Objective:

Determine the annual Hansen tree cover loss and emissions for 2001-2017 using the same criteria that Brazil uses for its official Amazon emissions statistics in order to compare Brazil’s official national record with GFW’s independent data

Data sources:

* PRODES primary forest annual boundaries from the end of 2006/beginning of 2007 to 2017: <http://www.dpi.inpe.br/prodesdigital/dadosn/mosaicos/2017/PDigital2000_2017_AMZ_gtif.zip>
  + This raster only goes back to deforestation in 2007, so it can only be used to reconstruct primary forest from 2007 to 2017
  + This raster did not align with Hansen pixels. Got this to align with Hansen pixels using gdal\_warp: C:\GIS\GFW\_Climate\Brazil\_emis\_comparison\PRODES>gdalwarp -t\_srs EPSG:4326 -tr 0.00025 0.00025 -tap PDigital2017\_AMZ\_30m.tif PDigital2017\_AMZ\_30m\_warp.tif
  + Verified that this raster aligned with Hansen pixels.
* PRODES primary forest annual boundaries from the end of 2000/beginning of 2001 to 2014: <http://www.dpi.inpe.br/prodesdigital/dadosn/mosaicos/2014/PDigital2000_2014_AMZ_gtif.zip>
  + This raster can be used to reconstruct primary forest from the start of 2001 to the start of 2014
  + This raster did not align with Hansen pixels. Its resolution is also coarser than the 2007-2016 PRODES raster. Got this to align with Hansen pixels using gdal\_warp: C:\GIS\GFW\_Climate\Brazil\_emis\_comparison\PRODES>gdalwarp -t\_srs EPSG:4326 -tr 0.00025 0.00025 -tap Prodes2014\_AMZ\_60m.tif Prodes2014\_AMZ\_60m\_warp.tif
  + Verified that this raster aligned with Hansen pixels.
* Hansen annual tree cover loss: s3://gfw2-data/forest\_change/hansen\_2017/
* Annual fire tree cover loss from MODIS: <ftp://fuoco.geog.umd.edu/MCD64A1/C6>
* Legal Amazon boundary: same as Liz and Mikaela used in their blog post. Zipped shapefile sent via Slack on 8/21/18. Currently at C:\GIS\Multi\_project\Brazil\_legal\_Amazon\.
  + This shapefile is clearly based on a raster but the raster it is from was not aligned with Hansen loss pixels. This caused shifting of the resulting layers during the analysis.
  + Converted this to a raster, aligning it with Hansen loss pixels [following these directions](https://support.esri.com/en/technical-article/000013221) (using the Processing Extent, Output Coordinates, and Raster Analysis environments) in Polygon to Raster conversion.
  + Verified that this raster aligned with Hansen pixels.

Processing methods:

* Clipped Hansen loss 2001-2017 to Brazil’s legal Amazon boundary (from Liz Goldman) so that all results would be constrained by that.
* Generated map of which Hansen loss pixels from 2001 to 2017 had burning that year or the preceding year. This method was developed by Sam Gibbes for the global forest carbon flux model.

1. Downloaded raw hdf files from ftp site to s3
   1. Ran carbon-budget/burn\_date/multi\_thread\_download\_upload\_raw\_hdf.py
   2. Saved raw hdf to s3:// gfw2-data/climate/carbon\_model/other\_emissions\_inputs/burn\_year/raw\_hdf/
2. Stacked raw hdf files by year. This produced a raster for each year (e.g., 2015) for each hdf tile (e.g., h06\_v11) that says the last day in each year that a pixel was burned
   1. Ran carbon-budget/burn\_date/multi\_thread\_stack\_ba\_hv.py
   2. Saved hdf stacked by year to s3://gfw2-data/climate/carbon\_model/other\_emissions\_inputs/burn\_year/burn\_year/
3. Clipped files to Hansen tile size, resolution, alignment, etc. This produced a Hansen-sized tile (e.g., 10N\_100E) for each year (e.g., 2012) with the last date each pixel was burned in that year.
   1. Ran carbon-budget/burn\_date/multi\_thread\_clip\_year\_tiles.py
   2. Saved s3:// gfw2-data/climate/carbon\_model/other\_emissions\_inputs/burn\_year/burn\_year\_10x10\_clip/
4. Flattened burn year tile to get most recent burn year:
   1. Ran carbon-budget/burn\_date/hansen\_burnyear.py
   2. Saved burn year tiles to s3://gfw-files/sam/carbon\_budget/burn\_loss\_year/
5. Initially, I tried putting the tiles into a mosaic dataset in a geodatabase. However, the script kept failing on the Con function for reasons I couldn’t figure out. So I merged all 12 burn year tiles in the study area into a single raster (ugggh!) and that still didn’t work. But when I created a raster attribute table for the merged burn year raster, the script did work on it. So, I used the merged burn year raster (burnyear\_merge\_20181013.tif) for the analysis.

