# AJAE MS 20231 Replication Package README

The code in this replication package downloads the raw data and constructs the final data sets in R for analysis in Stata. Some of the analysis is performed in R. There are 12 R files and X Stata files which generate the 5 figures and 4 tables in the paper. The replicator should expect the code to run for about 10 to 16 days depending on download speed. If the replicator would rather skip downloading all or some of the raw data, all of the intermediary data files directly used to generate to final data for analysis are also include in the replication package.

## Data Availability and Provenance Statements

### Statement about Rights

I certify that the author(s) of the manuscript have legitimate access to and permission to use the data used in this manuscript.

I certify that the author(s) of the manuscript have documented permission to redistribute/publish the data contained within this replication package. Appropriate permission are documented in the [LICENSE.txt](https://social-science-data-editors.github.io/template_README/LICENSE.txt) file.

The data is licensed under a Creative Commons/CC-BY-NC license. See LICENSE.txt for details.

### Summary of Availability

All data **are or were** publicly available. Some data that was originally downloaded in 2019 has since been updated at the source. In these few cases, a copy of the raw data used in the analysis is provided in the replication folder archive. In other cases where data is downloaded by the programs from sources that may revise this data in the future, a copy of the original data used in this article’s analysis is provided.

### Details on each Raw Data Source

U.S. state and county boundary shapefiles for 2018 were downloaded from the U.S. Census Bureau at <https://www2.census.gov/geo/tiger/TIGER2018/>. A copy of these shape files is provided as part of this archive. The data are in the public domain.

Datafiles: tl\_2018\_us\_state.shp and tl\_2018\_us\_county.shp in the data\_raw/US\_States directory.

PRFIP grid IDs for the rainfall index product within each county in CA were queried from the USDA RMA PRF support tool at <https://prodwebnlb.rma.usda.gov/apps/prf>. Specifically, we queried the site’s web api at <https://prodwebnlb.rma.usda.gov/apps/PrfWebApi/PrfExternalStates/GetSubCountiesByCountyAndState>. PRFIP grid geometry was queried from the sites’ internal map server at <https://prodwebnlb.rma.usda.gov/apps/prf/proxy/proxy.ashx?https://pdgis.rm.usda.net/arcgis/rest/services/ITM/PRF_Grid/MapServer/0/>. The data are in the public domain, but prospective users should limit query load to protect the public server resource. Code to download, clean, and merge the grid data is provided in code/01-get-county-state-grid-geometry.R within the archive.

PRFIP rainfall index values by interval and grid in CA from 1948 to 2019 were also queried from the USDA RMA PRF support tool web api at <https://prodwebnlb.rma.usda.gov/apps/PrfWebApi/PrfExternalIndexes/GetIndexValues>. The data are in the public domain, but prospective users should limit query load to protect the public server resource. Code to download, clean, and merge the index value data is provided in code/02-get-rainfail-index-values.R within the archive.

Daily NDVI measurements from 1981 to 2019 by year were downloaded from the NOAA CDR at <https://www.ncei.noaa.gov/data/avhrr-land-normalized-difference-vegetation-index/access/>. Additional documentation is available at <https://www.ncei.noaa.gov/access/metadata/landing-page/bin/iso?id=gov.noaa.ncdc:C00813>. The data are in the public domain. Code to download daily NDVI measurements in CA is provided in code/03-get-ndvi-noaa-cdr.R within the archive. The code to extract mean daily NDVI and compute bimonthly interval and March, April, May interval index values by PRFIP grid is also provided in code/04-compute-ndvi-index-values-by-grid.R within the archive. Note that NOAA may occasionally revise some of their recent observations, usually those from within the past 3 or 5 years.

2018 CA vegetation cover data and map including WHR types were obtained from the CAL FIRE FRAPS FVEG map server at <https://egis.fire.ca.gov/arcgis/rest/services/FRAP/fveg>. The data are in the public domain, but the 2018 data and map may no longer be available for download. A copy of the 2018 raster file is provided as part of this archive. Code to download the raster attribute table and extract the most common vegetation cover in each PRFIP grid is provided in code/05-get-cal\_fraps\_veg\_category.R within the archive.

