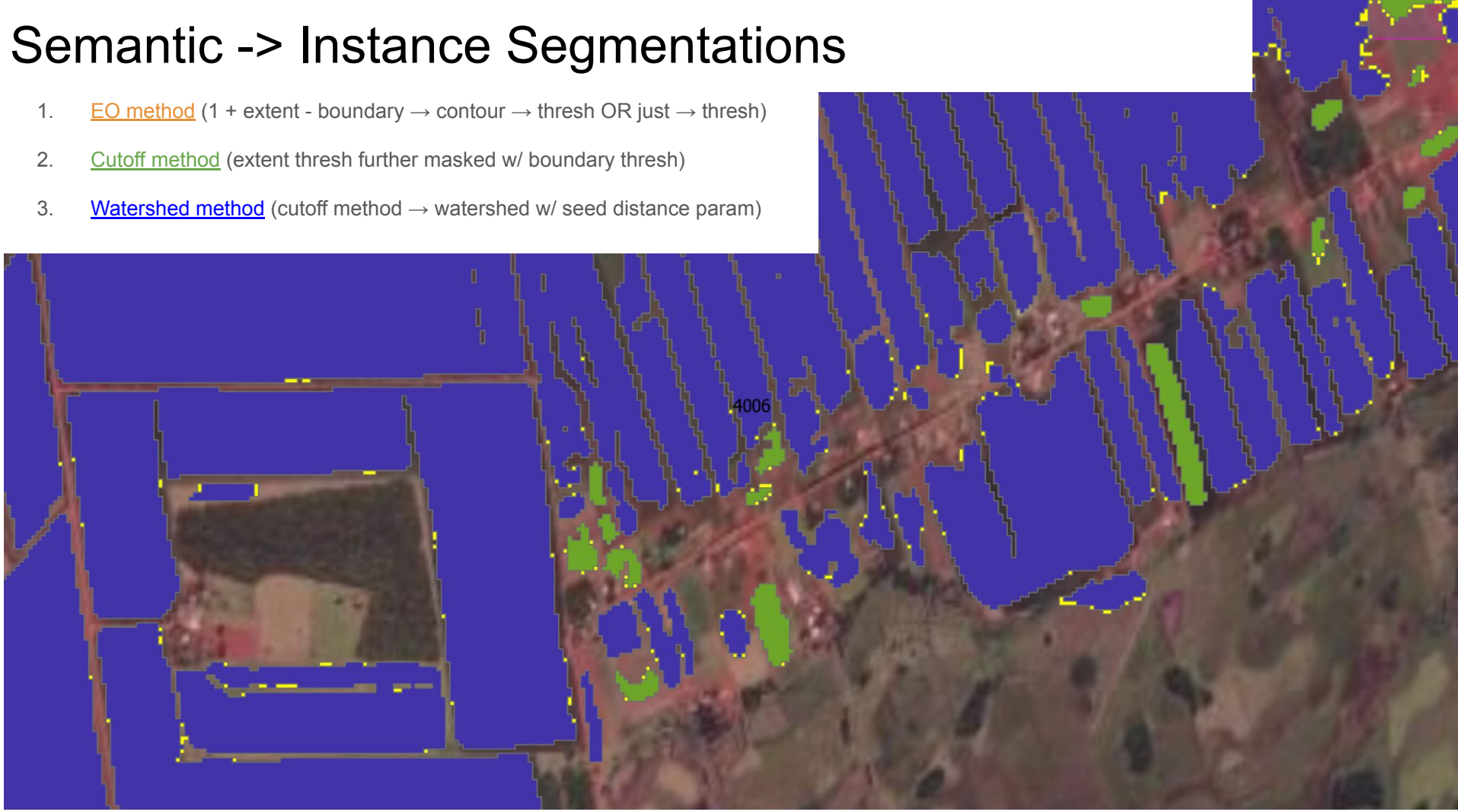
Cultionet inference composites - R: Extent proba, G: Border proba, B: distance to border

