Jan 2

**Desk**

Worked on R code KDD\_GDD\_Timeseries.R (turned DOY since Jan 1 to DOY since Aug 1 with harvest year/actual year and added a second y axis) and turned it into an R Notebook.

Looked at Planet Imagery in Drive folder Crop Visualization to summarize issues and plan next steps

Jan 3

**Desk**

Made research plan for January, in Jan 4 Meeting powerpoint

Jan 4 - 7

**Desk**

New GEE repo called LandCover to extend Jake’s soy classification; looked at Jake’s original classification code, put it in GEE file Soy Classification Jake mb2\_3, make plan for modifying Jake’s classification and evaluating them

Jan 14

**Desk**

In GEE file LandCover/Soy Classification Variations, set up functionality to export accuracy measurements for various land cover classifiers to choose among

Jan 15

**Desk**

Cloned Alex Carroll’s crop date progress GitHub onto Mac laptop, and cloned R bootcamp’s tutorial GitHub repo onto Mac laptop

**Linda office hours for career counseling**

Do informational interviews six months before graduation, look at it as low stakes; not to get a job but rather to find out about possible companies

Look at the professional development programs and groups they recommended (handouts they gave)

Start networking through friends and LinkedIn

Semester goal: draft a resume

Jan 16

**Desk**

GEE file Soy Classification Masking: try to mask out classified pixels by neighbors (i.e. a pixel in the middle of nowhere is masked)

GEE file Center Pivot Classification to train and classify location of center pivot irrigation

**Avery Skype**

* Land cover classification stuff
  + Jake’s soy\_pts\_agsat is the intersection of agsat and soy\_pts\_1, used it because agsat is very careful with location of points. However since soy\_pts\_agsat is only available in Mato Grosso, use soy\_pts\_1 instead and mask with mapbiomas3
  + Agsat has points outside of Mato Grosso too, we can try to get them later if needed
* Task due Mon Jan 21:
  + CSV of muni ID, CAR poly ID, year, soy\_yield.dta’s variables that are listed in the readme on Dropbox, soy area in centroid from our soy map, soy area outside of centroid and within each donut from our soy map, and same for ecosystem area (mapbiomas forest + savannah, classes 3 and 4)
  + Assume even distribution of legally clearable area in the polygon, if polygon area exceeds donut then include a ratio
  + Put R code to merge datasets in Dropbox
  + Look at Jake’s code ‘Forest Neighborhood’ for getting areas by donuts
  + Make the input shp file into a point file

Jan 17

**Desk**

* Worked in GEE files LandCover/Soy Classification Variations and Soy Classification Masking. Decided on different types of land use map assets and the levels of masking that need to be exported, realized that DC/SC classifications have wide differences among jake’s map and my soy map variations; without a trustworthy validation dataset that’s definitely well geolocated, it’s hard to know which to believe.
* Perhaps new training information, like date of peak EVI, and using multiple soymaps and making a consistency map will help. Planet imagery would definitely help with creating new training points.

**Meeting with Sally**

* Top priority: get $5000 of Planet Lab imagery. Ask Jen Dawson in Davis 6th floor, turn right from 602
  + For creating DC vs SC training points and validating the timing estimate
  + Get images within a soy cropped area; weight the image area downloaded by the total amount of soy in the area
  + Think about how to choose the Planet images to download
  + Think about how to do ‘manual’ classification of Planet: how to automate it?
    - Object oriented: chop images into polygons
    - Ask Iryna about how to do it
* Mapbiomas3 error: check using soy\_pts\_1 in Mato Grosso
* Need to consider the south as well
* If do consistency masking, think about whether the methods are independent. It’s OK if only pick 100% consistency, or produce a set of confidence levels with the land use map
* Test if accuracy is sensitive to number of SC vs DC points: see if we have enough data points. If accuracy isn’t sensitive to number of DC training points, then maybe there’s something about SC that’s hard to classify

Jan 18

**Desk**

Begin Avery’s work on Cerrado compliance. Uploaded 2001 and 2014 Cerrado CAR shapefiles, began code in GEE file AveryTasks/Cerrado\_Compliance and on desktop, R code in Avery Work/compliance

Jan 21

**Desk**

Update masking for connectedPixelCount in GEE file Soy Classification Masking. Did very close up maps of count from connectedPixelCount to see how the function works. Still have pixels not lining up but at least the masking is good. Re-create the new connectionMask soymaps.