Datafile: data\_raw/calfire\_landuse.tif

Elevation data was obtain from the USGS 3DEP program of The National Map (TNM) suite of products and services. See <https://www.usgs.gov/core-science-systems/ngp/3dep/about-3dep-products-services?qt-science_support_page_related_con=0#qt-science_support_page_related_con> for more info about this service. The authors used the TNM Downloader (v2.0) web app at <https://apps.nationalmap.gov/downloader/#/> to identify cells/blocks within and around California. A list of these cells/blocks and links to download elevation data at 1 arc second resolution was exported from the app and is provided in data\_raw/TNM\_cartExport\_20191231\_220907.txt within the archive. Code to download this data and extract mean elevation for each PRFIP grid is provided in code/06-get-ned-elevation.R within the archive. The data are in the public domain.

PRFIP county bases values and premium rates from 2011 to 2021 for each bimonthly index interval and grid were scraped from the USDA RMA Actuarial Information Browser (AIB) at <https://webapp.rma.usda.gov/apps/actuarialinformationbrowser/>. Code to scrape the AIB by PRFIP grid in CA is provided in code/07-get-AIB-data.R within the archive.

USDA RMA Summary of Business data was downloaded by year from <https://www.rma.usda.gov/en/Information-Tools/Summary-of-Business/State-County-Crop-Summary-of-Business>. Copies of the data in both the State/County/Crop/Coverage Level and Type/Practice/Unit layouts are provided as part of this archive. Code to clean and subset this data for the PRFIP and counties in CA is provided in code/07-get-AIB-data.R within the archive.R within the archive.

Datafiles: data\_raw/USDA\_RMA\_Summary\_of\_Business directory

## Data file list

The following list covers the raw data files which were directly used in our analysis. All of the intermediary R data files are generated by the R programs, but are also included in the replication folder. See the R programs as documentation for these intermediary files. Some processed files are included in both RDS and csv formats, but the list below only documents the RDS files. All csv files use a “.” to represent NA values for importing to Stata.

Note: The replication folder does not provide a codebook, but all variable names and most code was written with an attempt at best practices and should be self-documenting. For raw files and datasets, more documentation should also be available directly from the source.

|  |  |  |  |
| --- | --- | --- | --- |
| Data file | Source | Notes | Provided |
| data\_raw/US\_States | U.S. Census Bureau | County and State Shapefiles | Yes |
| data\_raw/USDA\_RMA\_Summary\_of\_Business | USDA RMA | SCCC and TPU layouts | Yes |
| data\_raw/calfire\_landuse.tif | CALFIRE FRAPS | Raster image | Yes |
| data\_raw/TNM\_cartExport\_20191231\_220907.txt | USGS TNM | Exported txt file | Yes |
|  |  |  |  |
|  |  |  |  |

## Computational requirements

The replication folder archive includes all the R code necessary to download and generate the main datasets for the analysis.

### Software Requirements

* R 4.1.1
  + See the sessionInfo.txt file for the most recent package and dependency versions used for this project. However, note that due to time constrains not all programs were rerun during the most recent session, so you may run into issues with backwards compatibility. This primarily effects the NDVI and elevation programs which take a considerable amount of time to run.
  + the file “0\_setup.R” will install all dependencies (latest version), and should be run once prior to running any and all other programs.
* Stata (code was last run with version XX)
  + TODO (packages?)