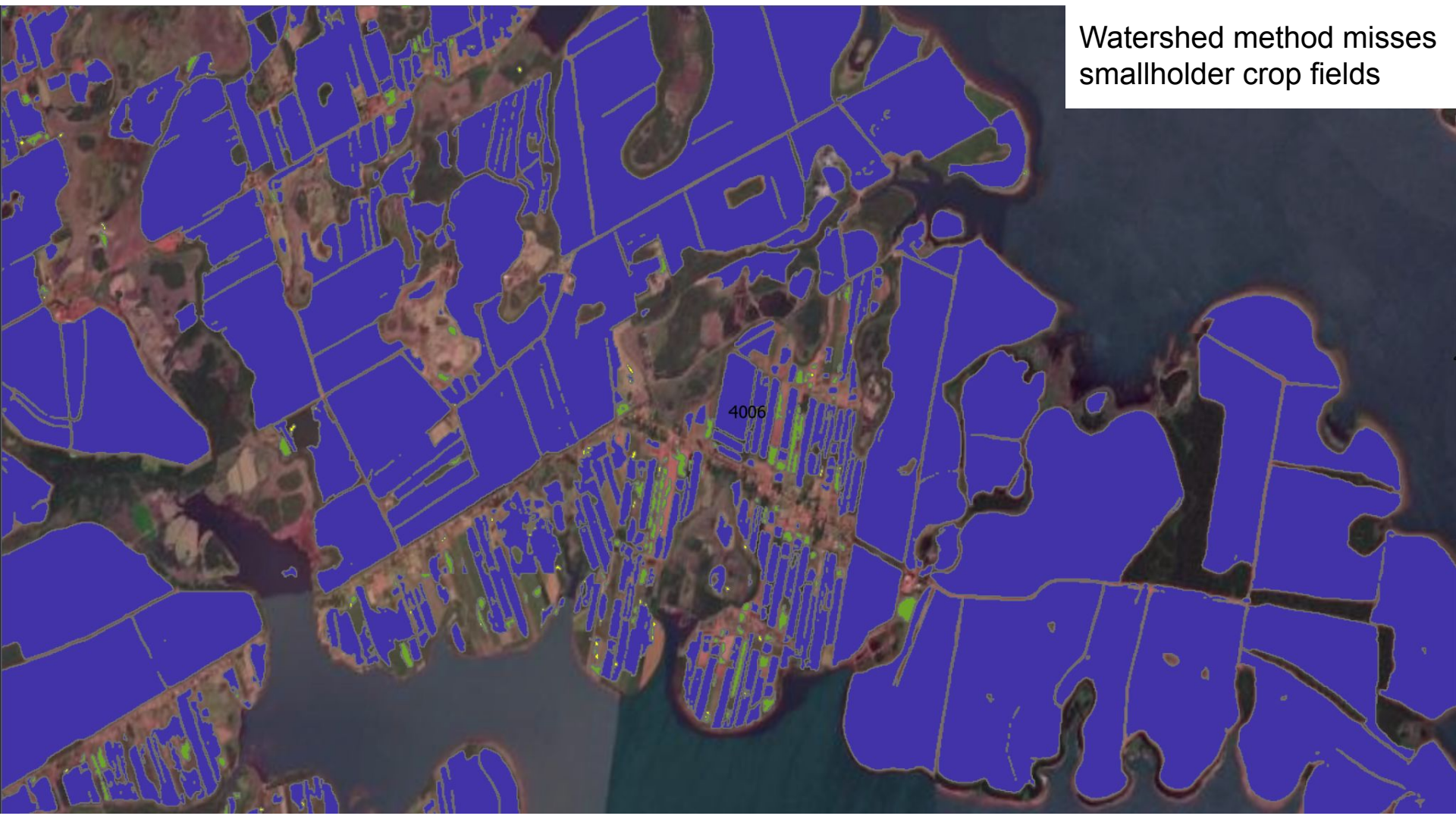



María Auxiliadora
La Candelaria
Virgen I. Concepción
Koe Rory
Calle Asentamiento Ri
La Victoria
María Auxiliadora
La Candelaria
Koe Rory
La Victoria

Semantic -> Instance Segmentations

1. EO method (1 + extent - boundary → contour → thresh OR just → thresh)
2. Cutoff method (extent thresh further masked w/ boundary thresh)
3. Watershed method (cutoff method → watershed w/ seed distance param)