* Created annual primary forest rasters from PRODES rasters for 2007 to 2017 using the PRODES 2017 raster:
  + Legend for PRODES rasters of different years at: <http://www.dpi.inpe.br/prodesdigital/dadosn/mosaicos/class_rgb.txt>.
  + Needed to create a separate primary forest raster for each year to include in the loss/emissions analysis for each year.
  + For Liz Goldman and Mikaela Weiss’s analysis of 2017 loss comparison using the 2017 PRODES raster, primary forests were codes 1, 16, and 24 (FLORESTA, d2017, r2017).
  + This is on the principal that to get primary forest at the start of 2017, you remove deforestation from all previous years, but not from 2017, which hadn’t occurred yet. I applied the same principle to get primary forest at the start of each year from 2007 to 2017. I confirmed this approach with Marcelo Matsumoto (forest GIS analyst in Brazil) during a phone call and subsequent e-mails.
  + More explicitly: to get primary forest for each year from 2007 to 2017, I included code 1 (FLORESTA) and the d and r codes from the year of interest to 2017. For example, primary forest in 2009 was codes 1 and 8-24 (FLORESTA, d2009-d2017, r2010-r2017) and primary forest in 2014 was codes 1, 13-16, and 21-24 (FLORESTA, d2014-d2017, r2014-r2017). I did not include cloud (code 5, nuvem) pixels in primary forest because there was no way of knowing what was there.
  + I implemented this in a function in Brazil\_emis\_comparison/utilities.py, which used the ArcPy Reclassify geoprocessing command to reclassify the 2017 PRODES raster into individual annual primary forest rasters. Maybe I could’ve written the overall analysis to not need to create and save the annual PRODES rasters but this was easier and I wanted to have the annual rasters for my records and potential additional analyses. I used two different reclassification commands (2007-2009 and 2010-2017) because of the addition of the residual deforestation classes for 2010-2017.
  + QC: Checked that 2007, 2012, and 2017 were producing the expected outputs. Made sure to include rasters from before and after 2010 because the PRODES reclassification command was different for those two categories.
* Created annual primary forest rasters from PRODES rasters for 2000 to 2014 using the PRODES 2014 raster:
  + Legend for PRODES rasters at: <http://www.dpi.inpe.br/prodesdigital/dadosn/mosaicos/class_rgb.txt>
  + Needed to create a separate primary forest raster for each year to include in the loss/emissions analysis for each year. Each primary forest map was at the beginning of that year, i.e. the deforestation of that year had not occurred yet.
  + From e-mailing with Marcelo Matsumoto, I learned that codes 21 (desflorestamento) and 17 (residuo) were deforestation in the most recent year (2014). Thus, they I treated them as if they were code d2014, which meant they remained as primary forest for each of the years I was creating (even 2014, because that is primary forest at the beginning of 2014, so d2014 hadn’t occurred yet). I did not include cloud (code 6, nuvem) pixels in primary forest because there was no way of knowing what was there.
  + Because the legend for the raster is all out of order (deforestation from consecutive years don’t have consecutive codes), there wasn’t an easy way to automate the creation of the annual primary forest rasters from this file, unlike for the 2007-2017 rasters. So I created each primary forest raster manually in ArcMap using the Reclassify tool. The Python snippets are below. In the tool environments, I set the output raster resolution to be the same as the Hansen raster (0.00025 degrees) (Raster Analysis menu).
  + arcpy.gp.Reclassify\_sa("Prodes2014\_AMZ\_60m\_warp\_low\_res.tif", "Value", "0 NODATA;1 NODATA;2 NODATA;3 NODATA;4 NODATA;5 1;6 NODATA;7 NODATA;8 NODATA;9 NODATA;10 NODATA;11 NODATA;12 NODATA;13 NODATA;14 NODATA;15 NODATA;16 NODATA;17 1;18 NODATA;19 NODATA;20 NODATA;21 1;22 NODATA;23 NODATA;24 NODATA", "C:/GIS/GFW\_Climate/Brazil\_emis\_comparison/PRODES/PRODES\_primary\_forest\_2014\_early\_raster\_NoData\_v4.tif", "DATA")
  + arcpy.gp.Reclassify\_sa("Prodes2014\_AMZ\_60m\_warp\_low\_res.tif", "Value", "0 NODATA;1 NODATA;2 NODATA;3 NODATA;4 NODATA;5 1;6 NODATA;7 NODATA;8 NODATA;9 NODATA;10 NODATA;11 NODATA;12 NODATA;13 1;14 NODATA;15 NODATA;16 NODATA;17 1;18 NODATA;19 NODATA;20 NODATA;21 1;22 NODATA;23 NODATA;24 NODATA", "C:/GIS/GFW\_Climate/Brazil\_emis\_comparison/PRODES/PRODES\_primary\_forest\_2013\_early\_raster.tif", "DATA")
  + arcpy.gp.Reclassify\_sa("Prodes2014\_AMZ\_60m\_warp\_low\_res.tif", "Value", "0 NODATA;1 NODATA;2 NODATA;3 NODATA;4 NODATA;5 1;6 NODATA;7 NODATA;8 NODATA;9 NODATA;10 NODATA;11 NODATA;12 NODATA;13 1;14 NODATA;15 1;16 NODATA;17 1;18 NODATA;19 NODATA;20 NODATA;21 1;22 NODATA;23 NODATA;24 NODATA", "C:/GIS/GFW\_Climate/Brazil\_emis\_comparison/PRODES/PRODES\_primary\_forest\_2012\_early\_raster.tif", "DATA")
  + arcpy.gp.Reclassify\_sa("Prodes2014\_AMZ\_60m\_warp\_low\_res.tif", "Value", "0 NODATA;1 NODATA;2 NODATA;3 NODATA;4 NODATA;5 1;6 NODATA;7 NODATA;8 NODATA;9 NODATA;10 NODATA;11 NODATA;12 NODATA;13 1;14 NODATA;15 1;16 NODATA;17 1;18 NODATA;19 NODATA;20 NODATA;21 1;22 1;23 NODATA;24 NODATA", "C:/GIS/GFW\_Climate/Brazil\_emis\_comparison/PRODES/PRODES\_primary\_forest\_2011\_early\_raster.tif", "DATA")
  + arcpy.gp.Reclassify\_sa("Prodes2014\_AMZ\_60m\_warp\_low\_res.tif", "Value", "0 NODATA;1 NODATA;2 NODATA;3 NODATA;4 NODATA;5 1;6 NODATA;7 NODATA;8 NODATA;9 NODATA;10 NODATA;11 NODATA;12 NODATA;13 1;14 NODATA;15 1;16 1;17 1;18 NODATA;19 NODATA;20 NODATA;21 1;22 1;23 NODATA;24 NODATA", "C:/GIS/GFW\_Climate/Brazil\_emis\_comparison/PRODES/PRODES\_primary\_forest\_2010\_early\_raster.tif", "DATA")