Jan 23

**Desk**

* Emailed Joe Mascaro about Planet educational license
* Requested quote for satellite images from landinfo.com
* Began center pivot training for Brazil:
  + tested center pivot code already developed for GEE (original code is in file LandCover/Center Pivot Example)
  + read papers on center pivot training

Jan 24

**Desk**

In GEE file LandCover/Center Pivot Classification, worked on masking out the falsely detected center pivots at the edge of Landsat images and water bodies.

**Meeting with Sally**

* Best option for center pivot: (1) for a section of image, identify the best periods of time when there’s maximum difference between rainfed and irrigated pixels (based on stdev of pixels in the image; (2) do edge detection over the selected periods of time for various kernel sizes (perhaps check on the area of center pivot classified pixels over different times and see how it changes); (3) check in Kansas whether setting different kernel sizes makes a huge difference in whether detection is successful. Look for yellow fields for center pivot detection – maybe the difference is more stark in dry season when there’s no natural vegetation
* Don’t go down machine learning route yet because it tends to overfit
* Cop out option: just throw out large areas with high density center pivot
* Look at what’s causing false positive in farms; false positive in natural vegetation is ok because will mask it out with mapbiomas3
* Figure out what I want to do with the timing data and whether false positives vs false negatives are the bigger thing to avoid
* Ask Iryna about center pivot classifications

Jan 28 - 29

**Desk**

In GEE file LandCover/Center Pivot Classification and Center Pivot Classification v2, explored the scale used in reduceToVectors() to get features of center pivot after doing kernels. Visualized what happens at each scale using reproject() function at various scales. 30m is the best scale for making sure that the reduceToVectors() features match up with the original detected kernel peaks. Higher scales produce features ‘out of thin air’ or that are not aligned at all with the original detected kernel peaks.

In GEE file LandCover/Center Pivot Classification v3, start looking at performance at different EVI composite dates, kernel sizes, region of Brazil. Different EVI composite dates didn’t do well; give up on classifying center pivot for now.

Jan 30

**Desk**

Looked at documentation for Planet Labs satellites, chose to use analytic tiles or scenes for Planet Scope and RapidEye (basically don’t use basic and try not to use visual because it’s color corrected); will still need to calibrate with sentinel-2

Create GEE file LandCover/Planet Image Selection to overlay all potential places to get Planet image

In GEE file LandCover/Soy vs Nonsoy Classificaiton, combined SC soy and DC soy into a single category to see how well Jake’s classification map distinguishes soy from nonsoy, based on soy\_pts1\_agsat training points

Jan 31

**Desk**

Looked at accuracy under the case of combining SC and DC soy into a single ‘soy’ category; accuracy increased, but false positives are still high (30 – 40%). Start by looking at planet images in Matopiba survey polygons, where we know it’s soy.

Started GEE file LandCover/Planet Image Selection to select polygons and export them as shp files in order to load them into Planet for image selection. Started PlanetLabs folder on desktop to hold images and shp files.

Started setting up Planet Labs – GEE connection on desktop

**Meeting with Sally**

* Center pivot stuff: later, might try to add in soil moisture and look at ‘out of season greenness’ (but don’t seem to have high contrast between rainfed vs irrigated agri)
* For soy vs nonsoy issue, see what the false positives (i.e. it’s really nonsoy but the classifications calls it soy) are – are they consistently SC or DC, do they have an obvious characteristic? Maybe will just need to include it in the uncertainty of timing estimates
* For choosing Planet images in Mato Grosso, find a really dense area of soy training points to select an area that’s probably soy
* To automate SC vs DC, create timeseries analysis (at pixel or field scale) and do timeseries reasonableness; or do a peak dectection (e.g. find largest EVI dates and see how they’re spaced out)
* Automate tabbing through images in time and visually identify crop timing, SC/DC; ask someone to double check my classification because it’s subjective; ask Iryna’s students to help?