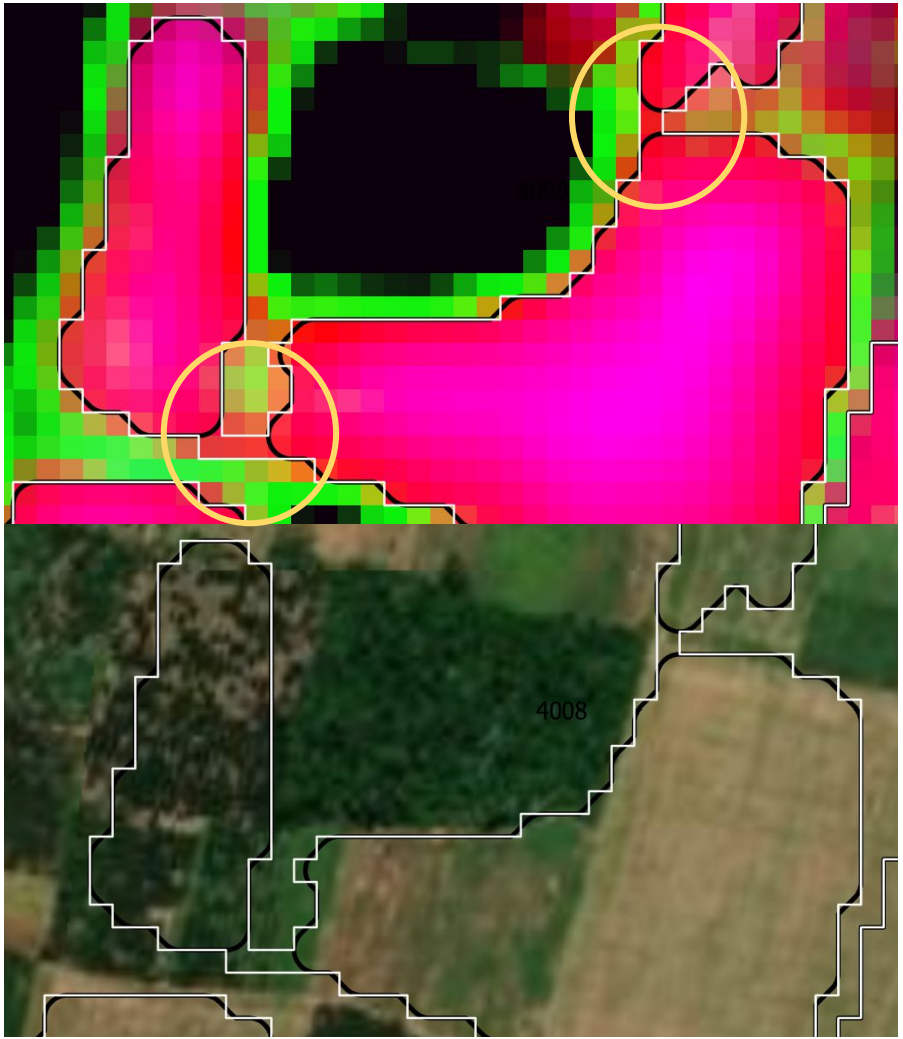

Watershed method misses
smallholder crop fields
4006


Contour (isolines) smoothes pixel borders but misses holes within fields
4012
ALTO PARANA DEPARTMENT
ITAPUA DEPARTMENT