  + arcpy.gp.Reclassify\_sa("Prodes2014\_AMZ\_60m\_warp\_low\_res.tif", "Value", "0 NODATA;1 NODATA;2 NODATA;3 NODATA;4 NODATA;5 1;6 NODATA;7 NODATA;8 1;9 NODATA;10 NODATA;11 NODATA;12 NODATA;13 1;14 NODATA;15 1;16 1;17 1;18 NODATA;19 NODATA;20 NODATA;21 1;22 1;23 NODATA;24 NODATA", "C:/GIS/GFW\_Climate/Brazil\_emis\_comparison/PRODES/PRODES\_primary\_forest\_2009\_early\_raster.tif", "DATA")
  + arcpy.gp.Reclassify\_sa("Prodes2014\_AMZ\_60m\_warp\_low\_res.tif", "Value", "0 NODATA;1 NODATA;2 NODATA;3 1;4 NODATA;5 1;6 NODATA;7 NODATA;8 1;9 NODATA;10 NODATA;11 NODATA;12 NODATA;13 1;14 NODATA;15 1;16 1;17 1;18 NODATA;19 NODATA;20 NODATA;21 1;22 1;23 NODATA;24 NODATA", "C:/GIS/GFW\_Climate/Brazil\_emis\_comparison/PRODES/PRODES\_primary\_forest\_2008\_early\_raster.tif", "DATA")
  + arcpy.gp.Reclassify\_sa("Prodes2014\_AMZ\_60m\_warp\_low\_res.tif", "Value", "0 NODATA;1 1;2 NODATA;3 1;4 NODATA;5 1;6 NODATA;7 NODATA;8 1;9 NODATA;10 NODATA;11 NODATA;12 NODATA;13 1;14 NODATA;15 1;16 1;17 1;18 NODATA;19 NODATA;20 NODATA;21 1;22 1;23 NODATA;24 NODATA", "C:/GIS/GFW\_Climate/Brazil\_emis\_comparison/PRODES/PRODES\_primary\_forest\_2007\_early\_raster.tif", "DATA")
  + arcpy.gp.Reclassify\_sa("Prodes2014\_AMZ\_60m\_warp\_low\_res.tif", "Value", "0 NODATA;1 1;2 NODATA;3 1;4 1;5 1;6 NODATA;7 NODATA;8 1;9 NODATA;10 NODATA;11 NODATA;12 NODATA;13 1;14 NODATA;15 1;16 1;17 1;18 NODATA;19 NODATA;20 NODATA;21 1;22 1;23 NODATA;24 NODATA", "C:/GIS/GFW\_Climate/Brazil\_emis\_comparison/PRODES/PRODES\_primary\_forest\_2006\_early\_raster.tif", "DATA")
  + arcpy.gp.Reclassify\_sa("Prodes2014\_AMZ\_60m\_warp\_low\_res.tif", "Value", "0 NODATA;1 1;2 NODATA;3 1;4 1;5 1;6 NODATA;7 NODATA;8 1;9 NODATA;10 NODATA;11 NODATA;12 NODATA;13 1;14 NODATA;15 1;16 1;17 1;18 NODATA;19 NODATA;20 NODATA;21 1;22 1;23 1;24 NODATA", "C:/GIS/GFW\_Climate/Brazil\_emis\_comparison/PRODES/PRODES\_primary\_forest\_2005\_early\_raster.tif", "DATA")
  + arcpy.gp.Reclassify\_sa("Prodes2014\_AMZ\_60m\_warp\_low\_res.tif", "Value", "0 NODATA;1 1;2 NODATA;3 1;4 1;5 1;6 NODATA;7 NODATA;8 1;9 NODATA;10 NODATA;11 NODATA;12 NODATA;13 1;14 NODATA;15 1;16 1;17 1;18 NODATA;19 1;20 NODATA;21 1;22 1;23 1;24 NODATA", "C:/GIS/GFW\_Climate/Brazil\_emis\_comparison/PRODES/PRODES\_primary\_forest\_2004\_early\_raster.tif", "DATA")
  + arcpy.gp.Reclassify\_sa("Prodes2014\_AMZ\_60m\_warp\_low\_res.tif", "Value", "0 NODATA;1 1;2 NODATA;3 1;4 1;5 1;6 NODATA;7 NODATA;8 1;9 NODATA;10 NODATA;11 NODATA;12 NODATA;13 1;14 1;15 1;16 1;17 1;18 NODATA;19 1;20 NODATA;21 1;22 1;23 1;24 NODATA", "C:/GIS/GFW\_Climate/Brazil\_emis\_comparison/PRODES/PRODES\_primary\_forest\_2003\_early\_raster.tif", "DATA")
  + arcpy.gp.Reclassify\_sa("Prodes2014\_AMZ\_60m\_warp\_low\_res.tif", "Value", "0 NODATA;1 1;2 NODATA;3 1;4 1;5 1;6 NODATA;7 NODATA;8 1;9 NODATA;10 NODATA;11 NODATA;12 1;13 1;14 1;15 1;16 1;17 1;18 NODATA;19 1;20 NODATA;21 1;22 1;23 1;24 NODATA", "C:/GIS/GFW\_Climate/Brazil\_emis\_comparison/PRODES/PRODES\_primary\_forest\_2002\_early\_raster.tif", "DATA")
  + arcpy.gp.Reclassify\_sa("Prodes2014\_AMZ\_60m\_warp\_low\_res.tif", "Value", "0 NODATA;1 1;2 NODATA;3 1;4 1;5 1;6 NODATA;7 NODATA;8 1;9 NODATA;10 NODATA;11 1;12 1;13 1;14 1;15 1;16 1;17 1;18 NODATA;19 1;20 NODATA;21 1;22 1;23 1;24 NODATA", "C:/GIS/GFW\_Climate/Brazil\_emis\_comparison/PRODES/PRODES\_primary\_forest\_2001\_early\_raster.tif", "DATA")
  + arcpy.gp.Reclassify\_sa("Prodes2014\_AMZ\_60m\_warp\_low\_res.tif", "Value", "0 NODATA;1 1;2 NODATA;3 1;4 1;5 1;6 NODATA;7 NODATA;8 1;9 NODATA;10 1;11 1;12 1;13 1;14 1;15 1;16 1;17 1;18 NODATA;19 1;20 NODATA;21 1;22 1;23 1;24 NODATA", "C:/GIS/GFW\_Climate/Brazil\_emis\_comparison/PRODES/PRODES\_primary\_forest\_2000\_early\_raster.tif", "DATA")
* Wrote code that applies the four PRODES emissions exclusion criteria that Liz and Mikaela used for each year to annual Hansen loss: 1) within the legal Amazon, 2) within PRODES primary forest for that year, 4) no fire that year or the preceding year, and 4) loss cluster was greater than 6.25 ha. The first was implemented by clipping Hansen loss to the legal Amazon before doing anything else. The latter three were implemented in the analysis script.
  + Ran this with Brazil\_emis\_comparison/Brazil\_GFW\_emissions\_comparison.py on my laptop.
  + The final output for this is a shapefile for each year with the Hansen loss that met these four criteria. This shapefile can then have its emissions calculated for comparison with Brazil’s official emissions statistics.
  + The order in which I applied the exclusions to Hansen loss was somewhat arbitrary, though I started with loss only in the legal Amazon so that all the following statistics were about the legal Amazon.
  + Note on the “no fire” exclusion: Hansen loss pixels were excluded if they had fire that year or the preceding year (except for loss 2001 pixels, which were only excluded if they had loss in 2001 since there is no record for fire areas in 2000). The reason for excluding loss with fires the year before (not just that year) is that the loss could have occurred the year before it was recorded and therefore been the same year as the fire; that is, fire from the year before could have caused the loss observed the following year. I developed this rule upon advice from Mikaela Weiss. When examining the output rasters, I noted that almost none of the raw loss pixels for a given year were on burn pixels from any year besides that year or the preceding year. In other words, among the loss pixels for a given year that were on burn pixels, they were almost always on burn pixels from that year or the preceding year.
  + The script creates an intermediate shapefile for each exclusion step so that emissions can be calculated after each exclusion is applied. This will allow me to see how much each exclusion affected the overall emissions (e.g., did excluding fire areas contribute a lot relative to the >6.25 ha exclusion?)
