Aug 1

**Desk**

Worked on GEE file “Timeseries Validation w Matopiba v6”, tried to validate plant/harvest estimates. The newest thing is that there’s an option to get rid of entire polygons if a certain percentage of its pixels don’t have reasonable quarter period. However, lots of polygons are gotten rid of – looking at specific pixel EVI TS, lot seem to be natural vegetation even though it’s a CAR poly field. Perhaps use agricultural area/total area, and legal reserve area that’s reported sometimes to see if there’s a relation between quality of estimate and percent area that’s farmed?

Aug 2

**Desk**

In GEE file “timeseries Validation w Matopiba v6”, correlated agri area (from Jake’s land use map) in each polygon to the percent bad pixels in each polygon.

Used Jake’s agri pixels only to estimate plant/harvest in each polygon – improved performance a little but there are still some extremely bad ones.

Looked at why specific polygons do so badly

Started Aug 15 lab meeting powerpoint

Aug 6

**Avery meeting**

* He recommended, for plant/harvest validation, to use specific planted area for CAR polys.
* Also tasked me with getting percentiles of plant, peak and harvest dates for 25km radius of chosen soy points, pooled over 2003 to 2016.

**Sally meeting**

* For plant/harvest estimates, need to propagate error due to incorrect underlying land use.
* Pin down how to select for agri pixels (test the land use classification with CAR polygon natural veg poly)
* Don’t spend too much more time fine tuning estimations to these methods

**Desk**

* Started GEE file Modis Phenology v4 to calculate LAI from Kross SR with C2 mask for Modis Aqua and Terra, for rally points. Put the resulting LAI timeseries for Rally points on GitHub
* Started GEE file Crop Dates Statistics to get quantiles of 2003-2016 first and second crop plant, peak, and harvest dates within a 25 km radius

Aug 7

**Desk**

Started GEE file Crop Dates Statistics v2 to get quantiles of 2003-2016 first and second crop plant, pea nad harvest dates within a 25 km radius. See PPT for this week for warnings on potential ways these estimates could be biased.

Aug 9

**Desk**

* In GEE file Modis Phenology v4, added ability to calculate plant dates across Brazil (i.e. doubleCrop\_timing\_full) and export the Rally point plant estimates
* Looked at using specific CAR categories (i.e. natural vegetation) and mapbiomas to constrain the areas we look at in estimating plant/harvest for Matopiba points. For CAR categories, see Avery’s list of what to mask out as non-agri and what to keep as agri. Download CAR polygons for the municipalities that overlap with Matopiba’s CAR polygons; Graham Jeffries provided the path to this data in one to one Slack message.
* New GEE file: “Land Use Classification” to look at overlaying CAR land classifications with Mapbiomas cropland layer. Use it to explore ways to classify agri land in order to validate planting/harvesting estimates later.
* In Crop Dates Statistics v2, added “unique identifier pair” for every pixel in Brazil based on lat and long, and worked on exporting every 2016 soy pixel’s 25km, 2003-2016 percentiles (need to break Brazil into regions first)

**Jake, Avery meeting**

* In Jake’s GEE, weather station LST -> LST to NSAT will be helpful.
* For percentiles of soy plant/peak/harvest,
  + first add in a unique id for every pixel in Brazil that’s an integer
  + sample all soy pixels in 2017 and make them into points
  + calc percentiles for 2003-2017, 25 km radius
  + calc percentiles for each separate year, 25 km radius

Aug 10

**Desk**

Worked on GEE script Crop Dates Statistics v2 – changed using an array of [0] when there’s no soy pixel in year/25km region to using an array of [2016 avg value] depending on single, doubleFirst or doubleSecond. For sampling soy points in 2016, did it at scale 1000 instead of 500 because at 500 there’s a chance that the lat/longs for two sampled points are the same

Got some temperature tasks from Sally

Aug 11

**Desk**

In GEE file Crop Dates Statistics v2, did random stratified sampling to get a manageable number of points

New GEE file: Modis Phenology v5, which uses new rally points and assigns each point to its rally-reported ID

Aug 12

**Desk**

Put randomly sampled soy plant/peak/harvest percentiles onto GitHub in project\_crop\_dates, but need to redo it to have spatially stratified sample every 5km – this is in GEE file Crop Dates Statistics v3

Worked on downloading and ingesting into GEE the specific CAR polys associated with legal reserve, natural veg, etc in Matopiba for validating plant/harvest

Aug 13

**Desk**

Created 30km stratified samples of plant/harvest percentiles in GEE file Crop Dates Statistics v3

Downloaded and ingested specific CAR poly data on natural vegetation, etc into GEE assets CARpolys\_landUse\_features and CARpolys\_landUse\_image

Aug 14

**Desk**

Looked at Jake’s GEE script Forest Sample to produce timeseries of solar radiation, top of atmosphere radiation and DOY. Made it my own GEE script called Forest Sample v2

Where I got points for the Xavier dataset: <https://utexas.app.box.com/v/Xavier-etal-IJOC-DATA/folder/51193938759>

Aug 15

**Desk**

Got Brazil biome shapefiles

Got municipality code – state matchup, use that to get the recommended plant/harvest for each state to replace very badly estimated pixels

Aug 16

**Desk**

Started GEE file Crop Dates Statistics v4 to do stratified samples of plant/harvest estimates from 2003-2016, over 25km circles, at the 5km level (so it requires automation of exporting and of stitching csvs back together)

Aug 17-22

**Desk**

Worked on Avery’s stuff, files are Forest Sample v2 in my repo and extract\_crop\_dates\_v2 in Avery’s repo on GEE

Aug 23

**Desk**

Looked at mapbiomas 2.3 and 3 with GEE file Mapbiomas Explore, made confusion matrices and looked at transitions, counted unusual/impossible transitions (i.e. forest to farm to forest again)

Aug 27

**Meeting with Ashok**

He suggests meeting with Michael d Sohn at lbl ([mdsohn@lbl.gov](mailto:mdsohn@lbl.gov)) for causal stats coaching

Why aren’t we looking at plant data in the US?

In addition to using NDVI peak and inflection to determine planting/harvest, use temperature to adjust the distance from inflection point to planting date?

Use soil moisture to see when it rained instead of rain data, to stay in high resolution

Are farmers slamming to one edge of the accepted credit-planting window to meet the insurance requirements for planting window?

Labor shortages, machine rentals, etc all affect planting date

Aug 31

Desk

Worked on adding CAR specific natural veg polygons as a way to select pixels for validating plant/harvest estimates in Matopiba (GEE: Timeseries Validation Matopiba v6), but found that (perhaps because