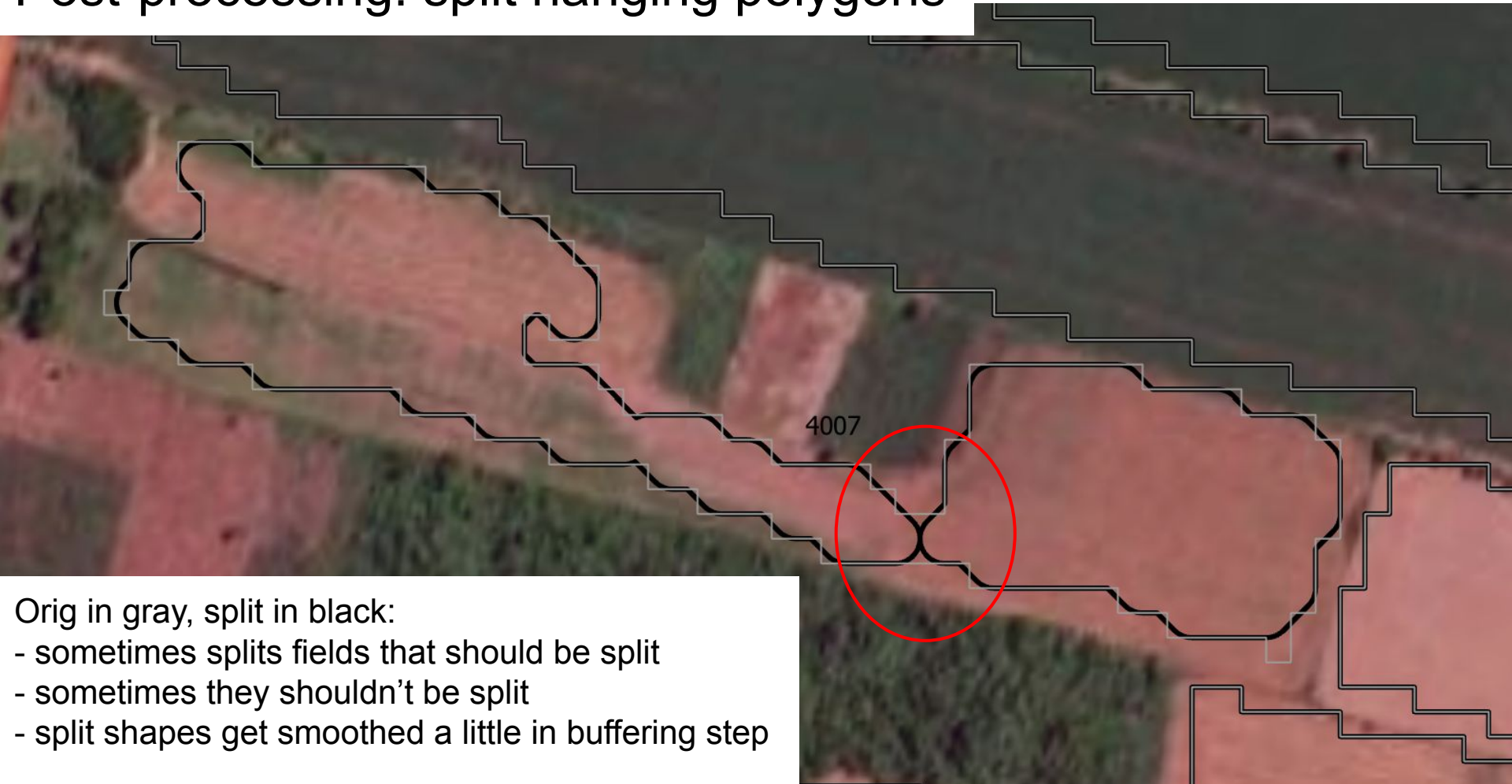

Contour (isolines) smoothes pixel borders but misses holes within fields
3703

Post-processing

1. Split hanging polygons
    a. Negative buffer (erode) vectors: If polygon is eroded to nothing or polygon isn't split, keep old (uneroded) geom. If split (eroded to multiple polygons), use new geom and rebuffer
2. Smooth
    a. <u>For integrating with pixel-based classification, skipping this step</u>



