Replication package for Olmstead and Sigman, "Droughts and Economic Activity"

Three directories:

1. Code: code, log files, and main analysis data sets
2. Data: data that was processed in GIS before being read into Stata
3. Output: figures and tables generated by the code

Processed data in "data" subdirectory:

1. Files from "Polygon to Raster" then "Raster to ASCII" in ArcGIS:
   (a) pfaf4_id.asc: based on Hydro1k shapefiles
   (b) admin1.asc: based on GADM geopackage files
2. Attribute tables following "spatial union" in ArcGIS of GRanD locations with areas:
   (a) grand_pfaf4_global.txt: union with Hydro1k shapefile using Pfafstetter level 4 subbasin (="pfaf4" below) areas
   (b) dam_admin1_join.csv: union with GADM geopackage using Administrative Level 1 areas
3. pfaf4char_global.txt: exported attribute table from Hydro1k shapefile
4. pop2000_pfaf4.txt: from ArcMap, summary of Gridded Population of World pop density within pfaf4 subbasin areas
5. pfaf4_country_match.txt: attribute table for spatial union of Hydro1k shapefile with country boundary shapefile
6. /level4_down_basins/ and /level4_up_basins/: Continent-level upstream and downstream subbasins from MS-Basic code that read Pfafstetter coding from Hydro1k

Main analysis data sets in "code" directory:

1. for_reg.dta: main analysis data set for all grid-cell level analyses
2. for_reg_pfaf4.dta: alternative data for subbasin-level analyses in Appendix I
3. for_reg_admin1.dta: alternative data for state/province-level analyses in Appendix J

Code to create the main analysis datasets:

1. **read_lights.R** (runtime: 3,850 sec): read .tif files from NOAA website, aggregate to .05 degrees, and save as ASCII (run before read_lights_drought.do)
2. **read_DSI.R** (runtime: 978 sec): read .hdf files from UMontana website, save as ASCII (run before read_lights_drought.do)
3. **read_lights_drought.do** (runtime: 3,907 sec): read annual lights and DSI ASCII grids into Stata
4. **read_ascii_cru.do** (runtime: 488 sec): read sc-PSDI, temperature, and precipitation data from Climate Research Unit (CRU) website. Monthly .5 degree aggregated to annual.
5. **read_cropland.R** (runtime: 2 sec): read .tif file and output ascii. (run before read_other_raster.do)
6. **read_other_rasters.do** (runtime: 858 sec): read cross-sectional raster data (pfaf4, cropland, irrigation)
7. **read_pfaf4char.do** (runtime: 6 sec): read in and create subbasin level characteristics, including dams and upstream/downstream characteristics
8. **merge_grid.do** (runtime: 181 sec): combine all data and create for_reg.do, the main analysis dataset at the .5 degree grid-cell level
9. **merge_pfaf4.do** (runtime: 71 sec): creates alternative aggregation to Pfafstetter Level 4 subbasin level, used in Appendix I
10. **merge_admin1.do** (runtime: 268 sec): creates alternative aggregation to GADM Administrative Level 1, used in Appendix J

Stata analysis code:

(All start from main analysis dataset for_reg.dta, except reg_nocell.do)

1. **reg_main.do** (runtime: 678 sec): analysis for all tables (except Table 1) in the paper, Figure 1, and appendix tables C and D
2. **reg_mfx.do** (runtime: 70 sec): translate effects into elasticities and GDP changes for Table 1 in paper and Appendix F
3. **reg_appendix.do** (runtime: 333 sec): additional appendix tables and figures, including those in Appendices A, B, E, G, H, K, and L
4. **reg_nocell.do** (runtime: 9 sec): results with alternative aggregation levels (pfaf4 and admin1) in Appendices I and J (uses for_reg_pfaf4.dta and for_reg_admin1.dta)

Notes on running code:

- Stata code was run in Stata Now Version 19 and R code in R version 4.4.3.

- Directory addresses are relative to “code” directory. “cd” to code directory before running.
- Run all “read_xx” code before “merge_xx” code
- Additional Stata package “ras2dta” needed for reading in ASCII grid files

Data sources:

1. Nighttime lights: https://www.ngdc.noaa.gov/eog/data/web_data/v4composites/
2. Remote sensed drought severity index (DSI): https://www.umt.edu/numerical-terradynamic-simulation-group/project/modis/dsi.php
3. Climate Research Unit Self-correcting Palmer DSI (sc-PDSI): https://crudata.uea.ac.uk/cru/data/drought/
4. Temperature and precipitation (CRU): https://data.ceda.ac.uk/badc/cru/data/cru_ts/cru_ts_4.08/data
5. Dams: Global Reservoirs and Dams (GRanD) v1: https://www.earthdata.nasa.gov/data/catalog/sedac-ciesin-sedac-grandv1-dams-1.01
6. Subbasin shapefile: Hydro1K: https://cmr.earthdata.nasa.gov/search/concepts/C1220567910-USGS_LTA.html
7. Groundwater: World Bank, 2025. A Global Dataset of Aquifer Typologies and Groundwater Resources. https://documents1.worldbank.org/curated/-en/099090825042039644/pdf/P178601-17404731-ebfc-4a21-9a85-3d5c421c211c.pdf
8. Country grid: GPW v4 national identifier grid: https://dataverse.harvard.edu/-dataset.xhtml?persistentId=doi:10.7910/DVN/UTN7K2
9. Flares: https://eogdata.mines.edu/global_flare_data
10. Urban extents: https://sedac.ciesin.columbia.edu/gpw/app/?code=ECU&file=grumpv1&data=urextent&type=ascii&resolut=30
11. Cropland: https://data.earthdata.nasa.gov/nasa-earth/human-dimensions/sedac-root/downloads/data/aglands/aglands-croplands-2000/gl-croplands-geotif.zip
12. Irrigation: FAO Aquastat https://www.fao.org/aquastat/en/geospatial-information/global-maps-irrigated-areas/latest-version/
13. Subnational administrative data: GADM https://gadm.org/download_world36.html
14. Population: Gridded Population of the World v4, https://www.earthdata.nasa.gov/data/catalog/sedac-ciesin-sedac-gpwv4-popdens-r11-4.11

List of output files (in "output" directory) with sources:

1. **reg_main.do** creates
   Figure1.emf
   table2_GW_dsi.rtf
   table3_dam_droughtcat.rtf
   table4_dam_hetero_dsi.rtf
   table5_upstream_dams_dsi.rtf
   table6_upstream_drought.rtf
   table7_landuse_irrig.rtf
   table8_country.rtf
   tableC1_dsi.rtf
   tableC2_pdsi.rtf
   tableD1_GW_pdsi.rtf
   tableD2_dam_droughtcat.rtf
   tableD3_dam_hetero_pdsi.rtf
   tableD4_upstream_dams_pdsi.rtf
2. **reg_mfx.do** creates
   table1.docx
   tableF1_mfx.docx
3. **reg_appendix.do** creates
   FigureA1.pdf
   FigureA2_binscatter.emf
   FigureE1_lags.emf
   FigureG1_growth.emf
   FigureH1_climate_controls_dsidsi.pdf
   FigureH2_compare_temp_precip.pdf
   tableB1_sumstat.rtf
   tableG1_dams_growth.rtf
   tableK1_country_heterogeneity.rtf
   tableL1_IVmeans.rtf
   tableL2_iv_dsi.rtf
   tableL3_iv_pdsi.rtf
4. **reg_nocell.do** creates:
   FigureI1_pfaf4.emf
   FigureJ1_admin1.emf
   tableI1_dam_pfaf4.rtf
   tableJ1_dam_admin1.rtf