### Memory and Runtime Requirements

#### Summary

Approximate time needed to reproduce the analyses on a standard (CURRENT YEAR) desktop machine:

10-16 days

#### Details

The code was last run on a laptop with the following setup:

OS: Windows 10 Home 64-bit (10.0, Build 19043)

Processor: Intel® Core™ i7-8550U CPU @ 1.80GHz (8 CPUs), ~2.0GHz

Memory: 16 GB

## Description of programs/code

* R
  + 00\_setup.R installs and loads the required packages for the project.
  + 01-get-county-state-grid-gemoetry.R downloads and processes all the CA spatial data and geometry using a consistent CRS.
  + 02-get-rainfall-index-values.R downloads and lightly processes historical rainfall measurements and index values used in the PRFIP for all grids in CA.
  + 03-get-ndvi-noaa-cdr.R downloads downloads and process daily rasters with NDVI measurements from NOAA.
  + 04-compute-ndvi-index-values-by-grid.R computes bimonthly NDVI index values using NOAA CDR measurements for all grids in CA. Also generates expected forage index values using NDVI observed March, April, and May NDVI measurements.
  + 05-get-cal\_fraps\_veg\_category.R processes the CALFIRE FRAPS 2018 vegetation map and categorizes each PRFIP grid in CA by rangeland vegetation
  + 06-get-ned-elevation.R downloads and processes USGS elevation data, identifying which PRFIP grids in CA are above or below 1524 meters.
  + 07-get-AIB-data.R scrapes the USDA RMA Actuarial Information Browser for historical county base values and premium rates for all PRFIP grid-county combinations in CA.
  + 08-get-summary\_of\_business\_data.R processes and combines the yearly USDA RMA Summary of Business files into a single RDS file for each layout
  + 09-compile-main-datasets.R compiles data downloaded and processed in all previous R programs to generate the final datasets to be analyzed in Stata or generate some tables and figures in R.
  + 10-generate\_tables.R generates a version of Table 1A in R.
  + 11-generate-maps-figures.R generates Figures 1, 2, and 3 in R.

### 

## Instructions to Replicators

This project uses both R and Stata. Data is download and processed in R and the majority of the analysis is done in Stata. The R programs should be executed before the Stata programs if you want to replicate the project from scratch.

R

* Programs should be executed individually in numerical order starting with 00-setup.R. However, not all programs need to be executed unless you want to download and process all the raw and intermediary data files.
* Sincesome programs can take a long time to run, we have also provided the outputs from every program. If, for example, you just want to quickly replicate generating the final datasets, then you just need to execute programs 00-setup.R and 09-compile-main-datasets.R in that order.
* Since the replicator has the option to omit executing some of the intermediary programs, we do not provide a single script that runs all R programs in sequence. We also do not recommend this approach because the individual programs can take a lot of time to execute and we haven’t implemented sophistacted error handling, e.g. the download programs may fail if you lose internet connection part way.
* Each program includes a command that clears objects from memory after they complete.
* Due to recent (as of 2020) updates to some of the R spatial packages (sf, raster, etc.) that use the PROJ library, you may get multiple warnings or other error messages about NA or empty CRS when running 03-get-ndvi-noaa-cdr.R to download the NDVI raster files. We’ve attempted to address this backwards compatibility problem by manually assigning a the original CRS to each file, but the messages will persist.

Stata

## List of tables and programs

The provided code reproduces:

|  |  |  |  |
| --- | --- | --- | --- |
| Figure/Table # | Program | Line Number | Output file |
| Table 1A | 10-generate\_tables.R | 26 | None, table was typed out by hand |
| Figure 1 | 11-generate-maps-figures.R | 35 | Figures/prf\_grid\_rangelands\_cbv.pdf |
| Figure 2 | 11-generate-maps-figures.R | 68 | Figures/rangeland\_RI\_NDVI\_elev\_leq\_1524\_density\_2011\_2018\_RI\_NOAA\_CDR.pdf and Figures/rangeland\_RI\_NDVI\_elev\_leq\_1524\_density\_2011\_2018\_NDVI\_NOAA\_CDR.pdf |
| Figure 3 | 11-generate-maps-figures.R | 144 | Figures/prf\_grid\_optimal\_combos.pdf |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

## References