  + The way I wrote the script, it needs to be run separately for the two PRODES inputs (early: 2001-2014, and late: 2007-2017). To switch between these, I commented out different blocks of code and changed the years of the for loop. It’s ungainly but I didn’t feel like coming up with anything else. This takes a few hours to run for each year, so for the 25 years (between the two PRODES inputs) it’s a lot of run time.
  + QC: Checked the processing of each intermediate tif for loss years 2001, 2007, and 2012. Checked that the outputs for 2007 based on the two PRODES inputs were different, as expected (since they were using different PRODES inputs).
* Ran prep\_for\_tsv\_creation.py on the shapefiles to create a new field that had the file name in it. That will be used for Hadoop.
* Copied all the shapefiles and tifs to s3.
* Created a m4.16xlarge spot machine and ran 1b\_Summary-AOIs-to-TSV/convert-AOI-to-tsv.py from <https://github.com/wri/gfw-annual-loss-processing>. This converts all the shapefiles into tsvs for input into Hadoop without intersecting them with administrative boundaries (GADM) since that’s not relevant for this project. This took just a few minutes.
* Created a large spark Hadoop cluster and ran Hadoop on the tsvs for all years at once but it got very slow at the end of stage 2.0 (incremented 1 step every 30 minutes or less) and so I stopped the run and split the tsvs into four sets (e.g., 2001-2005), which I ran separately. These still ran slowly near the end of stage 2.0 but they finished at least. Charlie theorized that Hadoop ran slowly at the end of stage 2.0 was because it was processing all the really complex shapes that hadn’t been done yet. For the year subsets, this left a few complex shapes but when I ran all years together, this left lots of really complex shapes that dragged out the entire process.
* For Hadoop on each set of years, I followed the instructions on <https://github.com/wri/gfw-annual-loss-processing>: python annual\_update.py --analysis-type loss --points-folder s3://gfw2-data/alerts-tsv/loss\_2017/ --output-folder s3://gfw-files/dgibbs/GFW\_Climate/Brazil\_emis\_comparison/full\_model\_201810/Hadoop\_output\_20190104/raw/2006\_09/ --polygons-folder s3://gfw-files/dgibbs/GFW\_Climate/Brazil\_emis\_comparison/full\_model\_201810/tsvs\_for\_Hadoop\_20190104/2006\_09/ --iterate-by None It took a few hours to run each set of years.
* Combined the Hadoop outputs from all sets of years into a single csv and post-processed them using <https://github.com/wri/gfw-annual-loss-processing/tree/master/2_Cumulate-Results-and-Create-API-Datasets>
* Converted loss area from hectares to million hectares and emissions from Mg biomass to Mt CO2 (divide by 100000, multiply by 0.5, multiply by 44/12). This produced the final loss area and emissions for each year under the area exclusion criteria using the old and new PRODES data.
* Compared the loss area in the legal Amazon without any other exclusion criteria against the loss that Liz Goldman reported in [this blog post](https://blog.globalforestwatch.org/data/technical-blog-comparing-gfws-2017-tree-cover-loss-data-to-official-estimates-in-brazil). As expected, loss areas were the same (given that they are Hansen loss in the same legal Amazon polygon): 3.47 million ha. To check that emissions were correct, I compared this analysis’s emissions in the legal Amazon without any other exclusions to emissions for all of Brazil reported on GFW Climate country pages (30% tcd). As expected, the emissions from this analysis were lower than the emissions for all of Brazil each year. I also checked emissions from this analysis against the emissions from the Amazonia area of interest on the GFW Climate Brazil country page (30% tcd). Values were very similar but not the same, which is to be expected given that “legal Amazon” is different from the “Amazonia” biome. Thus, I feel confident that my legal Amazon, exclusion-free loss and emissions values are consistent with standard GFW values.
* I also compared my 2017 loss area values with each exclusion criteria to the 2017 loss area values reported in Liz’s [blog post](https://blog.globalforestwatch.org/data/technical-blog-comparing-gfws-2017-tree-cover-loss-data-to-official-estimates-in-brazil) to check whether my exclusions were resulting in the same loss area as the exclusions she used. I found that excluding loss in PRODES primary forest (the largest exclusion of loss area) resulted in the same loss area: 1.87 million ha (46% of loss area). However, there are differences after that. Liz next applied the patch size exclusion (ignore loss of areas <6.25 ha), followed by the fire exclusion (ignore loss in areas with fire), resulting in 0.61 million ha of loss in 2017. I applied them in the opposite order: fire exclusion, then patch size exclusion, resulting in 0.79 million ha of loss in 2017. PRODES’ loss that year was 0.71 million ha, so Liz’s estimate and my estimate were about equally far from PRODES’ estimate.
* However, the difference between my estimate and Liz’s is relatively high (0.18 Mha). I have thought of two potential reasons for this, though there could be more: 1) I applied my last two exclusions in a different order from Liz, and 2) I used MODIS burned area (250 m product) instead of Liz using Hansen burned area (which is available only for 2017). To determine what caused this difference, I reran the 2017 analysis using the same order of loss exclusions that Liz did. Whatever differences there are between the output of my reordered analysis and Liz’s, I can attribute to using different burned area data.