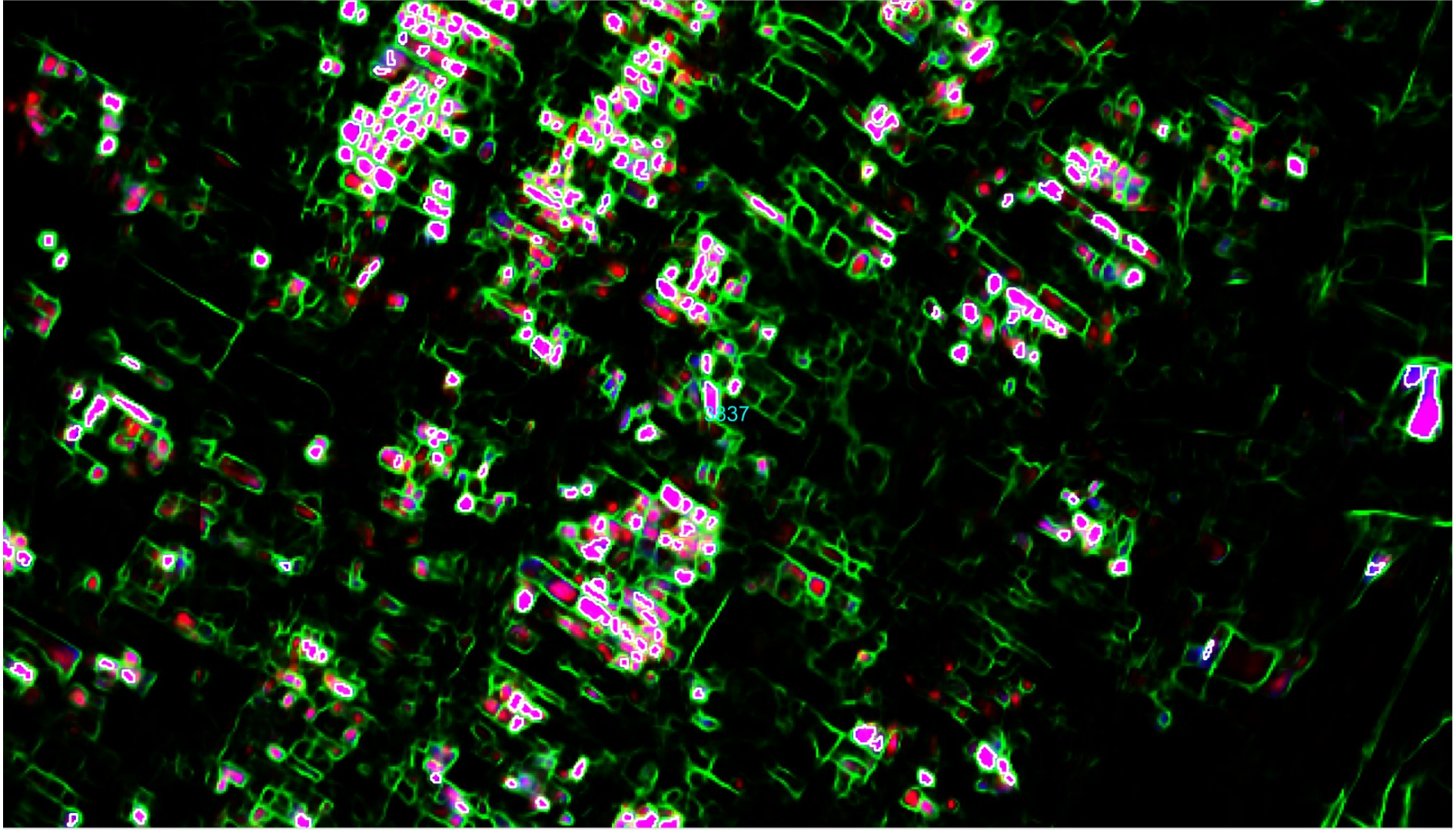
Post-processing: split hanging polygons

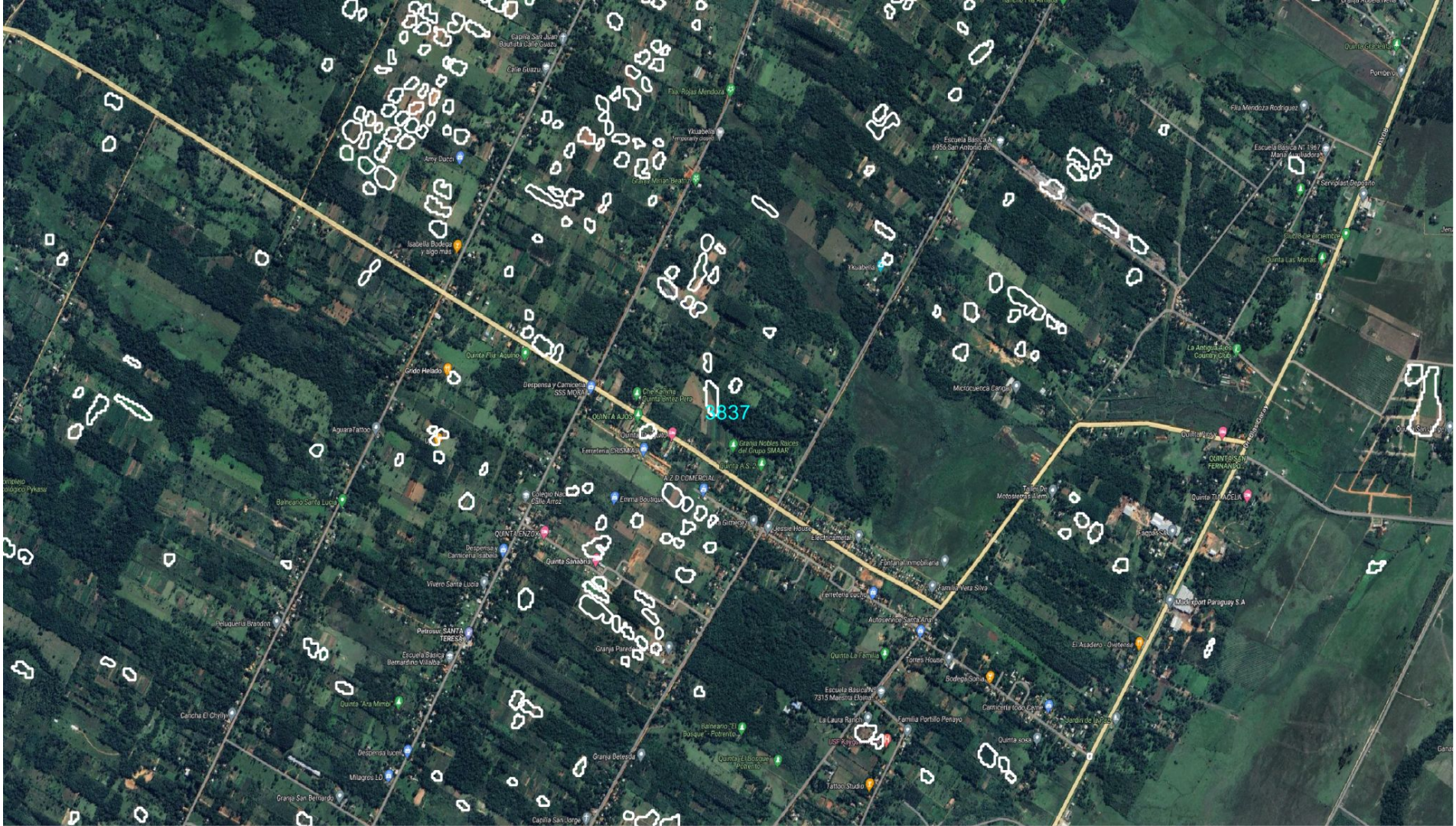


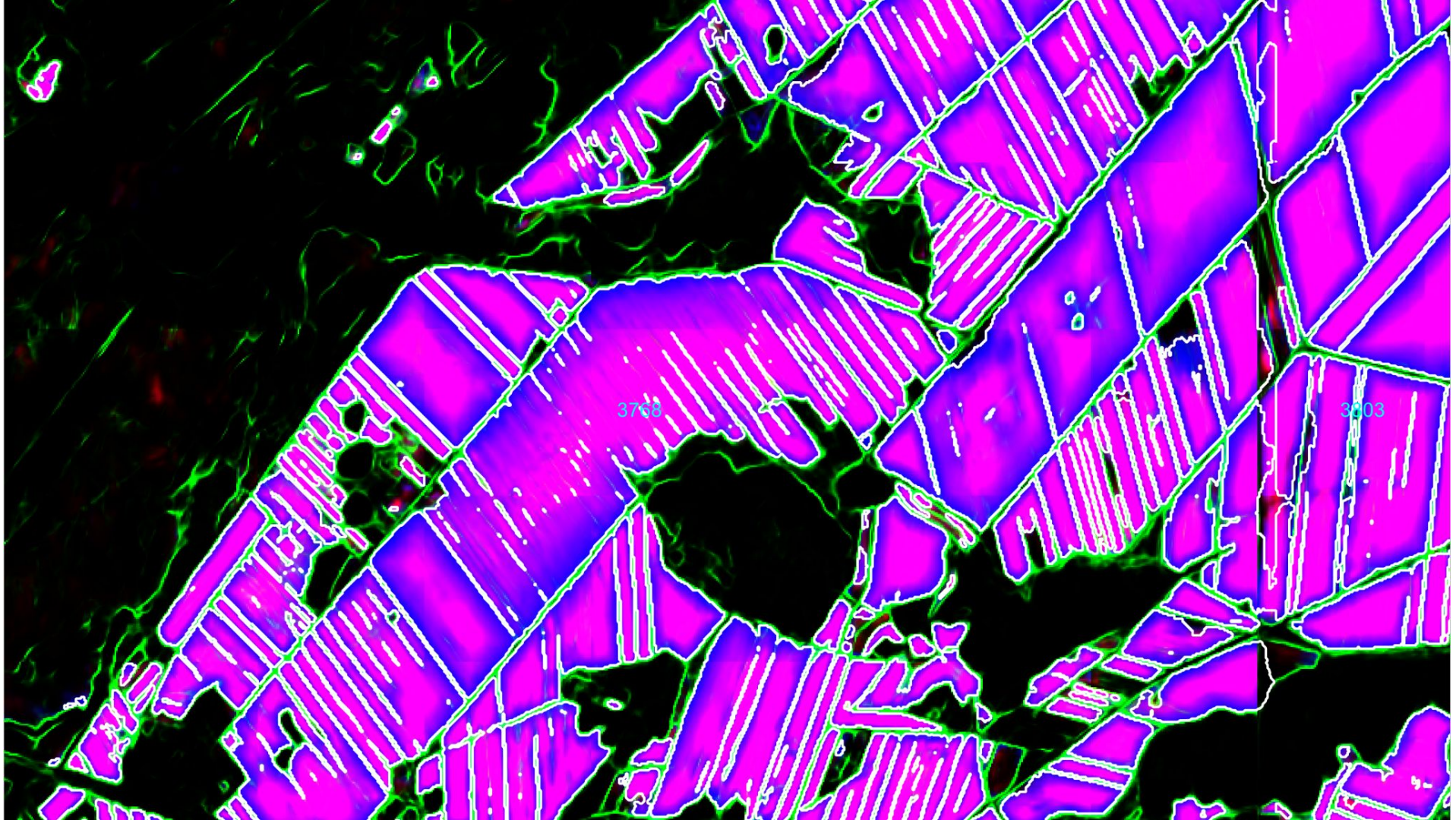
Orig in gray, split in black:
- sometimes splits fields that should be split
- sometimes they shouldn't be split
- split shapes get smoothed a little in buffering step


