November 12 - 13

**Desk**

Started Timeseries\_Sentinel, looked at VH and VV timeseries for soy pixels, found not amazing correlations between unsmoothed, uncleaned backscatter and EVI. Some pixels have very sparse backscatter data, but it may still be helpful in the rare cases when there’s a huge gap in EVI due to clouds during greenup. LAI of 1 corresponds to EVI of 0.25 and LAI of 3 corresponds to EVI of 0.5.

Calculated linear regression slope and Pearson correlation between EVI and VH, VV, also smoothed with a moving window to get averaged VH and VV on corresponding EVI observation dates in order to compare a smoothed version of VH and VV to smoothed EVI.

There seems to be a slight improvement in regressed slope and correlation when comparing smoothed Sentinel to smoothed EVI versus raw Sentinel to raw EVI but the improvement isn’t super obvious or spatially consistent so it’s worth trying VH, VV in both smoothed and unsmoothed forms? And decide on one using the ‘reasonableness’ of the estimate

Nov 14

**Desk**

In GEE file Timeseries\_Sentinel, calculated first derivative of VH and VV, calculated the max of the first derivative (had to separate single vs double cropped areas at this point). It looks like VH and HH are exactly the same for certain pixels, but not all pixels. The quarter periods derived from VH and HH are exactly the same in most cases, but only because the max of the first derivative happen to be the same (VH and HH timeseries look very similar). So, use either VH or VV to get quarter period, not both (because it would put too much trust into Sentinel timeseries). Pick either VH or VV based on the reasonableness of the estimate.

Note, there’s a weird thing where the smoothed first derivative of VV and VH goes all the way back to early August even though VV and VH timeseries and unsmoothed VV and VH first derivatives don’t go back there, so the smoothed first derivative appears to ‘change very quickly’ during August. At a 40 minute moving window, the first derivative around greenup isn’t affected by stuff going on in August so it’s okay.

Tried to mask out pixels where the max EVI minus amplitude/2 is greater than 0.6, a lot of area is masked out. Unsure yet whether the masked out places definitely have worse predictions.

Nov 15

**Desk**

In GEE file Timeseries\_Sentinel, made maps of ‘reasonableness’ of quarter period estimates from Sentinel and MODIS as the distance between the estimated quarter period and 40 days.

* Only choose inflection points for Sentinel within the window 60 to 10 days before MODIS-estimated peak day. Using 60 days as a cutoff seems to get rid of a lot of super early inflection point estimates. There seems to be a tendency for a lot of Sentinel derived estimates of inflection point to be around 10 days before the peak – perhaps this means VH and VV naturally behave differently, or means there needs to be more smoothing.
* After ‘doubly smoothing’ the first derivative of VH and VV, they seem to reach ‘maxes’ at different times – i.e. VV has higher quarter period than VH (VV’s inflection point is earlier than VH’s). Perhaps it’s better to take an average of VH and VV? VV seems to be more consistently close to MODIS derived inflection point than VH, which tends to be earlier or later.
* Important: get rid of the ‘extra filler points’ during derivative calculation; otherwise it will affect the estimates, tends to give quarter period around 58 (because the cutoff is 60, but the derivative keeps increasing beyond 60)

In GEE file Landsat Phenology v4, started getting LAI timeseries from Landsat. Saved in Drive, Gabriel\_Rally\_Landsat

**Meeting for AgroServe**

Two things for me to do

* scaling of point-scale estimates of damage to state wide and Cerrado wide estimates of damage
* tell detailed story about a farmer in MT and in Matopiba about historical and future climate predictions, include a cartoon to link the pieces

Nov 16

**Desk**

In GEE file Timeseries\_Sentinel\_v2, cleaned up the code. Need to make sure the averaging doesn’t add on extra information in August/September when there isn’t actually any VV, VH data there because it will mess up the derivative. Mask out at VV -> mean VV and derivative -> smoothed derivative.

NEED TO change decibels to power before doing averaging, etc! then turn power back into decibels before plotting.