###Get the loss area in the legal Amazon

arcpy.Project\_management(in\_dataset="legalAMZ\_loss\_2017", out\_dataset="C:/GIS/GFW\_Climate/Brazil\_emis\_comparison/Output\_exclusion\_order\_changed/legalAMZ\_loss\_2017\_reproj.shp", out\_coor\_system="PROJCS['World\_Eckert\_IV',GEOGCS['GCS\_WGS\_1984',DATUM['D\_WGS\_1984',SPHEROID['WGS\_1984',6378137.0,298.257223563]],PRIMEM['Greenwich',0.0],UNIT['Degree',0.0174532925199433]],PROJECTION['Eckert\_IV'],PARAMETER['False\_Easting',0.0],PARAMETER['False\_Northing',0.0],PARAMETER['Central\_Meridian',0.0],UNIT['Meter',1.0]]", transform\_method="", in\_coor\_system="GEOGCS['GCS\_WGS\_1984',DATUM['D\_WGS\_1984',SPHEROID['WGS\_1984',6378137.0,298.257223563]],PRIMEM['Greenwich',0.0],UNIT['Degree',0.0174532925199433]]", preserve\_shape="NO\_PRESERVE\_SHAPE", max\_deviation="", vertical="NO\_VERTICAL")

arcpy.AddGeometryAttributes\_management(Input\_Features="legalAMZ\_loss\_2017\_reproj", Geometry\_Properties="AREA", Length\_Unit="", Area\_Unit="HECTARES", Coordinate\_System="")

Then in the legalAMZ\_loss\_2017\_ reproj.shp attribute table, sum the area column to get the total area of loss in hectares. This is not exactly the same as what Hadoop calculates but it is close enough.