Letty Coiffure
3877
3912
Parque ecológico 1° de Marzo
UNIDAD DE SALUD FAMILIAR


3801
Granja Fila Bogado




3837


3768
3903


3768
3803
PK10
PASO KIPE

Chip accuracy

Metrics:

1. Number of field polygons
2. average field area
3. total crop area/1km x 1km chips

For each metric, find if ref value is greater or less than pred… then use different formula to make greater positive numbers=worse overprediction, more negative numbers=worse underprediction

- Underpredicting – ref > pred, make more negative: 1-(ref/pred)
- Overpedicting – when pred > greater, make more positive: (pred/ref)-1
- Perfect (no error) – ref = pred: 0

Split 950 ref chips by one w/ @ least 1 reference crop polygon (753) and w/out any (197)

Chip accuracy + comparison

- **chips w/ crops**, **chips w/out crops**, **combined**
- smaller # below 0 (more neg) = underpredicting more
- larger # above 0 (more pos) = overpredicting more
- Colors: saturation by version (rows)… more saturated = version w/ more error & gray = version w/ least error

version	chips wCrops error			chips noCrops error			normalized + chips wCrops			normalized + chips noCrops			sum norm 2*wCrops + noCrops err			sum all metrics err
	numFields	fieldArea	totalArea	numFields	avgArea	totalArea	numFields	fieldArea	totalArea	numFields	fieldArea	totalArea	numFields	fieldArea	totalArea	
EO_6th	-0.18	-671.42	-3,244.05	0.13	695.52	811.59	0.06	0.28	0.25	0.17	0.21	0.17	0.29	0.76	0.67	**1.72**
EO_7th	-0.19	-693.48	-3,318.35	0.10	556.98	665.27	0.07	0.30	0.26	0.09	0.12	0.10	0.22	0.72	0.63	**1.57**
EO_7.5	-0.19	-710.98	-3,392.42	0.10	507.68	599.16	0.07	0.31	0.28	0.09	0.09	0.08	0.23	0.72	0.63	**1.57**
EO_8	-0.20	-711.95	-3,396.98	0.09	478.23	548.30	0.09	0.31	0.28	0.04	0.08	0.05	0.22	0.70	0.61	**1.53**
EO_8.5	-0.21	-711.05	-3,392.63	0.07	431.43	485.06	0.10	0.31	0.28	0.00	0.05	0.03	0.20	0.67	0.58	**1.45**
EO_9	-0.23	-722.24	-3,403.96	0.07	354.42	427.11	0.13	0.32	0.28	0.00	0.00	0.00	0.27	0.64	0.56	**1.47**
cutoff_b30_e40	-0.16	-642.78	-3,215.24	0.17	763.59	1,011.70	0.01	0.25	0.24	0.29	0.25	0.26	0.32	0.75	0.74	**1.81**
cutoff_b40_e20	-0.15	-370.00	-1,918.92	0.28	1,215.25	1,860.40	0.00	0.00	0.00	0.59	0.53	0.93	0.59	0.53	0.93	**2.05**
cutoff_b40_e40	-0.22	-642.67	-3,215.15	0.17	771.22	1,021.96	0.12	0.25	0.24	0.29	0.26	0.26	0.53	0.76	0.75	**2.03**
cutoff_b40_e50	-0.25	-668.19	-3,240.84	0.13	717.44	836.25	0.18	0.27	0.25	0.17	0.22	0.18	0.53	0.77	0.68	**1.97**
cutoff_b40_e60	-0.26	-683.70	-3,271.55	0.10	568.27	682.03	0.19	0.29	0.25	0.09	0.13	0.11	0.46	0.71	0.62	**1.79**
cutoff_b50_e40	-0.31	-642.56	-3,215.12	0.17	776.05	1,029.06	0.27	0.25	0.24	0.29	0.26	0.26	0.83	0.76	0.75	**2.34**
Wtrshd_b40_e20_s15	-0.33	-382.75	-1,931.86	0.24	1,271.98	1,783.24	0.31	0.01	0.00	0.49	0.56	0.59	1.12	0.59	0.60	**2.31**
Wtrshd_b40_e60_s10	-0.31	-683.66	-3,271.57	0.10	568.27	682.03	0.27	0.29	0.25	0.09	0.13	0.11	0.63	0.71	0.62	**1.96**
Wtrshd_b40_e60_s15	-0.42	-684.50	-3,275.98	0.09	561.47	666.13	0.47	0.29	0.25	0.06	0.13	0.10	1.01	0.70	0.61	**2.32**