Nov 20

**Meeting with Avery**

Crop dates summary stats at the state level: look at timeseries of plant date, use R’s ggridges package on a sampled grid from GEE. Look at how crop dates vary in space and time. Look at constraining planting date estimates to only the good ones; they tend to be quite late.

Property level: pick on property from Matopiba survey and one from MT. Describe how those polygons are different/similar to their neighbors, include yield information.

Get Brazil CAR polygons from GitHub

Get Dave’s supersample points and use it and Rally to pick properties for property specific stuff

Nov 21

**Desk**

From Tufts cluster, downloaded CAR polys by the state and ingested them into GEE

In GEE file Crop\_Vis, begin gathering CAR polys and Dave’s soy points for picking polygons for visualization

Nov 26 - 27

**Desk**

In GEE file Crop\_Vis,

* Picked out two CAR polys in Matopiba and four CAR polys in Mato Grosso to look at their individual farm’s plant/harvest behavior over time and compare them to their neighbors. Picked out these polygons based on their overlap with Rally points (for Mato Grosso), quality of Matopiba survey data (for Matopiba), amount and temporal consistency of soy planted area, overlap with Dave’s sample points.
* For a test ‘chosen’ CAR poly, got the neighboring CAR polys and reduce regions to get avg plant/harvest, stdev plant/harvest, and total planted area of single and double crop. Also set up the output table to contain the distance of the given polygon to the farm of interest and total area of the polygon. This is to make eventually summarizing the data easier.

Nov 28

**Desk**

In GEE file Crop\_Vis, added timeseries of farm of interest and the surrounding farms. Surrounding farms have smaller year to year change in harvest and planting dates.

Nov 29

**Meeting with Avery**

For the maps, use new color scheme (viridis in R), not rainbows

For each parcel, regress planting anomaly from the median in a 50km circle versus percent natural vegetation; place less weight on polygons with a lot of natural vegetation.

For more ‘adjustment’ of estimated planting and harvest, do some sort of quantile-quantile mapping with IMEA’s crop progress reports. The reports are for Mato Grosso and regions within Mato Gross (as defined in <http://www.imea.com.br/upload/publicacoes/arquivos/justificativamapa.pdf>)

The actual crop progress data is in <http://www.imea.com.br/imea-site/relatorios-mercado-detalhe?c=4&s=7> and <http://www.imea.com.br/imea-site/relatorios-mercado-detalhe?c=4&s=8>

**Desk**

In GEE file Crop\_Vis\_v2, created a way to ‘fill in’ colors of the CAR polygons based on planting and harvest date. Also added exporting of histogram information.

R file Crop Timing Vis.R: start doing ggplot and ggridges plot of many years’ worth of histograms.

Nov 30

**Meeting with Sally**

Create a semester plan for upcoming semester

Look up commercial satellites for validation, ask Avery and buy them

**Meeting with Iryna**

On Landsat-MODIS fusion: Eli Melaas (Mark Friedl), paper in RS of Environment

Can do plot over regions rather than individual pixels; use Landsat alone to map soy and then only work with MODIS pixels that have above 80% soy. This can avoid needing a Landsat-MODIS fusion method.

Greg Biging did MODIS-Landsat fusion, see (Zhong et al, 2011) paper in international journal of RS and (Zhong et al, 2012)

Paper in Biogeosciences, 2014 by Klostermann (evaluating remote sensing of…)

Paper in Ecological Applications by Toomey and Friedl, in 2015, title starts with ‘greenness…’

For Sentinel 1, look specifically at Sentinel 1 lit (not RADARSAT, etc) for noise elimination. Don’t use raw SAR, calibrate before using it because it needs region specific processing. Might also use Sentinel to enhance soy mapping instead of just as timeseries analysis.

Look up ENVISAT

DigitalGlobe MapMart: look at high res commercial images. Not very high time resolution

Planet Labs satellites: ask for data for educational purposes, has images from 2015 and later, can work with images using Python API