###Get the area of the PRODES primary forest exclusion loss area in the legal Amazon

arcpy.Project\_management(in\_dataset="legalAMZ\_loss\_2017\_PRODES", out\_dataset="C:/GIS/GFW\_Climate/Brazil\_emis\_comparison/Output\_exclusion\_order\_changed/legalAMZ\_loss\_2017\_PRODES\_reproj.shp", out\_coor\_system="PROJCS['World\_Eckert\_IV',GEOGCS['GCS\_WGS\_1984',DATUM['D\_WGS\_1984',SPHEROID['WGS\_1984',6378137.0,298.257223563]],PRIMEM['Greenwich',0.0],UNIT['Degree',0.0174532925199433]],PROJECTION['Eckert\_IV'],PARAMETER['False\_Easting',0.0],PARAMETER['False\_Northing',0.0],PARAMETER['Central\_Meridian',0.0],UNIT['Meter',1.0]]", transform\_method="", in\_coor\_system="GEOGCS['GCS\_WGS\_1984',DATUM['D\_WGS\_1984',SPHEROID['WGS\_1984',6378137.0,298.257223563]],PRIMEM['Greenwich',0.0],UNIT['Degree',0.0174532925199433]]", preserve\_shape="NO\_PRESERVE\_SHAPE", max\_deviation="", vertical="NO\_VERTICAL")

arcpy.AddGeometryAttributes\_management(Input\_Features="legalAMZ\_loss\_2017\_PRODES\_reproj", Geometry\_Properties="AREA", Length\_Unit="", Area\_Unit="HECTARES", Coordinate\_System="")

Then in the legalAMZ\_loss\_2017\_PRODES\_reproj.shp attribute table, sum the area column to get the total area of loss in hectares. This is not exactly the same as what Hadoop calculates but it is close enough.

###Get the area of the PRODES primary forest, >6.25 ha loss chunks, in the legal Amazon

arcpy.gp.RegionGroup\_sa("legalAMZ\_loss\_2017\_PRODES.tif", "C:/GIS/GFW\_Climate/Brazil\_emis\_comparison/Output\_exclusion\_order\_changed/legalAMZ\_loss\_2017\_PRODES\_neighbor.tif", "EIGHT", "WITHIN", "NO\_LINK", "")

arcpy.RasterToPolygon\_conversion(in\_raster="legalAMZ\_loss\_2017\_PRODES\_neighbor.tif", out\_polygon\_features="C:/GIS/GFW\_Climate/Brazil\_emis\_comparison/Output\_exclusion\_order\_changed/legalAMZ\_loss\_2017\_PRODES\_neighbor.shp", simplify="NO\_SIMPLIFY", raster\_field="Value")

arcpy.Project\_management(in\_dataset="legalAMZ\_loss\_2017\_PRODES\_neighbor", out\_dataset="C:/GIS/GFW\_Climate/Brazil\_emis\_comparison/Output\_exclusion\_order\_changed/legalAMZ\_loss\_2017\_PRODES\_neighbor\_reproj.shp", out\_coor\_system="PROJCS['World\_Eckert\_IV',GEOGCS['GCS\_WGS\_1984',DATUM['D\_WGS\_1984',SPHEROID['WGS\_1984',6378137.0,298.257223563]],PRIMEM['Greenwich',0.0],UNIT['Degree',0.0174532925199433]],PROJECTION['Eckert\_IV'],PARAMETER['False\_Easting',0.0],PARAMETER['False\_Northing',0.0],PARAMETER['Central\_Meridian',0.0],UNIT['Meter',1.0]]", transform\_method="", in\_coor\_system="GEOGCS['GCS\_WGS\_1984',DATUM['D\_WGS\_1984',SPHEROID['WGS\_1984',6378137.0,298.257223563]],PRIMEM['Greenwich',0.0],UNIT['Degree',0.0174532925199433]]", preserve\_shape="NO\_PRESERVE\_SHAPE", max\_deviation="", vertical="NO\_VERTICAL")

arcpy.AddGeometryAttributes\_management(Input\_Features="legalAMZ\_loss\_2017\_PRODES\_neighbor\_reproj", Geometry\_Properties="AREA", Length\_Unit="", Area\_Unit="HECTARES", Coordinate\_System="")

Then selected all the features greater than 6.25 hectares and saved them to legalAMZ\_loss\_2017\_PRODES\_neighbor\_reproj\_areas\_large.shp

###Get the area of the PRODES primary forest, >6.25 ha loss chunks, without fire, in the legal Amazon

arcpy.gp.Reclassify\_sa("burnyear\_merge\_20181013.tif", "Value", "0 NODATA;1 NODATA;2 NODATA;3 NODATA;4 NODATA;5 NODATA;6 NODATA;7 NODATA;8 NODATA;9 NODATA;10 NODATA;11 NODATA;12 NODATA;13 NODATA;14 NODATA;15 NODATA;16 1;17 1", "C:/GIS/GFW\_Climate/Brazil\_emis\_comparison/Output\_exclusion\_order\_changed/burnyear\_2016\_2017.tif", "DATA")

arcpy.RasterToPolygon\_conversion(in\_raster="burnyear\_2016\_2017.tif", out\_polygon\_features="C:/GIS/GFW\_Climate/Brazil\_emis\_comparison/Output\_exclusion\_order\_changed/burnyear\_2016\_2017.shp", simplify="NO\_SIMPLIFY", raster\_field="Value")