(value - row max)/ (row max + row min)

(value - row min)/ (row max + row min) → normalize

2 * wCrops + noCrops

Combine w/Crop + noCrop errors →

Integrating with RF classification

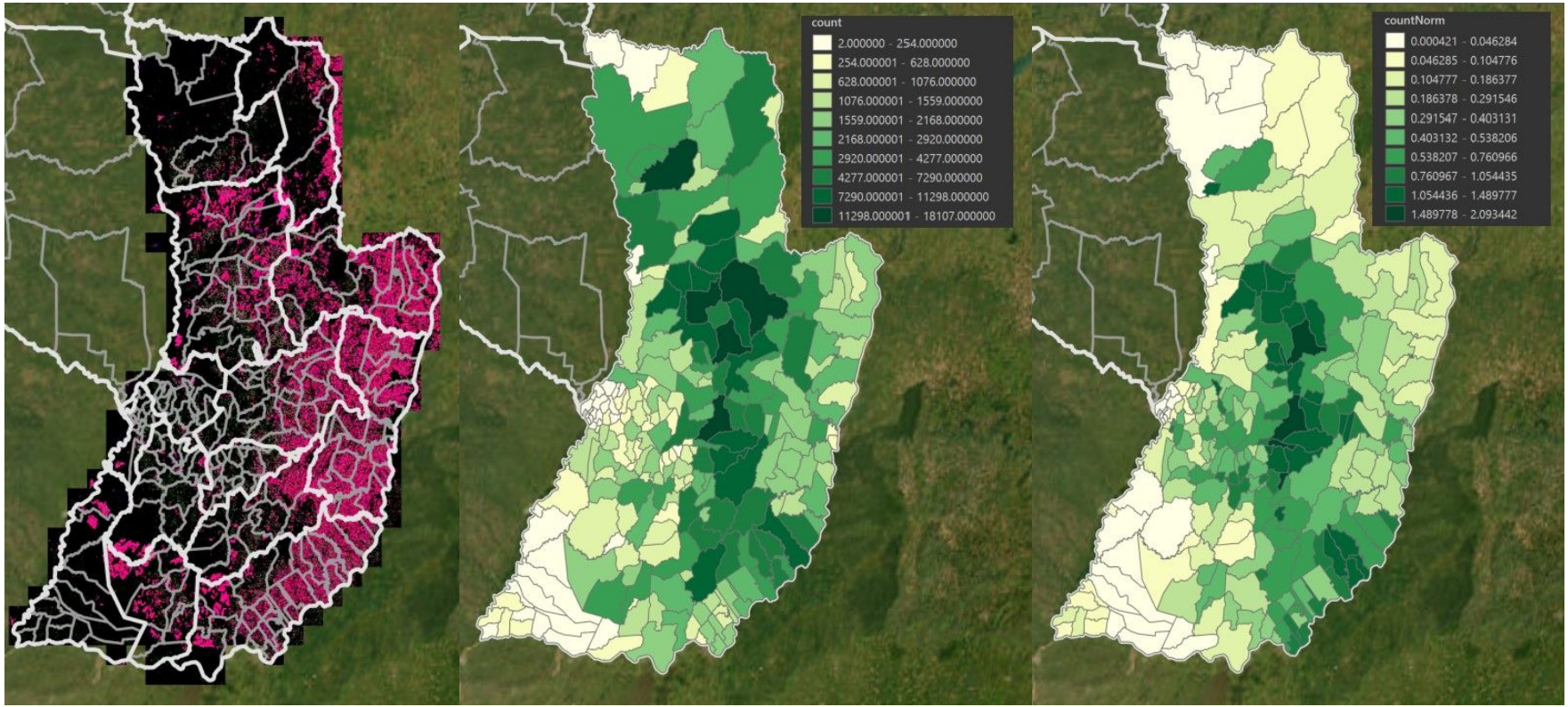
1. **Rasters** for RF predictive features:

- Crop extent probability (cultionet prediction band 3)
- Distance to crop field border (cultionet prediction band 1)
- Individual crop field area
- Individual crop field area/perimeter ratio
- Texture: StDev of each field polygon’s avg Nov & Dec GCVI (5 smoothed dates)

2. **Vectors** of individual fields for majority pixel class attribution

Field size stats by district

L: # of fields, R: # of fields/area



Field size stats by district

Blue: avg field size, Orange: st dev of field size