arcpy.Project\_management(in\_dataset="burnyear\_2016\_2017", out\_dataset="C:/GIS/GFW\_Climate/Brazil\_emis\_comparison/Output\_exclusion\_order\_changed/burnyear\_2016\_2017\_reproj.shp", out\_coor\_system="PROJCS['World\_Eckert\_IV',GEOGCS['GCS\_WGS\_1984',DATUM['D\_WGS\_1984',SPHEROID['WGS\_1984',6378137.0,298.257223563]],PRIMEM['Greenwich',0.0],UNIT['Degree',0.0174532925199433]],PROJECTION['Eckert\_IV'],PARAMETER['False\_Easting',0.0],PARAMETER['False\_Northing',0.0],PARAMETER['Central\_Meridian',0.0],UNIT['Meter',1.0]]", transform\_method="", in\_coor\_system="GEOGCS['GCS\_WGS\_1984',DATUM['D\_WGS\_1984',SPHEROID['WGS\_1984',6378137.0,298.257223563]],PRIMEM['Greenwich',0.0],UNIT['Degree',0.0174532925199433]]", preserve\_shape="NO\_PRESERVE\_SHAPE", max\_deviation="", vertical="NO\_VERTICAL")

arcpy.Erase\_analysis(in\_features="legalAMZ\_loss\_2017\_PRODES\_neighbor\_reproj\_areas\_large", erase\_features="burnyear\_2016\_2017\_reproj", out\_feature\_class="C:/GIS/GFW\_Climate/Brazil\_emis\_comparison/Output\_exclusion\_order\_changed/legalAMZ\_loss\_2017\_PRODES\_neighbor\_reproj\_areas\_large\_no\_fire.shp", cluster\_tolerance="")

Deleted existing area column in legalAMZ\_loss\_2017\_PRODES\_neighbor\_reproj\_areas\_large\_no\_fire.shp (areas calculated for /legalAMZ\_loss\_2017\_PRODES\_neighbor\_reproj\_areas\_large.shp)

arcpy.AddGeometryAttributes\_management(Input\_Features="legalAMZ\_loss\_2017\_PRODES\_neighbor\_reproj\_areas\_large\_no\_fire", Geometry\_Properties="AREA", Length\_Unit="", Area\_Unit="HECTARES", Coordinate\_System="")

* I figured out the reason for the difference between Liz’s analysis for 2017 and when I tried to repeat her analysis for 2017 in the same order (2/22/19). There were basically two reasons:
* 1. Liz calculated the area of the loss patches before applying any exclusion and never recalculated loss patch areas before applying the patch size exclusion (i.e. after applying the PRODES exclusion).
* Thus, her loss patch sizes were larger than the actual loss patches after loss in non-PRODES primary forest was removed.
* On the other hand, I calculated loss patch areas after excluding loss in non-PRODES primary forest, so when I did the patch size exclusion, it was on the current loss patches, not the original patches.
* 2. Liz separately processed the rasters for fire and non-fire loss for 2017.
* That meant she calculated the size of loss patches separately for fire and non-fire loss.
* That mattered when fire and non-fire loss patches were connected and individually were loss than 6.25 ha but combined were more than 6.25 ha.
* It resulted in Liz excluding some loss patches that were larger than 6.25 ha based on the loss size exclusion.
* The non-fire portions of these loss patches would then be kept through the fire exclusion step because their non-fire loss size was >6.25 ha.
* I, on the other hand, used a single fire/non-fire loss raster for 2017, so if a patch was >6.25 ha, it would be kept through the area exclusion, even if part of it was from fire.
* Takehome: difference between Liz's results and my repeating her methods is pretty much explained. Both methods are defensible and we can proceed with the analysis in the confidence that my methods are sound, though somewhat different from Liz’s.
* Obtained PRODES annual deforestation data from [http://www.obt.inpe.br/prodes/dashboard/prodes-rates.html#](http://www.obt.inpe.br/prodes/dashboard/prodes-rates.html). Each year’s data has its exact value. Copied that into my analysis spreadsheet for graphing against the tree cover loss data I calculated.
* Obtained deforestation emissions data from SEEG v6 (Tasso Azavedo). We considered using SEEG emissions for the NYDF 2018 update but Tasso hadn’t released the 2017 emissions numbers yet, so we didn’t. Tasso sent the link to SEEG v6, released in October 2018. I had to add the emissions for 2001 to 2009 to the pivot table; in the default view, only emissions from 2010 were activated.. Spreadsheet from <https://storage.googleapis.com/mapbiomas/Tabelao_SEEG_VI_GERAL-BR_UF_2018_11_19-SITE.xlsx>, then went to tab “consulta Brasil’ and used the row called Alteracoes de uso do solo”. (The general SEEG website that Tasso sent (<http://seeg.eco.br>) didn’t seem to have working links.) This data is for all of Brazil, though, not just for the legal Amazon.

Resources for understanding Brazil’s deforestation and emissions targets:

Brazil’s 2010 (or 2012?) Nationally Appropriate Mitigation Action: <https://unfccc.int/files/bodies/awg-lca/application/pdf/20120518_brazil_1038.pdf> Shows 564 Mt CO2e reduction from Amazon deforestation by 2020, though I don’t know what the two columns in the table represent. From all sources, a reduction of 36.1%-38.9% of emissions from 2020. Also has a graph of annual Amazon deforestation, though it doesn’t say what the area is (legal Amazon or Amazon biome).

WRI blog article on Brazil’s emissions reduction pledge, 11/2019: <https://www.wri.org/blog/2009/11/brazil-pledges-ambitious-emissions-reductions#fn:1> Goal includes reduction of deforestation in the Amazon by 80% by 2020 relative to 2005, I think. It doesn’t say what the 2005 baseline or 2020 targets are. Notes that there is debate over the economic growth rate predictions that Brazil used for estimating BAU emission (4% vs. 5% vs. 6% annual growth).

Table similar to the Brazil NAMA pdf, but from 11/13/2009: <http://ecen.com/eee75/eee75p/metas_gee_brasil.htm> (Came from searching for the second reference in the reference list in <https://www.wri.org/blog/2009/11/brazil-pledges-ambitious-emissions-reductions>) Again, not clear why there are two mostly identical columns under “Amplituda da reducao 2020”. Shows the 564 reduction below BAU for 2020 but the 1084 Mt BAU emissions do not agree with the table in the press release that John Cannon sent (947.6 Mt). Maybe this is an earlier version of BAU?

Actual text of Decree 7390 from 2010 (12/9/2010), setting emissions goals: <https://www2.camara.leg.br/legin/fed/decret/2010/decreto-7390-9-dezembro-2010-609643-norma-pe.html> (download the Word document, but also saved to my project folder). In table in section 1.1.2 in the annex, shows that deforestation in the legal Amazon (paragraph specifically talks about legal Amazon) under BAU is 1,953,500 ha, C emissions/ha are 132.3 (seems to come from the Second Inventory) (=485.5 Mt CO2/ha), and total emissions are 947,642,850 tons CO2e (or 947 Mt CO2). In addition to providing the original source for the 947 Mt CO2e BAU emissions, it also shows what the corresponding deforestation under BAU is supposed to be, which I hadn’t seen before. Legal Amazon deforestation in 2020 under BAU is the average of deforestation from 1996 to 2005, using PRODES data (according to 1.1.1 in the Annex). The PRODES data do not exactly match the PRODES data I have but they are similar enough; I know that PRODES has been revised over time. I’m not including the PRODES time series in 1.1.1 in my comparison spreadsheet because I believe it to be superseded by the more recent PRODES data… Also, this now clarifies how Brazil calculated its 2016 and 2017 emissions statistics for the Amazon that it used in the press release table (383 Mt CO2 and 337 Mt CO2). They are simply the PRODES loss areas those years (0.7893 Mha and 0.6947 Mha) multiplied by the CO2 density (485 Mt CO2/ha). If they simply multiplied PRODES loss by the CO2 density in 2016 and 2017 for the press release, then I assume that’s legitimate to do in preceding years, so this means I can calculate “official” emissions for the legal Amazon back to 2001!

An overview of Brazil’s climate legislation, orders, and authorities: <https://gettingthedealthrough.com/area/42/jurisdiction/6/climate-regulation-brazil/> I don’t think the information is directly relevant but it’s good background (in English).

National Policy on Climate Change declaration, 12/29/2009: <https://www.preventionweb.net/files/12488_BrazilNationalpolicyEN.pdf> (In English) The only specific numbers are that GHG emissions will be reduced by between 36.1 and 38.9% relative to 2020 BAU emissions. Doesn’t say what the BAU emissions are but I think I’ve established that pretty well from other sources.

Brazilian website about Brazil’s proclamation that it achieved its goal: <http://www.observatoriodoclima.eco.br/emissoes-brasil-nao-justificam-otimismo-com-meta-de-2020/> The first paragraph in the “Metodologia de calculo” section says that the deforestation goal was for 80% less deforestation than the average deforestation between 1996 and 2005, resulting in 189 Mt CO2e for the Amazon. I checked the math and that works out: their deforestation goal was 0.39 Mha, which corresponds to 189 Mt CO2 using their CO2 density of 485.5 Mt CO2/ha (from decree 7390/2010). It also says that deforestation in the Amazon is still at 6,957 km^2, which is 78% higher than what the target suggests. These numbers work with my deforestation target numbers: PRODES 2017 is 0.69 Mha (according to their website) and the 2020 target value I have (80% less than the 2006-2015 average) is 0.39 Mha, which is 78% less than the PRODES 2017 value. So, that’s good—it means that my confusion about the 2017 PRODES deforestation being higher than the 2020 target deforestation rate is justified (that is a valid comparison). It also supports my confusion about Brazil’s statement that the 2017 reported emissions being higher than the 80% reduction emissions target means they reached their goal; this article also says that doesn’t make sense. The article also notes that the announcement wasn’t accompanied by an explanation of how they reached that conclusion.

Talked with Daniela from CLUA about Brazil’s declaration. She told me that Brazil’s official GHG tracker is SIRENE: <http://sirene.mctic.gov.br/portal/opencms/paineis/2018/08/24/Emissoes_em_dioxido_de_carbono_equivalente_por_setor.html> SIRENE was started in response to SEEG’s success. Both SIRENE and SEEG report emissions for aboveground biomass only. She thought that the timing of Brazil’s announcement of its success was probably related to its trying to get money from the Green Climate Fund. She did not know why Brazil said it had achieved its emissions reduction target when the reduction was less than the 2010 decree. We looked at SIRENE together and didn’t see any emissions data for the Amazon specifically.